Aalto University
School of Business

Google Search Queries and Their Impact on Mutual Fund Performance and Flow

Finance

Master's thesis

Anh Dao

2016

Department of Finance
Aalto University
School of Business

Aalto University School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Anh Tuan Dao

Title of thesis Google search queries and their impact on mutual fund performance and flow

Degree Master of Science in Economics and Business Administration

Degree programme Finance

Thesis advisor(s) Peter Nyberg

Year of approval 2016 **Number of pages** 62 **Language** English

Abstract

This study presents a novel way to interpret the effects of investor attention on mutual fund flow and performance, by using Google Trends. Google Trends is an internet service that shows the aggregated search volumes on specific keywords. Based on its nature as being representative of investors' search frequency, Google Trends acts as a direct measure for investor attention. The paper then seeks to discover a relationship between a change in search interests and a change in mutual fund flow and performance.

The data used in this study come mainly from two sources. The mutual fund data is collected from CRSP database, with important characteristics such as monthly total net assets, monthly returns, age and expense ratio. The search data is collected from Google Trends' database using a web crawling program, with the chosen mutual funds' tickers as keywords. The final sample includes 235 mutual funds in the U.S in the period 2006 – 2015.

The study shows a significant negative link between a change in search interests and sample fund's short-term performance. The result is opposite for flow, which displays a positive correlation with investors' search volumes, although the data is not statistically significant. Furthermore, the study also reveals search interests would generate a bigger impact on flows into smaller funds as well as well-performing funds, but again the result is not resolute due to statistical insignificance.

Keywords mutual fund flow, mutual fund performance, investor attention, search volume index, internet search

Contents

1. Introduction 3
1.1. Motivation and background of the study 3
1.1.1. On Google Trends 3
1.1.2. On mutual funds 6
1.2. Research questions and hypotheses 7
1.3. Contribution 8
1.4. Overview of the main findings 9
1.5. Limitations of the study 10
1.6. Structure of the paper 10
2. Literature review 11
2.1. SVI and its relevance in finance 11
2.1.1. SVI literature in general 11
2.1.2. SVI literature in the field of finance 13
2.2. The determinants of fund flow 15
2.2.1. Market elements as determinants of fund flow 15
2.2.2. Investor attention as a determinant of fund flow 17
3. Data and sample 20
3.1. Mutual fund data 20
3.2. SVI data 21
4. Methodology 24
4.1. Required parameters 24
4.2. Regression methods and hypothesis testing 25
5. Empirical findings 28
5.1. Sample descriptive statistics 28
5.2. General findings from data characteristics 29
5.3. Cross-correlation matrix 33
5.4. Regression results 35
5.4.1. SVI and fund performance 35
5.4.2. SVI and fund flow 37
5.4.3. The influence of fund size, performance and age on SVI and fund flow 43

6. Conclusion 50
6.1. Summary of general findings 50
6.2. Suggestions for future researches 51
References 52
Appendix 55
Appendix 1: The code for the Google Trends web crawling program 55
Appendix 2: The list of chosen mutual funds 55

List of Figures and Tables

Figure 1: Search Engines' market shares - March 2016. Source: NetMarketshare.com.........................4
Figure 2: Search Engines' historical market shares 5
Figure 3: iPhone SVI data from Google Trends (2004-2016)..6
Figure 4: BlackRock Equity Dividend Fund's SVI data, weekly, in the period 2006 - 2015 22
Figure 5: The percentage of funds in the sample with SVI over zero, 2006 - 2015 30
Figure 6: Co-movement of SVI and fund flow in 2007..31

Table 1: Variable Definitions 21
Table 2: Descriptive statistics of the sample data...29
Table 3: Summary of methods used for testing hypotheses 25
Table 4: Descriptive statistics of total net assets, monthly flow and monthly return variables..........32
Table 5: Cross-correlation matrix of the variables 34
Table 6: SVI-Return regression results 35
Table 7: SVI-Flow regression results 38
Table 8: SVI-Flow regression results, on a yearly basis 41
Table 9: SVI-flow regression results with interaction between lagged TNA and lagged ASVI 45
Table 10: SVI-flow regression results with interaction between lagged TNA and lagged ASVI, on a yearly basis
....46
Table 11: SVI-flow regression results with interaction between lagged return and lagged ASVI......47
Table 12: SVI-flow regression results with interaction between lagged return and lagged ASVI, on a yearly basis
....49

1. Introduction

1.1. Motivation and background of the study

Information has always been one of, if not the, most important elements in financial investment. Information is the source of investor attention, a vital feature of financial markets, yet we still do not have a direct empirical proxy for it. Indirect proxies exist, such as abnormal returns (Barber & Odean, 2008), trade volume (Barber & Odean, 2008) or news headlines (Barber & Odean, 2008; Yuan, 2008). One problem with these measures is that they all assume if a stock experienced an extreme change in returns, volumes or was mentioned heavily in news, an investor would have paid attention to these events. This assumption is flawed since such thing cannot be guaranteed with the surplus amount of information in this day and age. With the rise of the internet, there comes the advantage of measuring interest data in real-time, based on search queries. Such advance has begun to contribute to many kinds of studies from different disciplines and areas, one of which is the field of financial investment. Da, et al. (2011) were the pioneers in suggesting a direct measure of investor attention: Google search queries, with the assumption that if an investor searches for a stock on the internet, she expresses an interest in that stock. It is a groundbreaking paper and also confirms the predictive power of internet searches and data.

1.1.1. On Google Trends

Google is an American technology company with a wide portfolio of Internet-related services and products. Some examples of their offerings are search engine (Google Search), online advertisements (Google AdWords), cloud services (Google Drive) and software (Android). Being one of the world's most valuable public company, Google holds itself to the mission of organizing the world's information and make it universally accessible and useful (Google[1], 2016). In finance research, one can find themselves using some Google products such as Google Scholar or Google Finance. Although there are

[1] https://www.google.com/intl/en/about/

many search engines with similar functionalities, Google's dominating market share (Figure 1 and Figure 2) makes it an obvious choice for the purpose of this study.

Google Trends (www.google.com/trends) is a specialized web facility that shows how often a specific term has been searched, relative to the total search volume across different world regions and languages. From the data of their search engine Google Search, Google aggregates the search frequency data, represented by Search Volume Index (SVI). In other words, SVI is the number of searches for a specific term, scaled by the time series' average of the predefined time period. Google Trends' SVI data is available publicly for any search term with sufficient search volume on a weekly and monthly basis. Trends' data can be categorized with a number of different criteria: country of origin (results of 'football' in the U.S. is different from 'football' in the U.K), an industry ('Tesla' in automobile versus 'Tesla' in history) and so on. Trends' data is available from 2004 to the previous week.

Figure 1: Search Engines' market shares - March 2016.

The graph shows 8 popular search engines' market share in the U.S. market. Source: NetMarketshare.com

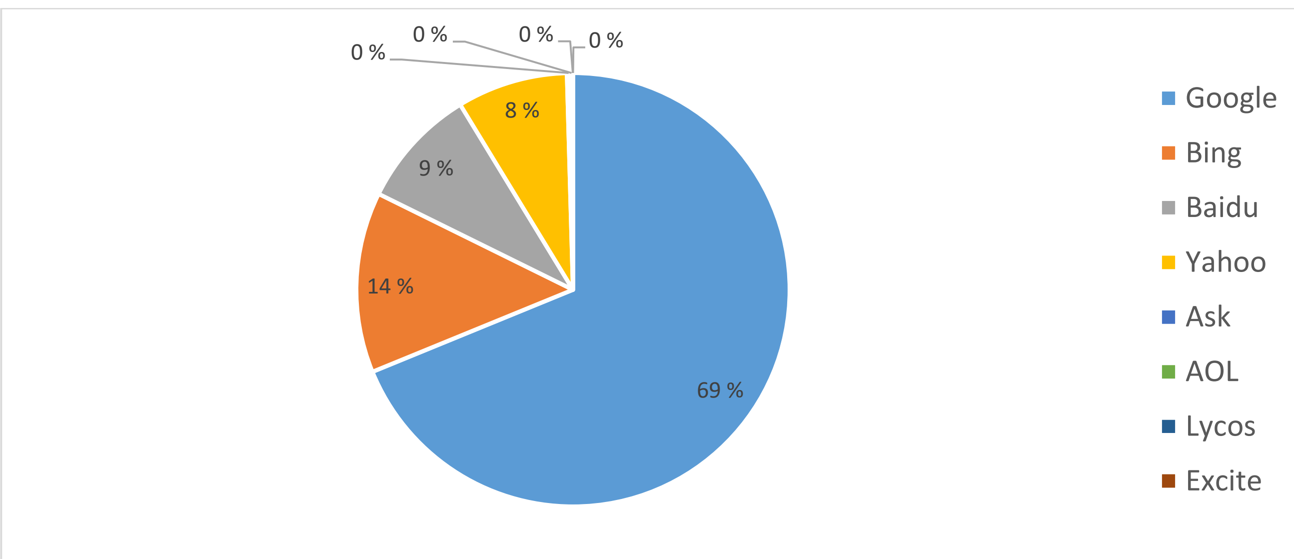
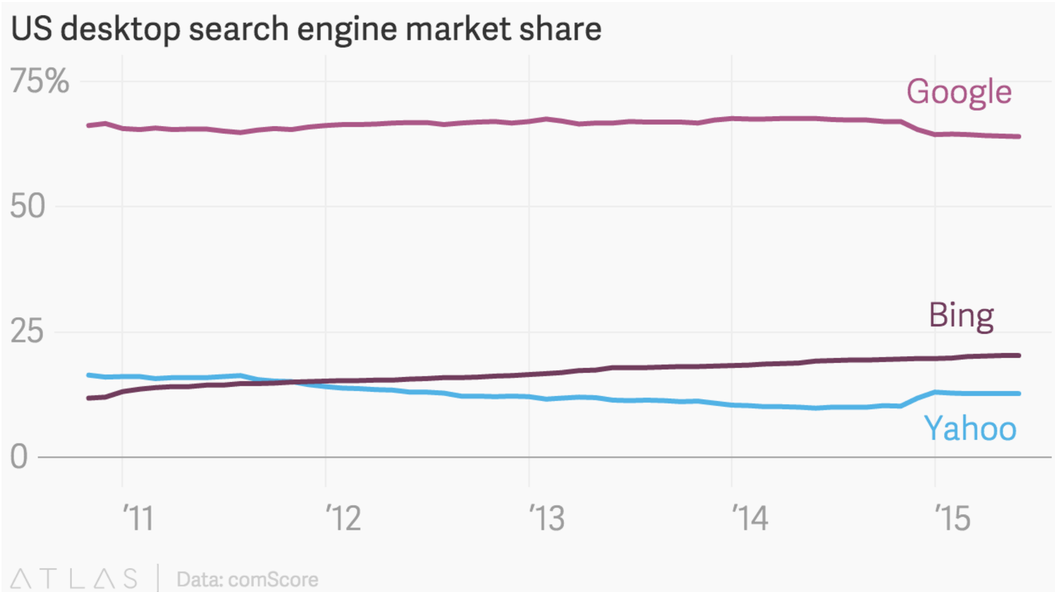


Figure 2: Search Engines' historical market shares.

The chart shows the historical development of three most popular search engines' market share in the U.S market

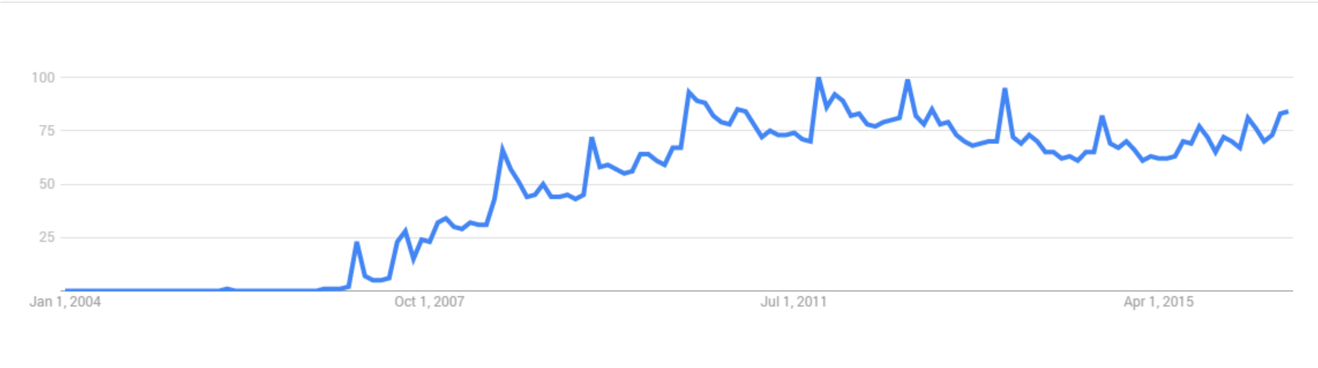


To better understand how the algorithms behind Google Trends work, we can look at the example in Figure 3 below. Figure 3 shows the SVI data for the search term 'iPhone' in the category of 'smartphone' from all over the world. Every time the term is entered into Google Search, it is registered as a search volume entry. Google aggregates these entries to come up with a total, then scales it relative to the highest search volume (equaling 100 SVI) during the time period. In Figure 3, we can see the highest search volume for 'iPhone' occurred in October 2011 and September 2012, the announcement time for iPhone 4S and iPhone 5, respectively.

It is clear from the chart that search interest in iPhone started in 2007 – the year the first iPhone was introduced. In the following years, interest peaked mostly during September and October, the time of the year when Apple announced its new iPhone lineup. Most recently there was an announcement of a new iPhone in March 2016 – an unusual time – and SVI perfectly captures interests then, with a notable increase in the month.

Figure 3: iPhone SVI data from Google Trends (2004-2016).

The chart shows the relative development of the search interests of the term "iPhone" from 2004 to 2016. The highest point implies a Search Volume Index of 100, meaning the highest volume the term has been ever searched for. Source: Google Trends



1.1.2. On mutual funds

Mutual funds now account for a majority of household investments and savings (ICI Study, 2006). Over the past years there have been numerous studies on the causes for mutual fund flow variations. Previous studies mostly focus on the relation between fund flow and market return. We now know that flow into equity funds has a positive correlation with concurrent and subsequent market returns (Warther, 1995), and that daily flow moves positively with concurrent and previous day's market return, but not with lagged flow (Edelen & Warner, 2001). There are also studies on different elements in relation with fund flows, such as Morningstar ratings, whose upgrades will result in a significant increase in fund flows (Guercio & Tkac, 2008), or market volatility which has a negative correlation with a shock in fund flow (Cao et al, 2007). In general, it is an ongoing debate to discover the determinants of fund flows. However, most of these determinants act as a vessel for information – the sole resource upon which investors base their investment decisions. As a result, there rises the need for a better and more versatile measure for information and attention, especially in relation to mutual fund flows.

1.2. Research questions and hypotheses

Since Google Trends' reliability and popularity started to rise, there have been numerous studies looking to discover its predictive power. On their study about predictability of search trends, Shimshoni, et al. (2009) find that over half of the most popular Google search queries had a predictive power in 12-month ahead forecast. Ginsberg, et al. (2009) manage to track influenza-like illness in a population and accurately estimate the current level of influenza activity in many regions of the U.S, thanks to the data on search queries. In the academia of other fields such as psychology or culture, SVI has also been widely utilized as a new perspective for research. However, within finance, the exploration of SVI has been relatively limited. Da, et al. (2011) was the pioneer in using SVI as a proxy for investor's attention and has since then sparked different follow-up studies on its predictive power in finance. This paper aims to follow the footstep of the significant study by Da, et al.

In this paper, I plan to combine the idea of SVI by Da, et al. (2011) with mutual funds, instead of stocks. Specifically, I want to study the effects of investor attention, measured by search queries, on respective fund flow. To my knowledge, this is the first study to make a connection between investor attention and mutual fund flow in such ways. The primary aim of this paper is to further understand if the predictive power of search queries data also applies for the mutual fund case, and also hopefully add to the existing literature on the reasons behind fluctuations in mutual fund flow.

With the core objective of validating Google Trends' predictive ability in the case of mutual funds, many questions arise. It is common sense to assume that the more people search for a specific fund on the internet, the more likely they are to make a change in their investment decision with that fund. It could be that they discover some appealing aspects of the fund and decide to put money into the fund; or it could be that they are disappointed with recent news about the fund and withdraw their money. The two first hypotheses of the paper aim to confirm if there is any correlation between a change in SVI and a change in fund flow and performance. These hypotheses closely follow the similar hypotheses presented in Da, et al. (2011).

H1: A change in a fund's SVI affects the changes in the fund's returns.

H2: A change in a fund's SVI affects the changes in the fund flow.

If there existed such a relationship, one could raise the question: would different characteristics of the funds cause an impact on SVI changes' effects? Indro, et al. (1999) determined that fund size affects the mutual fund performance, since funds must attain a minimum size to achieve sufficient returns to justify for their costs of acquiring and trading on information. As a result, it is reasonable to assume that fund size has a role in the correlation between SVI changes and fund flow variations. This is to say, a fund with bigger net asset usually has more prevalent information, resulting in a possibly higher search volume. Thus, the third hypothesis tests if this holds true.

H3: Fund size influences the impact of search interests on fund flow

A similar argument could apply for funds' past performance. Performance is one of the top drivers for mutual fund flows along fund size and past flows (Patel, et al., 1994), so naturally it can be expected that it will influence the impact of SVI changes on fund flows.

H4: Past performance influences the impact of search interests on fund flow

All of the hypotheses are studied using regression analysis with significance tests. The methodology is further explained in Section 4 and the findings are presented in Section 5.

1.3. Contribution

This is one of, if not the, first studies attempting to examine the relation between SVI and mutual fund flow, with the former acting as a proxy for investor attention. The most notable contribution of this paper is to add a new perspective on the research areas revolving around fund flows' determinants. It introduced a new explanatory factor for fluctuations in fund flows, as well as offer new insights into the behavioral aspect of investors' search activities. Furthermore, the paper is also another foray into search volume literature, which has recently been on the rise in the internet age.

1.4. Overview of the main findings

Based on the main sample of 235 mutual funds from the U.S., I discover that SVI has a statistically significant explanatory power for a fund's subsequent performance. Specifically, lagged abnormal SVI is inversely correlated to fund returns: for a 1% increase in search interests, a sample fund's return drops by approximately 0.13%. This is an unexpected finding, since it is against the common notion that rising search interests would lead to a better fund performance. Nevertheless, it means hypothesis H1 can be accepted, confirming that a change in SVI would cause a change in a fund's returns.

Using fund flow as the dependent variable in the regression models, I conduct the analysis to identify the relation between SVI and fund flow. The main result suggests that flow is positively correlated with investors' search volumes, as for every percentage increase in SVI, fund flow is higher by 4.8%. The same regression is also conducted on a yearly basis, and this robustness check confirms the general positive connection between SVI and fund flow. The only exception is the year 2008, where an increase in search interests would lead to a drop in flow. Considering that this is the year of the beginning of the financial distress, this inverse relation can be explained by a surge in search interests for *bad* news, leading to investors' withdrawal of money from funds. However, all of the mentioned interpretation are unfortunately not backed by statistically significant data, so hypothesis H2 cannot be accepted.

An additional analysis is conducted to study the influence of fund size and performance on SVI's impact on fund flow. By interacting the lagged abnormal SVI variable with a fund's total net asset and returns, I can effectively study their influence. The key results show that investors' search interests would generate a bigger impact on flows into smaller and lesser known funds. As for fund performance, for most of the studied time period, better performing funds have their flow more affected by a change in search interests. The exception comes during the years of 2009 – 2010, where worse performing funds' flow are more strongly influenced by the fluctuation in search volumes. Considering that these are the years of the financial distress, investors are likely more prone to negative news and make informed withdrawal from the bad performing funds, causing their flows to be strongly affected. Unfortunately, these results are not statistically significant at any confidence interval, so both hypothesis H2 and H3 cannot be accepted either.

To sum up, out of the four hypotheses presented in the paper, only the first hypothesis about the relation between SVI and fund performance can be accepted. All other hypotheses related to SVI impact on fund flow are not backed by statistically significant results and thus cannot be accepted.

1.5. Limitations of the study

This thesis will likely have many limitations. First of all, data from Google is stored by them and it is impossible to see the logic behind their algorithms when they generate search interests. Second, search terms, which are funds' names, can still contain noises and cause some bias for the results. Third, SVI data is relative to the highest search volumes in each year, and it is impossible to do any analysis with absolute search volumes data. Last but not least, data from Google Trends can only present the relative search volumes of a search term, not its nature as being a positive search or a negative search. This limitation will hamper some of the interpretation of results in the following sections.

1.6. Structure of the paper

The paper is organized in the following manner. The next section, Section 2, *Literature Review*, goes deeper into SVI and its application in the field of finance. It also provides a theoretical background on investor attention and information and how crucial a role they play in the financial market. Finally, an overview of previous studies on fund flows' determinants is presented.

Section 3, *Data*, introduces the unique dataset acquired for this study, as well as explain the data collection and cleaning process. Section 4, *Methodology*, details the specific steps, with clear equations, on the regression analyses performed. Section 5, *Empirical Findings*, presents all the technical results acquired from the analyses. Finally, Section 6, *Conclusion*, wraps up the study and gives suggestions for further studies and researches on the subject.

2. Literature review

This section goes into more details about the literature relevant to the paper. It is divided into two smaller parts. The first part gives more background information about SVI in academic researches and its application in the field of finance. The second part looks into mutual fund flow and its various determinants and drivers, especially investor attention.

2.1. SVI and its relevance in finance

In this part I will discuss in more details the literature surrounding the use of SVI in many fields of researches. The first smaller chapter will summarize how SVI has been used in general studies, and the second one will review how SVI has been implemented in finance researches.

2.1.1. SVI literature in general

Google Trends was launched by Google to the public in May 2006 and it has gone through various modifications and iterations with new features introduced every now and then. In 2008, Google introduced a major breakthrough by integrating the CSV extracting functionality into its Trends product. This development kick started a major wave of new researches and statistical analysis surrounding Search Volume Index (SVI), which were becoming more and more reliable thanks to Google's nurture. The rise of the new trend in studying search volume has been noticeably tremendous, considering the relatively young age since the inception of the service.

On the other hand, the idea of using internet search behaviors as a predictor of common real-world events dates back to an earlier time. One of the most notable studies is the analysis of Web Access logs by Johnson, et al. (2004). It examines the correlation between the frequency of influenza-related articles' access and the influenza data from the Center for Disease Control and Prevention (CDC). The result was statistically strong, but there was inconclusiveness on the timeliness of the data. After the introduction of Google Trends, Ginsberg, et al. (2009) expand on the influenza study and discovered that search data for 45 terms related to influenza can effectively predict the outbreak one or two weeks before CDC's own reports. This study is often regarded as the cornerstone for search volume studies and major evidence for Google Trends' predictive power. Interestingly, Google has implemented Google FluTrends[2], based on Ginsberg et al. (2009), to its own Trends service, once again showing the importance of the study. Since then, further studies in the field of health and medicine have been published. Some examples are the significant correlation between SVI and kidney stone disease in the U.S (Willard & Nguyen, 2013) and the U.S population's interests in skin cancer's strong association with melanoma outcomes (Bloom, et al., 2015).

SVI is also on the rise in other fields of research. Google's Chief Economist Hal Varian ran a study and discovered that simple regression models that include relevant Google Trends variables tend to outperform models without ones by 5% to 20% (Choi & Varian, 2011). A good example is in the field of tourism. As described by Choi and Varian, Google is one popular tool used for travel planning, thus they make the assumption that an increase in travel-location-related queries may lead to a rise in that location's number of travelers. They examine Hong Kong as the arrival location, with visitors from all over the world. The analysis shows a strikingly high R-squared of 73.3%, effectively demonstrating the predictive power of Google Trends.

One notable application of SVI is how it can be used to predict people's behaviors. Goel, et al. (2010) provide concrete evidences by studying search queries related to feature films, video games and music. They found out a significant link between search volume and box-office for the films, revenues of the games and rank of songs on Billboard respectively. At some points SVI can even predict future sales several weeks forward. The effect is strongest in movies and weakest in music. Judge and Hand (2010) confirm a similar occurrence in the UK cinemas. In addition to usual consumer behaviors, SVI has also been used to calculate political movements. Lui, et al. (2011), however, discover that SVI is a poor predictor of voting results in both the 2008 and 2012 U.S elections. The reason is explained as a voter's attention toward a specific candidate does not translate into an interest to vote; it might as well be a sign of following up on negative news about the candidate. This example shows that although SVI can be a strong addition in any forecasting model, its use should be considered thoroughly before implementation.

[2] https://www.google.org/flutrends

Perhaps one of the most helpful application of SVI is how it can be used to effectively predict various economic indicators in real time. Choi and Varian (2011) confirmed that search data can predict home sales, automobile sales as well as unemployment rate in the U.S. The power of Google Trends extends even further than the New World. Askitas and Zimmermann (2009) demonstrate a strong correlation between relevant keyword searches and unemployment rates using monthly German data. Other studies have also shown similarly valid link between job-seeking queries and unemployment payments (Baker & Fradkin, 2011). Another important economic indicator that SVI has been studied with is inflation. A new measure called Google Inflation Search Index (GISI) is employed against 36 different, more traditional survey measures and TIPS spread. It came on top with the lowest forecast error (Guzman, 2011). Commercial real estate forecasting is also hugely benefited from search volume data, since forecasting models with SVI elements strongly outperform the ones without SVI (Dietzel, et al., 2014). Finally, one of the more important findings is related to measuring economic uncertainty with SVI. Dzielinski (2012) proposes a new ex-ante measuring methods using SVI of the word "economy". His underlying reasoning is that a higher level of economic uncertainty increases the demand for more information. This intuition is confirmed when an indicator, with SVI component derived for the US, increases after the subprime crisis, peaks at the fall of Lehman Brothers and drops down since. In general, SVI has proven its flexibility and reliability as one of the core elements in improving existing economic forecasting models.

2.1.2. SVI literature in the field of finance

In finance SVI has been primarily used as a proxy for attention, with Da, et al. (2011) being one of the most notable and also the main influence of this paper. Given that existing measures of attention such as turnover, extreme returns or news headlines are all indirect proxies, they proposed SVI as a direct measure and conducted an SVI analysis with a sample of Russell 3000 stocks from 2004 – 2008. Their first analysis demonstrates that SVI has a certain correlation with existing aforementioned proxies, but it is much better at capturing investor's attention in a timely fashion. This can be explained that investors are likely to pay more attention to stocks, e.g. by searching for them on the Internet, well ahead of any news, headlines or announcements that cause extreme returns and high turnover, which are unlikely to happen without an already existing attention. Their second analysis goes into detail which types of investors SVI are more likely to capture attention. By cross-referencing retail orders that went through a market center such as Madoff Investment Securities, the study shows a strong correlation between Madoff orders and SVI, suggesting that SVI better captures the attention of individual investors. This is a reasonable result, since individual investors are more likely to search about their financial investments in Google while institutional ones do so on professional platforms like Bloomberg. The final analysis was done to test the attention-induced price pressure hypothesis proposed by Barber and Odean (2008). The result is interesting: a rise in SVI for Russell 3000 stocks can predict higher stock prices in the next two weeks and a later reversal within the year. For IPO stocks, SVI also plays a part in the extreme first-day returns and long-term underperformance. This is to say, successful IPOs with high level of SVI changes and higher abnormal returns underperform successful ones with lower SVI changes, since the latter are not subject to price pressure.

The study by Da et al. (2011) paved way for many more studies attempting to better understand the relation between SVI and the stock market. One main theme is using SVI as a proxy of information acquisition around earnings announcements. Drake, et al. (2011) discover that Google search volume soars about two weeks prior to the announcement and peaks at the announcement, with a magnitude as high as other important corporate events (e.g. M&A). Furthermore, when the search volume increases, it can predict pre-announcement price changes, causing a less likely price response when the news is announced. Another study surrounding earnings is by Fricke, et al. (2014), who find out that SVI can predict stock market reaction to earnings surprises, although it does not predict the analyst-based earnings surprises themselves. One notable finding is that SVI reduces post-earnings announcement drift (PEAD) up to 40 days after the announcement. It is explained that SVI improves the information flow towards uninformed investors, thus reducing information asymmetries and increasing market efficiency in general.

Other elements of the stock market have been also studied with SVI. One popular claim is that Google Trends can predict future price returns. Analyzing stocks in the S&P100 index, Challet and Ayed (2014) confirm there is indeed a boost in weekly forecasting performance model, compared to the ones using previous price returns, but the difference is not significant. In addition, the choice of keywords, i.e. what to search with Google, plays a key role in deciding the final result. In another study, Challet and Ayed (2014) come back to the keyword question. Using an industry-grade backtest system, they manage to verify that random finance-related keywords do not convey more exploitable predictive information.

However, by using suitable keywords can yield profitable investing strategies. Bijl, et al. (2015) perform a similar analysis on the stocks from S&P1500 index to see if they can replicate the findings from Challet and Ayed (2014). Their results show a small and positive relationship between daily searches and abnormal returns but a negative relationship between weekly searches and abnormal returns. Another interesting finding is that searches contain more predictive power in the time of crises, and in recent years search volumes have become a better predictor of abnormal returns. Wuoristo (2012) examines the phenomenon in the U.K and finds a strong correlation between an increase in search volume and a change in trading volume as well as short-term price run. Takeda and Wakao (2014) discover a similar pattern for the Japanese financial market, but with a weaker effect on stock returns.

2.2. The determinants of fund flow

What lie behind the volatile changes in mutual fund flows have continued to be a hot theme among mutual fund researches. Many studies have identified a number of different drivers, from internal element such as past performance to external force such as Morningstar ratings. In this part I will go through the main findings of previous researches about the determinants of fund flows.

2.2.1. Market elements as determinants of fund flow

One of the more popular theory is that fund flows are highly sensitive to past performance. In particular, many researchers agree on an asymmetric relationship between mutual fund flows and past performance: Funds with better performance enjoy a large new money inflow, while the ones with poor performance suffer outflow (Ippolito (1992), Gruber (1996), Sirri & Tufano (1998)). Huang, et al. (2007) put this idea to the test, with an incorporation of participation costs into their model. Their results show that fund flows are sensitive to performance and become even more so when the participation costs are not high. Market returns also play a crucial role in the decision-making process of mutual fund investors. Edelen and Warner (2001) discover a concurrent positive correlation between market returns and aggregate flow into U.S equity funds. For example, days with unexpected positive (negative) flows have a statistically significant link to preceding abnormal market returns of 25 (-25) basis points. However, both flow and market returns are both driven by the arrival of new information and thus it is only a possibility that there exists a causal relationship between them. In a subsequent revision of the paper, Edelen and Warner complement their results with an analysis of lead-lag daily flow-return. They conclude that in longer term, fund flows respond to index returns, or the information driving returns, rather than the other way around, with a lag of one day. However, within the trading day, the main relation appears to be returns reacting to flows, or flow-induced trades, indicating a price impact phenomenon.

Another interesting addition to fund flow – fund's performance relationship is the inclusion of tax. All of the findings from the studies in the previous paragraph use a fund's pretax returns as the proxy for its performance. Bergstresser and Poterba (2001) tackle the problem in a slightly different way: seeking any relationship between after-tax returns that taxable investors earn on equity funds and subsequent cash inflows into respective funds. The study argues that the inclusion of personal tax parameters is significant as mutual funds have become a more prominent channel for individual equity investment (Kennickell, et al., 2000). Using a sample of retail equity mutual funds over 1993 – 1999, they first confirm a similar correlation between pretax returns and net inflows, then go on to perform the same analysis using after-tax returns. The main result is that after-tax returns strongly outperform pretax returns in explaining net inflows; also, funds with larger stocks of unrealized capital gains experience smaller net inflows, similar reliant on their past return performance.

Volatility is an essential component of any financial market, so obviously there are attempts to research the correlation between equity mutual fund flows and market volatility. Cao, et al. (2007) set out to examine the dynamic relationship between these two elements and investigate whether market volatility is associated with concurrent and past aggregate flow. Using sample of mutual funds from 1998 to 2003, their primary finding is that daily market volatility is negatively correlated to concurrent and lagged flow. Also, an impulse response analysis shows that a shock in fund flow has a negative impact on market volatility. This is to say, an inflow shock corresponds to lower market volatility, and an outflow shock higher volatility. This effect is strongly significant over a period of ten days after the shock and gradually disappears. Furthermore, daily fund flow is also negatively related to lagged market volatility, giving evidence that fund investors time market volatility to their benefits.

Mutual fund expense also has a strong influence on mutual fund flow. Barber, et al.'s (2003) is one of the more popular studies on this topic. They argue that since there are far more mutual funds that investors can possibly carefully consider, they often make a choice based on some basic factors: low fees are preferable to high fees or past returns are just poor indicators of future returns. Over the time mutual funds become a common household investment, investors have become increasingly aware of mutual fund costs and grown averse to them. Barber, et al. reason that since front-end fees are more straightforward and in-the-face type of fees than operating expenses calculated behind the scenes, investors learn to avoid these upfront costs but not necessarily the operating expenses. Using fund flows data from 1970 – 1999 to perform a cross-section regression of fund flows on front-end fees, the researchers confirm their hypothesis, identifying a significant negative relation between these two elements. On the other hand, they are not able to pinpoint any concrete relation between fund flows and operating expenses, once again confirming their intuition.

Last but not least, an external force like Morningstar ratings can also have a substantial impact on mutual fund flows. Morningstar is the leading information agency in the fund marketplace and offer a one-to-five-star rating system with a monthly update frequency. Since all of their ratings are public, it is safe to say a rating change is a noticeable and apparent event. Thus, it is expected that investors will react in light of the news. Guercio and Tkac (2008) confirm this hypothesis, showing a significant positive abnormal flow after rating upgrades and a completely reverse pattern for rating downgrades. The results typically range from 13-30% of normal fund flow. This finding corroborates the claim of Morningstar's ability to influence fund investors and their investment allocations.

2.2.2. Investor attention as a determinant of fund flow

We have briefly discussed the many different determinants of fund flows, and in this part the focus is investor attention. Attention effects of individual investors are important elements of any financial market, yet most of its multiple facets have not been properly understood. One primary challenge is that attention is a scarce cognitive resource (Kahneman, 1973) and it cannot be expected for investors to take into account every piece of news related to their investments. Even if the news is publicly available, it is not certain that investors have acquired such information. Huberman and Regev (2001) identify a 300% daily return of a drug company's stock, due to a news article published in the New York Times, although that same story has been revealed half a year earlier in Nature magazine. As investors are not likely to acquire investment information from a science magazine, this new information had not been incorporated into their decision-making process before it came into New York Times. This is to say, the assumption that investors have undivided attention to their assets is not at all reliable. It leads to the question: so what would be a better way to depict and capture investor attention?

In the past decades, media coverage had been the primary method for investors to receive information about their holdings as well as other information related to the funds. Barber and Odean (2008) suggested that news is a primary mechanism for attention and that investors buy stocks that captured their attention. As a result, media coverage appropriately acts as a good proxy for investor attention. Sirri and Tufano (1998) discover that investors drive their attention and flock to funds praised by news spotlight. Kaniel, et al. (2007) find out that media coverage has a significant effect on money flows into the funds, specifically an additional 1.2% net flow for a month with a news article. The authors also went into more details, analyzing the different effects of positive and negative news and discovered that both have a significant correlation to eventual fund flows. The former is linked to an average of 1.5% increase in flow while the latter 1% decrease in flow.

Another interesting result from Kaniel, et al. (2007) is that media coverage has different effects on funds with different sizes, ages and past performance. It seems to contribute more to flows into smaller and lesser known funds and performance strongly enhances the media's effects on flows. Additionally, the longer the fund has existed, the weaker the effects from news, which is also a sign of investor learning. All their results are consistent with the hypothesis that investors acquire information through media coverage and that different level of knowledge about the funds result in different investment behaviors.

Solomon, et al. (2014) expand the same concept of Kaniel, et al. (2007) by looking deeper into how media coverage can affect the way investors allocate their holdings in a fund. The stocks in the fund with high past returns tend to attract extra flows from investors, but this effect is only visible when these stocks have been recently featured in the news. In contrast, the stocks with no major feature have no effect on flows. This is because when facing a long list of possible investment options, investors appear to respond only to those with the most available positive information, i.e. most featured on the news. As a result, funds that hold high-visibility winner stocks attract much greater capital flow than their counterparts with less visible holdings. The effect is opposite for funds that hold loser stocks. The more high-profile their loser holdings are, the more severe the attrition of flow is.

On the other hand, Solomon, et al. (2014) also find it doubtful that media coverage actually provides any valuable information for the investors. Although investors react to news-featured stocks in their funds, they do not make the decision in the right direction. That is to say, investors usually do not make a good decision to buy or sell their shares of a fund in light of the news related to that fund, but they always do something. This reaction creates an incentive for fund managers to window dress their portfolios by only holding media-covered winner stocks. However, there is no evidence that this behavior is penalized by the investors.

3. Data and sample

In this section, the data and study sample of the paper are presented. The first subsection discusses the mutual fund data, gathered from CRSP. The second subsection shows how SVI data is collected and cleaned from Google Trends. They also detail the process of eliminating unfit observations and finalizing the sample selection.

3.1. Mutual fund data

To fit the purposes of this analysis, the fund data needs to be in a timeframe where the Internet is the essential way for investors to access information. According to an ICI Study in 2006, nearly 73% of mutual fund shareholders used the Internet for financial information related to their investments, compared to only 7% in 1997. As a result, the main dataset will include U.S. mutual funds in the CRSP mutual fund database that existed in the 2006 – 2015 period.

Using CRSP's data extraction, I collected every possible fund's information that is available in the database. This initial process gathers 73014 mutual funds. One thing to note is that these 73014 funds include different classes of the same fund (1048576 observations). To make the analysis simpler and more straightforward, for each different fund I only include class A if it has more than one classes available. The choice of using class A instead of other classes is due to its typical benefits. Class A shares usually have lower fees and are designed for investors with long-term investments. This fits nicely with the purpose of the analysis, since I aim to study the SVI effects on the fund performance over a long period of time. If a fund only has full data for a single class, I will include that class in the sample and remove the rest with missing data. After this step, in the sample remains 3277 unique mutual funds.

The next step is to collect a number of different parameters of each fund required for the analysis. Using the list of 3277 unique mutual funds, I again take advantage of CRSP and extract the following data for each fund: monthly total net assets, monthly returns, age and annual expense ratio. Table 1 describes these variables in more details. This step generates 199897 fund-month observations. The only interval available in CRSP for mutual funds is monthly data, so the final analysis will be done on a monthly basis.

After acquiring the dataset with the defined parameters, I eliminate the funds that have missing data in any of the parameters. 213 funds are removed from the sample after this step, resulting in a 3064-fund sample. This is the sample ready for the SVI data collection process, detailed in Section 3.2.

Table 1: Variable Definitions

Variables	**Definition**
Total net assets (monthly)	The total value of the fund's assets (securities in its portfolio), less any liabilities.
Return (monthly)	The percentage profit of the fund's portfolio during the month
Age	The time in years the fund has existed
Expense ratio	The annual fee that the fund charges its shareholders. It expresses the percentage of assets deducted each fiscal year for all types of fund expenses.

3.2. SVI data

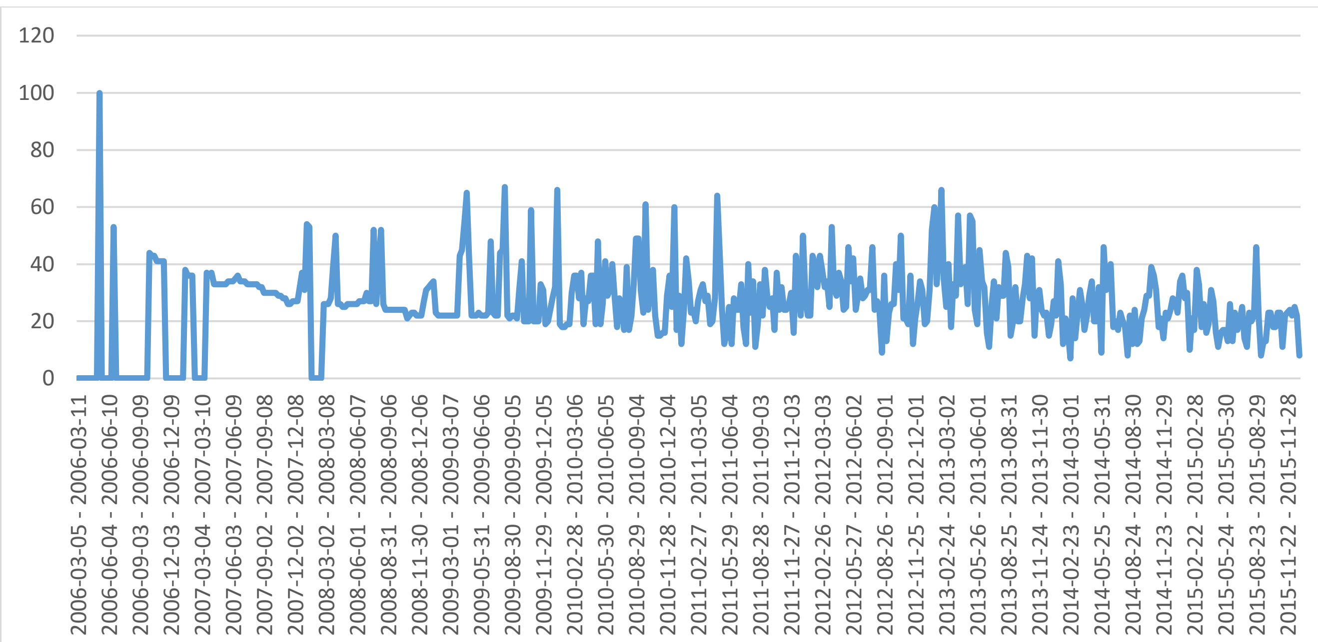
SVI data is provided by Google Trends service, whose data is available from 2004 till the previous week. However, due to its privacy and anonymity policies, Google will only display results if there are sufficient queries for a specific search term. This is to say, no SVI information is available for funds that generate too little or no search volumes. After inputting the term into the search bar, Google Trends will display a graph of SVI over the chosen time period, with the option to download a CSV file. The process is straightforward, but the numerous choices for the search terms are one of the major complication in the SVI data collection.

For each fund, there are usually two unique identifiers for itself: the fund's name and the fund's trade ticker. For example, Vanguard 500 Index Fund, aside from its name, can be pinpointed using its ticker VFINX. The first question arises: which would be a better choice for capturing investor attention for a given fund?

In stocks' cases, Da, et al. (2011) argue that using a company's name as the search term could be troublesome and biased, since investors may search for the name with other intentions than investing.

"Noisy" companies' names such as Apple or Tesla do not help either. However, in mutual funds' cases, these might not be a problem. First of all, funds exist for the sole purpose of attracting investments, so if a person searches for a particular fund, it is highly likely that she is at least interested in investing in the fund. Second, funds often have relatively long and detailed names that describe what the fund's primary strategy is (e.g. Prudential Jennison Small Company Fund). As a result, it is safe to say that one does not accidentally search the exact combination of words with a completely different objective in mind. In short, using a fund's name as a search term does not cause any significant problem that might disrupt the analysis. Figure 4 shows an example with BlackRock Equity Dividend Fund's SVI data, collected from Google Trends using its full name.

Figure 4: BlackRock Equity Dividend Fund's SVI data, weekly, in the period 2006 - 2015



Using the funds' tickers as search terms is another option. Da, et al. (2011) explain that using stocks' tickers in this case are less ambiguous, since searching for "AAPL" likely means an interest in the stock of Apple rather than its products. In short, users who search for tickers are the group of people that perfectly captures the purposes of the analysis. The same logic can apply in mutual funds' case, since every fund also has a unique ticker as its identifier. Furthermore, different from stocks' tickers, mutual funds' tickers comprise more characters in a more unique combination (for example, GuideStone Growth

Allocation Funds - GCOZX) so it is safe to say that they would not be confused with another unintentional meaning.

To collect data on all 3064 mutual funds in the sample first constructed in Section 3.1, I write a simple web crawling program using Python that automatically input the search term into Google Trends' search bar and download the appropriate CSV files of the queried SVI data (see Appendix 1 for the code). I first run this program with the list of sampled funds' names for the period 2006 - 2015. This step generates a total of 1568768 weekly observations for all the funds. Unfortunately, as mentioned before, Google Trends returns empty results when the queries do not generate sufficient search volumes in the chosen time period. Hence, the majority of my queries do not return valid SVI data. Out of 3064 funds queried, only 235 funds (or 7.67%) have a valid SVI data, which translate to 120832 weekly observations. This is a big limitation of the sample construction, but there is simply no way around Google's truncation of the results.

I repeat the same process for the funds' tickers, and have even worse luck with their SVI results. Out of 3064 tickers queried, only a handful of 13 tickers (or 0.4%) return a valid SVI data. This amount is too low for any reliable analysis, so I decide to dismiss the use of tickers in the analysis altogether.

One important thing to note is that SVI data comes on a weekly basis, which is the unchangeable default format from Google Trends for long time periods. Since all of the mutual fund sample's data is monthly, I need to consolidate SVI results from weekly to monthly. This step can be done with a simple arithmetic average of the four weeks making up the month.

$$SVI_m = \frac{SVI_{w1} + SVI_{w2} + SVI_{w3} + SVI_{w4}}{4}$$

in which SVI_m is the calculated SVI for a specific month, and SVI_ws are the four weeks in the month.

In summary, the final sample includes 235 mutual funds. The list of chosen funds is presented in Appendix 2.

4. Methodology

This study, for the most part, replicates the methodology employed by Da, et al. (2011) and adapt it for mutual funds. This section will go through all the hypotheses in this paper and further explain the process to test them.

4.1. Required parameters

There are two important values that need to be derived from the collected data: abnormal SVI and Fund Flow.

Abnormal SVI (ASVI) serves the purpose of identifying the attention peaks in the search volume data. It is calculated as the log of SVI during the month minus the log median SVI during the previous two months. Intuitively, the logarithm of median SVIs captures the "normal" level of attention that is robust to recent jumps or outliers. This is similar to the methods used by Da, et al. (2011), confirmed to be robust to the length of rolling window (1 month, 2 months and so on). ASVI has a considerable advantage that it is not affected by time trends and other less frequent seasonality. Furthermore, a large ASVI clearly indicates a jump in investor attention, thus is really helpful in this study. The formula below depicts how ASVI can be calculated from the SVI data:

$$ASVI_{i,t} = \log SVI_{i,t} - \log Median(SVI_{i,t-1}, \dots, SVI_{i,t-2})$$

in which, $ASVI_{i,\,t}$: abnormal SVI of fund i, in month t

$ASVI_{i,\,t}$: SVI of fund I, in month t

The next value to be derived is fund's net flow, which lies in the center of this study. Based on the formula used by Kaniel, et al. (2007), I calculate net flows into fund i over period t using the equation:

$$Flow_{i,t} = \frac{TNA_{i,t} - [TNA_{i,t-1} * (1 + r_{i,t})]}{TNA_{i,t-1}}$$

where $TNA_{i,\,t}$ is fund i's total net assets at the end of period t and $r_{i,\,t}$ is its return over period t. In this formula, flow is expressed as a percentage, and positive value indicates an increase in money flow into the fund. Average Market Flow (MF) value is also calculated using the average flow for all funds in the sample over period t.

4.2. Regression methods and hypothesis testing

There are four primary hypotheses that are taken into consideration in this paper. Table 2 briefly summarizes their purposes, the methods used for testing and the study from which the hypothesis draw reference.

Table 2: Summary of methods used for testing hypotheses

Main method	**OLS regression / Da, et al (2011)**
H1	The relationship between changes in SVI and changes in fund's return

Main method	**OLS regression / Kaniel, et al. (2007)**
H2	The relationship between changes in SVI and changes in fund flow
H3	Fund size's influence on results
H4	Fund performance's influence on results

The first hypothesis (H1) examines the relation between the changes in a fund's SVI and returns. It can be analyzed using OLS regression, with the fund's returns being the dependent variable and ASVI being one of the independent variables. The complete formula is shown below.

H1: A change in a fund's SVI affects the changes in the fund's returns.

$$R_t = \beta_0 + \beta_1 ASVI_{t-1} + \beta_2 r_{t-1} + \beta_3 \log(TNA_{t-1}) + \beta_4 * small\ dummy * r_{t-1} + \beta_5 * small\ dummy + \varepsilon_t$$

where R_t is the fund's return in month t

$ASVI_{t-1}$ is abnormal SVI in month t-1

r_{t-1} is the fund's return in month t-1 (or lagged returns)

TNA_{t-1} is the total net asset of the fund in month t-1

small dummy takes on the value of zero if the fund's TNA exceeds the median fund's TNA.

This hypothesis is derived from a similar one tested in Da, et al. (2011), in which ASVI is regressed against Russell 3000 Stock Returns. The main idea behind H1 is simply to validate if the statistically significant results found in the case of stocks still apply for funds. The fund's previous returns and total net assets are used as control variables, since they usually play a role in determining next period's returns (Carhart, 1997). The small dummy, used to represent smaller-sized funds in the sample, is added to account for the interaction between size and performance of a fund. This analysis will be performed with R_t and R_{t+1} to identify if search volumes have any effect on both the current month and one month ahead.

The second hypothesis closely follows the formula used by Kaniel, et al. (2007). In the study, he examines the relation between fund flows and news variables, controlled by known determinants of fund flows such as lagged flows, size, performance and fund expense. This paper applies the same methodology, but instead of using news, I utilize SVI as the variable for investor attention in the formula.

H2: A change in a fund's SVI affects the changes the fund flow.

$$Flow_t = \beta_0 + \beta_1 MF_t + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 \log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \varepsilon_t$$

where $Flow_t$ is the fund's flow in month t

MF is the Market Flow (average of all funds' flows) in month t

r_{t-1} is the fund's return in month t-1 (or lagged returns)

TNA_{t-1} is the total net asset of the fund in month t-1

Age is the number of years the fund has existed

ExpenseRatio$_{t-1}$ is the expense ratio of the fund in month t-1

σ_{t-1} is the volatility of the funds' returns in month t-1

ASVI$_{t-1}$ is the abnormal SVI of the fund in month t-1

The third, fourth and fifth hypotheses are still replications of Kaniel, et al. (2007)'s methodologies to study the influence of the fund's characteristics (size, past performance) on the SVI effects on fund flow. Similar to the previous hypotheses, I replace Kaniel, et al.'s media coverage variables with SVI as the variable for investor attention. The other control variables are similar to H2.

H3: Fund size influences the impact of search interests on fund flow

$$Flow_t = \beta_0 + \beta_1 MF + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 \log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \beta_9 [\log(TNA_{t-1}) * ASVI_{t-1}] + \varepsilon_t$$

H4: Past performance influences the impact of search interests on fund flow

$$Flow_t = \beta_0 + \beta_1 MF + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 \log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \beta_9 [\mathrm{r}_{t-1} * ASVI_{t-1}] + \varepsilon_t$$

For H2 to H4, robustness checks are also performed by doing the regression where data is sorted on a yearly basis to see how the SVI effects change over the course of a number of years.

5. Empirical findings

This section presents the main empirical results of the study. The first and second part discuss the general findings from the sample descriptive statistics and an analysis of the sample quality and co-movement of SVI-flow. Additionally, I also perform a univariate analysis of three most important fund characteristics in the sample: total net assets, fund flow and fund return. The third part investigates the correlations between the different parameters used in the overall analysis and try to understand the relation between them. The fourth section presents the regression results and checks the validity of the hypotheses.

5.1. Sample descriptive statistics

The descriptive statistics of all studies variables are presented in Table 3. As the table shows, the monthly returns of the funds are relatively low with a mean of 0.5% for the whole period. TNA figure has a high standard deviation, indicating that the chosen funds are distributed among many different fund sizes, which is a positive thing since the study aims to capture SVI effects for both small and large fund. ASVI data seems to display a similar pattern, as it ranges from an extreme drop in interests (minimum ASVI of -2.916) and a considerable rise in interests (maximum ASVI of 2.921). In addition, the mean ASVI suggests that for the whole period, investors are losing interests on the funds more often. However, flow data is positive, both for each fund and the average market. The mean monthly flow for the whole period is 6.9% while for the market, the number is 16.3%. This is an unexpected result. We have observed that on an average scale, investors are losing interests in their fund investments, while in fact positive flows are prevalent. This observation is seemingly against the common intuition that the more investors search for a fund, the more likely they are to invest in that fund. In the *Empirical Findings* section when we go into a deeper analysis, we will see if this observation still holds statistically.

Table 3: Descriptive statistics of the sample data.
The table shows the relevant statistics of the variables: total net assets (TNA), fund returns, fund's expense ratio, fund's abnormal search volume index (ASVI), fund flow, average market fund flow and fund's age. The results are calculated from a sample of 235 U.S mutual funds in the period 2006 – 2015

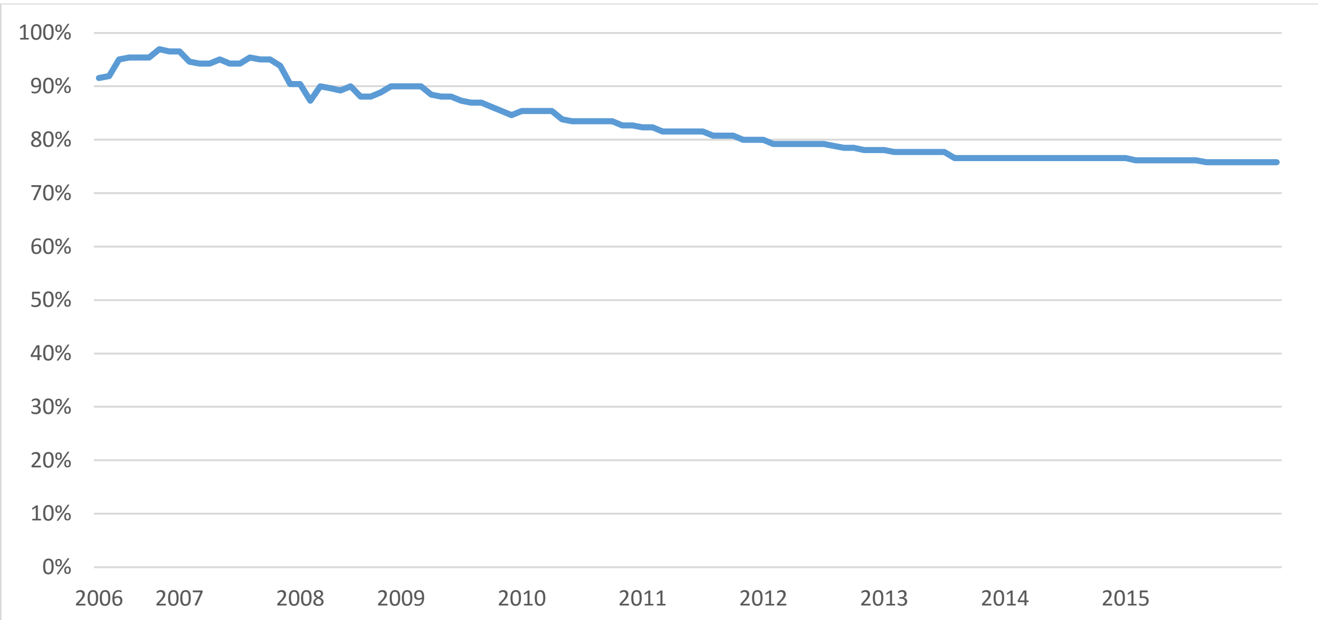
	TNA	Returns	Expense Ratio	ASVI	Flow	Average Market Flow	Age
Mean	6901.383	0.005	0.015	-0.056	0.069	0.163	20.10
Std Deviation	15311.4	0.048	0.008	0.513	60.13	4.219	16.547
Min	-99	-0.4	0	-2.916	-990.974	-8.469	0
Max	143000	0.381	0.027	2.921	9053.955	37.678	86
Observations	25546	25534	25546	25310	25309	24904	25546

5.2. General findings from data characteristics

This study uses a novel dataset when combining funds' data and their corresponding monthly search volumes. As a result, it is important to look at the quality and structure of the dataset. The first thing to look at is the SVI data from Google Trends. As described in previous sections, Google Trends will return a zero result if a search term does not generate enough volumes in the predefined period. Therefore, it is important to know how many funds in our sample generate decent search volume.

Figure 5 shows the percentage of the fund sample with SVI over zero in the period 2006 – 2015. From the chart, it is noticeable that the amount of SVI results over zero is relatively high, always staying above 75%. However, the number of funds with sufficient search volumes shows a downward slope throughout the years. This is most likely due to two factors. First, investors may have reduced their search interests in mutual funds starting from the year 2009 – the time of the financial distress' aftermath. This pattern is shown more clearly in Figure below. Incidentally, net flow into mutual funds experience a similar pattern (Investment Company Institute, 2015), so the implications are that reduced search interests can have a link to net flow. Second, the number of mutual funds in the U.S. after 2008 decreases noticeably before a slow recovery in 2012 (Statista, 2015). This partly explains the downward slope in Figure 5, because there are fewer funds available and funds that used to exist will always generate zero SVI after their demise. Funds closed out in a fairly rapid pace in the period 2008 – 2012 (Statista, 2015), so the chart depicts this trend quite effectively.

Figure 5: The percentage of funds in the sample with SVI over zero, in the period 2006 - 2015



The next general test is to look at the co-movement of SVI and fund flow to find out if the raw data can provide any indication on SVI's predictive power. Figure 6 depicts the movement of average flow and SVI, by aggregating respective data from all companies for each month during 2007 – 2015. SVI overall experiences a downward slope, indicating an evident drop in investors' search interests, especially after the year 2009. Average flow has a more stable development through the years, although it has several extreme drops and spikes. Specifically, the worst drop came in 2009, where average flow is under -$8 million, a drop of -800% compared to the previous year, indicating a gargantuan money withdrawal from funds. Interestingly, SVI and flow seems to follow a similar pattern. In 2011, there were extreme drops in fund flow of nearly -$6 million (approximately -500% compared to the previous year), and SVI also experienced a decrease of approximately -33%. They then both stabilized through the rest of 2011 before a big fall in the beginning of 2012. This similar movement indicates a possible link between flow and SVI: they seem to affect each other and move in the same manner.

Figure 6: Co-movement of SVI and fund flow in 2007.
Average flow is calculated as the average of the sampled 235 mutual fund flows in each studied timeframe. Average SVI is calculated as the average of the sampled 235 mutual fund's SVI in each studied timeframe.

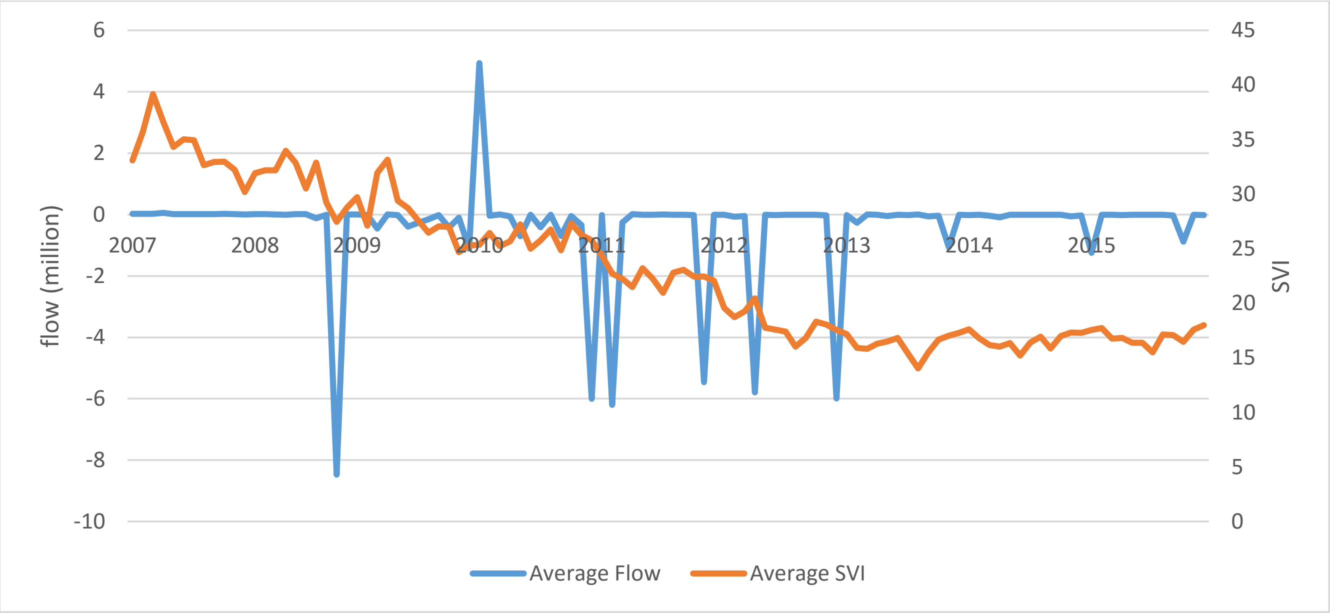


Table 4 shows a more detailed overview of the most important variables for this paper: total net assets, returns and flow. The funds are categorized into high- and low-search volume groups to more clearly show the apparent effects of SVI. The table also provides t-statistics for the difference in the means of the two groups.

First, similar to the previous observation regarding TNA, the funds in the full sample vary greatly in size, ranging from \$24.3 million in the 10^{th} percentile to nearly \$20 billion in the 90^{th} percentile. What's interesting is the figures from the high- and low-search-volume groups. High SVI group has a much lower mean TNA of only \$4.9 billion compared to the \$7.1 billion of the funds with low SVI. The t-statistics for this difference in mean is highly significant at the 1% level, suggesting that somehow smaller funds generate a higher search interests from the investors. I explore this in later sections, where I investigate whether fund size has a role in SVI impact on fund flow.

Second, the figures for monthly flows are in line with common intuition, where the mean flow of the high-search volume group reaches 10.9%, which is relatively higher than the low-search volume group (6.5%). However, this difference in means is not statistically significant. Another noteworthy point is that although these two groups differ strongly in the 10^{th} and 90^{th} percentile, the median fund in both groups receive almost no net flow in the month. This indicates a sign that search interests may not have that big of an impact on monthly flow.

Third, it is shown in the table that funds with lower search interests generally perform better than funds with higher search interests, with a mean return of 0.5% and 0.2% respectively. This is especially noticeable at the higher end of the sample (see 90th percentile), where funds with low search volumes outperform their counterparts with higher search volumes. The t-statistics show that this difference is statistically significant, proving an interesting point: investors tend to search for funds with worse performance. It can however be explained that investors are more prone to searching the *bad* news about the funds rather than the good news. In the later section, I will explore if performance is indeed an influence on SVI impact.

Table 4: Descriptive statistics of total net assets, monthly flow and monthly return variables. The data is analyzed from a sample of 235 U.S mutual funds in the period 2006 – 2015, together with their Search Volume Index (SVI) data collected from Google Trends service.

Variables	**Obs**	**Mean**	**Std Dev**	**10th Perc.**	**Median**	**90th Perc.**
Total Net Assets						
Full sample	25546	6901.383	15311.4	24.3	1041.6	19976.4
High search interests (SVI>50)	2019	4896.674	12786.56	3.5	273.5	12615.7
Low search interests (SVI<50)	23489	7081.321	15506.1	39.6	1136.9	20541.8
t-value for difference in mean		-6.333***				
Monthly Flow						
Full sample	25309	0.069	60.129	-0.032	-0.004	0.067
High search interests (SVI>50)	1953	0.109	3.104	-0.027	0	0.053
Low search interests (SVI<50)	23320	0.065	62.635	-0.033	-0.004	0.034
t-value for difference in mean		0.03				
Total Monthly Return						
Full sample	25534	0.005	0.048	-0.049	0.006	0.056
High search interests (SVI>50)	2017	0.002	0.042	-0.047	0.006	0.043
Low search interests (SVI<50)	23479	0.005	0.048	-0.05	0.005	0.057
t-value for difference in mean		-2.838***				

***significant at the 1% level **significant at the 5% level *significant at the 1% level

5.3. Cross-correlation matrix

The values for the correlation between the variables collected in the sample are presented in Table 5. The variables are on a monthly basis, with no lag in the data. The matrix shows some interesting perspectives on the relations between the variables.

First, the correlation between returns and other variables is relatively quite low, with the most notable correlation with expense ratio of -0.056, followed by average market flow of -0.046. This is in line with the results from Da, et al. (2011), who conclude that price movement is related to multiple factors.

Second, the logarithm of TNA is positively correlated with age, showing a value of 0.43, suggesting that the older funds usually have a much larger size. Moreover, net asset seems to also be related to SVI, with a figure of -0.156. This result indicates that smaller funds generally receive more search interests from the investors. Such observation agrees with the previous analysis in section 5.1 and also the findings from Kaniel, et al. (2007). As for other variables, log(TNA) seems to have a fairly weak correlation with them, with the lowest being flow (-0.008), which is a sign that fund size does not really affect the flow into the fund.

Third, it is noticeable that flow has a weak link to other variables, except for average market flow. The latter observation suggests that flow into individual funds may very well be affected by general market condition. However, other variables cannot really be used to explain flow, including SVI which we are trying to use to measure flow. In the later section a more detailed analysis on this aspect will be presented.

Last but not least, expense ratio seems to have a strong link to both market flow and SVI, with the values staying at 0.218 and 0.211 respectively. It can be explained as a sign of high interests from investors for funds with lower fees. In other words, lower fees attract more general investments (higher market flow) as well as more interests from the investors (higher SVI). This result agrees with the findings by Barber, et al. (2003), who discover that investors are more drawn towards funds with lower upfront fees, hence the more positive flows and interests.

Table 5: Cross-correlation matrix of the variables used in the main analyses. The data is collected and analyzed for a sample of 235 U.S mutual funds in the period 2006 – 2015.

	Return	Standard deviation	log(TNA)	Flow	Market flow	SVI	ASVI	Age	Expense Ratio
Return	1								
Standard deviation	-0.0774	1							
log(TNA)	0.0218	-0.0097	1						
Flow	0.0036	0.0094	-0.0083	1					
Market flow	-0.0461	-0.0904	-0.0408	0.0673	1				
SVI	-0.0312	-0.0291	-0.1553	0.0097	0.096	1			
ASVI	0.0057	-0.0087	0.0206	0.0036	-0.0313	0.2959	1		
Age	0.0104	-0.085	0.4303	-0.0135	-0.0361	-0.0432	0.0397	1	
Expense Ratio	-0.0559	0.0648	-0.0994	0.0142	0.2177	0.2111	-0.0207	-0.092	1

5.4. Regression results

This section outlines the important results from the regression analyses and discuss the possible reasons behind the numbers. The first subsection will revisit hypothesis H1 and discuss, based on the test results, if there is any relation between SVI and fund returns. The second subsection will analyze the most important hypothesis in this paper, H1, about whether search queries have any effect on mutual fund flow. The final subsection will include a different interaction term in each hypothesis to study the influence of fund size, performance and age on the SVI impact on fund flow.

5.4.1. SVI and fund performance

This regression is performed to test the first hypothesis H1. The main results are presented in Table 6.

H1: A change in a fund's SVI affects the changes in the fund's returns.

Table 6 depicts the key results from the regression with the dependent variables on two time periods: fund returns on month t and month t+1. The independent variables used are lagged ASVI, lagged return, lagged TNA, small dummy and the interaction term between small dummy and lagged return.

Based on the regression, it is easy to see that lagged ASVI has a strong link to fund returns, both in the same month and one month ahead. The result is statistically significant at the 5% level. Interestingly, lagged ASVI's coefficient has a negative sign in both time periods (-0.0013 and -0.0014 respectively), indicating an inverse relationship between search interests and fund returns. In other words, the regression suggests that when there is a drop in investor's search volumes related to a specific fund, its return in the same and next month will actually experience a slight increase, and vice versa. This is an unexpected finding and is not in line with the price pressure theory by Barber and Odean (2008), which suggests that an increase in attention leads to higher net buying and temporarily higher returns before stabilization. The analysis also depicts a strong positive correlation with lagged return. Its coefficient is noticeably high and statistically significant at the 1% level. This correlation helps confirming the theory on persistence of mutual fund performance (Carhart, 1997), which is a rational observation, thus improving the validity of the regression result.

Table 6: SVI-Return regression results.

The table shows the statistical relationship between the dependent variable, fund return, with the independent variables lagged Abnormal Search Volume Index (ASVI), lagged return, the logarithm of lagged Total Net Assets (TNA), the dummy variable for lagged return and the dummy variable for small funds. The regression is done based on the equation below, for the sample of 235 U.S mutual funds in the period 2006 - 2015:

$$R_t = \beta_0 + \beta_1 ASVI_{t-1} + \beta_2 r_{t-1} + \beta_3 log(TNA_{t-1}) + \beta_4 * small\ dummy * r_{t-1} + \beta_5 * small\ dummy + \varepsilon_t$$

where R_t is the fund's return in month t

$ASVI_{t-1}$ is abnormal SVI in month t-1

r_{t-1} is the fund's return in month t-1 (or lagged returns)

TNA_{t-1} is the total net asset of the fund in month t-1

small dummy takes on the value of zero if the fund's TNA exceeds the median fund's TNA.

	t		t+1	
Return	Coef.	Std. Err.	Coef.	Std. Err.
Lagged ASVI	-0.0013**	0.0006	-0.0014**	0.0006
Lagged Return	0.1260***	0.0089	0.0037	0.0090
Lagged TNA (log)	-0.0007***	0.0002	-0.0002	0.0002
Small Dummy x Lagged Return	0.0473***	0.0126	-0.0131	0.0127
Small Dummy	-0.0047***	0.0010	-0.0008	0.0010
_cons	0.0114***	0.0017	0.0071***	0.0017
R-squared	0.0239		R-squared	0.0003

***significant at the 1% level **significant at the 5% level *significant at the 1% level

All other independent variables also have a strong explanatory power for fund return, being strongly significant at the 1% level. The regression suggests that fund size has an important role in its continuous performance. The small dummy has a negative sign, indicating that smaller funds seem to experience a lesser return stream of approximately -0.47% than larger funds. However, the interaction term between the small dummy and lagged return stays at 0.0473. Such positive interaction shows that future returns of small funds are more affected by past returns that larger funds. To put it in another way, past returns play a stronger part in determining future returns for small-sized funds, but in general they perform worse than large-sized funds. Interestingly, this finding again is not in line with popular results from previous studies that fund size erodes performance (Chen, et al., 2004). One explanation for this unexpected result could be that the nature of search interests, calculated by SVI, cannot be clearly defined. This is to say, based on search volumes alone, it is impossible to determine whether the searches are about positive or negative news. For previous studies, e.g. Kaniel, et al. (2007), it has been a standard practice to differentiate between positive and negative interests, and they result in positive and negative performance changes respectively. If this theory is assumed to be true, it can be deducted that my sample for search volumes contains more negative news than positive ones. Thus, it is likely that investors tend to search for negative news on the internet. This is a noteworthy occurrence that would need further investigation outside the scope of this paper. On the other hand, the results for month t+1 are not as significant, so there is no concrete conclusion about past returns' effects on returns of one month ahead.

To sum up, based on the analyzed results, H1 hypothesis can be accepted. This means the regression finds a simple link: a change in a fund's search volumes has a recognizable inverse effect on its returns, both on the same month and one month ahead. This result is different from the common notion that search interests increase mutual fund returns, adding a new insight into how investor interests can affect fund performance. Nevertheless, H1 regression analysis confirms there is a connection between SVI and fund performance, giving more grounds on the notion that SVI can affect mutual funds. In the next section I will explore SVI effects on mutual fund flow with a similar regression, followed by a year-by-year analysis to confirm the test's robustness throughout the years.

5.4.2. SVI and fund flow

This section describes the key results from the regression conducted to test hypothesis H2.

H2: A change in a fund's SVI affects the changes the fund flow.

Table 7 depicts the numbers from the regression conducted with fund flow as the dependent variable. The independent variables are average market flow, lagged flow, lagged return, lagged TNA (log), fund age, lagged expense ratio, fund return volatility and lagged ASVI.

Table 7: SVI-Flow regression results

The table shows the statistical relationship between the dependent variable, fund flow, with the independent variables average market flow, lagged flow, lagged return, the logarithm of lagged Total Net Assets (TNA), fund's age, lagged expense ratio, volatility and lagged Abnormal Search Volume Index (ASVI). The regression is done based on the equation below, for the sample of 235 U.S mutual funds in the period 2006 - 2015:

$$Flow_t = \beta_0 + \beta_1 MF_t + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \varepsilon_t$$

where $Flow_t$ is the fund's flow in month t

MF is the Market Flow (average of all funds' flows) in month t

r_{t-1} is the fund's return in month t-1 (or lagged returns)

TNA_{t-1} is the total net asset of the fund in month t-1

Age is the number of years the fund has existed

$ExpenseRatio_{t-1}$ is the expense ratio of the fund in month t-1

σ_{t-1} is the volatility of the funds' returns in month t-1

$ASVI_{t-1}$ is the abnormal SVI of the fund in month t-1

Flow	Coef.	Std. Err.	t	P>t	95% Confidence Interval	
Average market flow	0.0882***	0.0195	4.5200	0.0000	0.0499	0.1265
Lagged flow	0.0000	0.0013	0.0100	0.9950	-0.0025	0.0025
Lagged return	4.7546**	1.6598	2.8600	0.0040	1.5013	8.0079
Lagged TNA (log)	0.0250	0.0344	0.7300	0.4670	-0.0424	0.0924
Age	-0.0032	0.0053	-0.6100	0.5450	-0.0136	0.0072
Lagged expense ratio	1.0491	10.5252	0.1000	0.9210	-19.5809	21.6791
Volatility	0.4899	29.7182	0.0200	0.9870	-57.7595	58.7394
Lagged ASVI	0.0478	0.1542	0.3100	0.7560	-0.2544	0.3501
_cons	-0.1561	0.2874	-0.5400	0.5870	-0.7194	0.4073
R-squared	0.0014					

***significant at the 1% level　　**significant at the 5% level　　*significant at the 1% level

There are a few noteworthy findings based on the results. First, for every percentage increase in average market flow, the sample fund flow rises by 8.82%. The effect is statistically significant at the 1% level, with a fairly low standard error of 2%. It suggests a marketwide phenomenon where the amount of flows into a specific fund has a strong dependence on flows into other funds in the market. Second, lagged flow seems to have no effect at all on the current fund flow, but the result is not significant, so there could not be any definite conclusion of the relation between lagged and current flow. Third, lagged return is highly positively linked to current flow, with a 475% increase in flow for every 1% increase in return. This outcome is statistically significant at the 5% level. The finding is in line with previous findings that mutual fund flows are highly sensitive to past performance.

All other dependent variables do not show any significant explanatory power for mutual fund flow. The lagged TNA log variable results in a positive coefficient, indicating that flows seem to be more focused on larger funds. Specifically, for every log point increase in TNA of a fund, flow into that fund grows by 2.5%. Similar patterns are also found with lagged expense ratio and volatility, so it can be said that money flows more into funds with higher fees and risks. This observation is against previous papers detailed in the Literature Review section, but the numbers shown are not statistically significant, so it is difficult to say if there is any new reliable finding in this regard.

The key variable to test H2 is the lagged SVI variable, which has a slightly positive coefficient. This result suggests that flow is positively correlated with investors' search volumes, as expected in the beginning. Unfortunately, the t-statistics reveal its insignificance, making it impossible to accept H2 that a change in SVI would have an effect on mutual fund flow.

Table 8 performs the same regression in more details, specifically on a yearly basis. The purpose of this analysis is a robustness check as well as to study the independent variables' movement throughout the years.

It is easy to see that average market flow has a continuously positive relation with a sample fund flow, with its coefficients being positive from 2006 – 2012. This is to say, for most of the time, individual fund flows seem to follow the market movement. The best example would be the year 2008. In 2008, a 1% increase in average market flow leads to more than 100% increase in a sample fund flow, also being significant at the 1% level. Similar significant results are also found in the year 2007 and 2010, with the figures being 71.3% and 54.5% respectively. After 2012, the figures start to move into the negative-sign zone, indicating a sample fund flow does not necessarily follow marketwide movement as before.

However, the coefficients are mostly close to zero and they are not as statistically significant either. As a result, from the observations with the average market flow variable, we can identify a visible pattern that investors' money flows closely follow the market condition until 2012, when this effect lessens and individual fund flow does not depend as much on the market movement.

Lagged flow shares a similarly positive correlation with fund flow, with the coefficients in all years being positive or at least close to zero. The most prominent year is 2007, where the regression produces a statistically significant result at the 1% level. This is also the time where past flow has the strongest influence on current flow: for every percentage increase in past flow, it leads to a 6.9% increase in current flow. For other years, the figures have largely hovered around the zero point, indicating that past flow most of the time does not necessarily indicate the movement of current flow. However, the majority of these results are not significant at any level, thus it is not possible to draw any empirical conclusion about the relationship between past and current fund flow.

Lagged return, or past performance, experiences a generally positive correlation with flow over the years. A notable example would be the recent year of 2015, where 1% increase in past performance is linked with a 14.3% additional flow. This result is statistically significant at the 1% level. Other similar years are 2007, 2011, 2013, where past performance also enjoy a positive connection with flows at a few confidence intervals (10%, 5% and 1% respectively). This observation confirms the finding from the previous Table. However, the most interesting numbers would be the huge negative coefficients in the year 2008 and 2010. A 1% drop in return is associated with a huge additional flow of 470% in 2008 and 397% in 2010. Although both of them are not statistically significant, they are still worth some attention. This phenomenon in 2008 could be explained partly with beginning of the financial crisis. During the first half of 2008, the market had not collapsed yet, so mutual funds still enjoyed a healthy stream of fund flow. When the distress arrived at the latter half of the year, fund returns dropped so vehemently and rapidly that investors did not manage to adjust their holdings in the fund accordingly. The year 2008 is the case where overall fund flow is still largely positive but fund returns are decimated mostly at the end of the year. As a consequence, the regression fails to capture such extreme changes, resulting in the final result being an inverse relation between past performance and flow.

Table 8: SVI-Flow regression results, on a yearly basis.

The table shows the statistical relationship between the dependent variable, fund flow, with the independent variables average market flow, lagged flow, lagged return, the logarithm of lagged Total Net Assets (TNA), fund's age, lagged expense ratio, volatility and lagged Abnormal Search Volume Index (ASVI). The regression is done based on the equation below, on a yearly basis for the period 2006 – 2015 for the sample of 235 U.S mutual funds:

$$Flow_t = \beta_0 + \beta_1 MF_t + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \varepsilon_t$$

where $Flow_t$ is the fund's flow in month t
MF is the Market Flow (average of all funds' flows) in month t
r_{t-1} is the fund's return in month t-1 (or lagged returns)
TNA_{t-1} is the total net asset of the fund in month t-1
Age is the number of years the fund has existed
$ExpenseRatio_{t-1}$ is the expense ratio of the fund in month t-1
σ_{t-1} is the volatility of the funds' returns in month t-1
$ASVI_{t-1}$ is the abnormal SVI of the fund in month t-1

Flow	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Market flow	0.0586	0.7128*	1.0445***	0.0655	0.5451	0.0003	-0.0007	-0.0012	-0.0294	-0.0059*
Lagged flow	0.0001	0.0689***	0.0106	0.0084	-0.0006	0.0010***	0.0000	-0.0001	0.0038	0.0001
Lagged return	0.9471	0.2167*	-4.7037	0.4495	-3.9753	0.0995**	0.0393	0.1757***	0.0810	0.1432
Lagged TNA (log)	-0.6300***	-0.006***	1.1200***	0.0393***	-0.4200**	-0.0008	-0.0009	-0.0008	0.0017*	0.0006
Age	0.0180	-0.0005*	-0.0704*	-0.0029*	0.0183	-0.0004***	-0.0003**	0.0000	0.0000	0.0001*
Lagged expense ratio	-1492.6910	0.5516	-4.4518	12.3129	-47.9759	0.4791	-0.8830	-0.0826	0.2490	-0.7079
Volatility	-0.2792	-0.6936	40.4577	5.0113	-217.5784	-1.5193**	-1.2280	-0.5837	-1.1928	-0.0164
Lagged ASVI	0.1311	0.0020	-0.1130	0.0084	0.0521	0.0015	0.0031	0.0060	0.0000	0.0001
_cons	42.9638	0.0389***	-6.2346***	-0.5024	4.4833**	0.0074	0.0127	-0.0013	-0.0252***	-0.0137**
R-squared	0.0134	0.0216	0.0173	0.0079	0.0047	0.0669	0.0044	0.0063	0.0044	0.0146

***significant at the 1% level **significant at the 5% level *significant at the 1% level

Lagged TNA, contrary to the general finding from Table, has a strong presence as a predictor of flow on a yearly basis. Until 2009, it has mostly maintained a positive correlation with flow, indicating that before the financial distress, investors are more prone to investment into larger funds. Specifically, for every log point increase in lagged TNA, a sample fund flow received 63% more flows in 2006. The same numbers for 2007 – 2009 are -0.5%, 112% and 3.93% respectively. All of these figures are statistically significant at the 1% level. On the other hand, after 2009, fund size starts to decline with flow, meaning investors tend to stray away from large flows after the financial crisis. The most notable example is in 2010, where a single log point increase in lagged TNA would cause flow to drop by 42%, which is significant at the 5% level. This inverse trend continues for four years, although the numbers after 2010 are not as significant. Fund size starts to regain its positive relevance with flow again only recently, during the past 2 years, but its coefficients are still closer to zero. All of these findings partly agree with the popular notion that fund size erodes performance. In this paper, thanks to the yearly analysis, the regression suggests an interesting result: in good market condition, investors are more likely to invest their worth into bigger funds, but during the times of crises, smaller-sized funds are a preferable choice.

As for fund age, it does not seem to have a strong effect on fund flow. Most of the coefficients over the years have largely stayed around the zero mark, except for the year 2008. In 2008, a fund having existed 1 year longer than a sample fund would receive around 7% less flow, and this result is significant at the 10% level. Lagged expense ratio's coefficients also mostly lie in the negative zones, indicating that fund flows are usually higher for funds with lower fees. This observation agrees with previous literature, but the results are not statistically significant. Thus, it is not possible to draw any conclusion regarding the relation between fund fees and flow from this dataset.

The variable of most relevance to the purpose of this study is lagged ASVI, which depicts the changes in search volumes initiated by the investors. Looking at the regression results, it is easy to see that a change in search interests are positively linked to subsequent mutual fund flow. For example, in 2006, an increase of 1% in search interests raises the next month's flow by 13.1%. An interesting example would be the year 2008, where an increase in search interests would lead to a drop in flow. With the fact that this was the year of the beginning of the financial distress, it can be explained that the inverse relation between these two elements are caused by a surge in search interests for *bad* news. In other words, during bad times, investors tend to search more for negative news, increasing the abnormal search volumes, but the negativity urges them to withdraw money from their funds. In such a case, a rise in search interests would be linked to a decline in flow. Nevertheless, throughout the years lagged ASVI mostly returns a positive coefficient, indicating that a general increase in search volumes would lead to an increase in flow as well. Unfortunately, no figures for lagged ASVI can achieve statistical significance in the t-test, so it is impossible to confirm the above observation empirically. As a result, the yearly regression analysis also leads to a rejection of hypothesis H2, validating the finding from the overall regression.

5.4.3. The influence of fund size, performance and age on SVI and fund flow

This section describes the three elements: fund size, performance, age and their interaction with ASVI in the regression analysis. The similar regression as above will be performed each time using a different interaction term to depict the elements' influence on the SVI impact on fund flow. In other words, this part attempts to answer the question: would the flow of small-sized funds be more affected by SVI than large-sized funds? Do worse-performing funds and older funds experience a similar effect? Answering these questions would also reveal whether the hypotheses H3, H4, H5 can be accepted.

H3: Fund size influences the impact of search interests on fund flow

H4: Past performance influences the impact of search interests on fund flow

Table 9 shows the regression results with the interaction term between lagged TNA (log) and lagged ASVI. This interaction term's values show how lagged TNA plays a role in determining lagged ASVI's impact on fund flow. Based on the results, fund size has a slight influence on search interest's impact on fund flow. Specifically, the interaction term has a negative coefficient of -0.06. The smaller percentage flow resulting from a change in SVI for the larger funds (or larger TNA) suggests that search interests play a more significant role for smaller-sized funds. However, the coefficient is not statistically significant at any level, thus overall there is no reliable empirical evidence on fund size's influence on SVI' effects.

Table 10 demonstrates the results from the same regression, done on a yearly basis from 2006 – 2015. For the first few years in the analysis (2006 and 2008), the interaction term has relatively high negative coefficients. For example, in 2006, a fund with a single log point of TNA higher than a sample fund would have its SVI effects on flow reduced by 21.4%. This negative pattern continues throughout the analyzed period, with the only exception in 2012 where the coefficient stays slightly above zero. As a result, the overview finding would be that in general larger-sized funds' flow would not be as affected by SVI as smaller-sized funds. In other words, the regression suggests that investors' search interests would generate a bigger impact on flow for smaller and lesser known funds (with smaller TNA). Unfortunately, this outcome cannot be empirically validated, since all of the interaction term's coefficients are not significant at any confidence interval. As a result, there is not enough evidence to accept hypothesis H3.

Table 11 shows the results from the regression where I interact the lagged return variable with lagged ASVI. The data demonstrates that in general, flows resulting from abnormal changes in SVI for better performing funds are lower than worse performing funds. Specifically, the interaction term has a coefficient of -0.199, indicating that investors' search interests generate a bigger impact for funds that are doing badly. This observation seems to be in line with the previous finding that investors may be more likely to search for negative news. Funds that are performing badly receive extra coverage from the media (Kaniel, et al., 2007), thus increasing the chance an investor can search for such news on the internet. When search interests for *bad* news about the funds increase, it is natural to see their flow dropping since investors are making an informed decision about their withdrawal. Again, this important coefficient is not statistically significant, so such finding cannot be backed with empirical evidence from this dataset.

Table 9: SVI-flow regression results with interaction between lagged TNA and lagged ASVI.

The table shows the statistical relationship between the dependent variable, fund flow, with the independent variables average market flow, lagged flow, lagged return, the logarithm of lagged Total Net Assets (TNA), fund's age, lagged expense ratio, volatility, lagged Abnormal Search Volume Index (ASVI) and the interaction term between the logarithm of lagged TNA with lagged ASVI. The regression is done based on the equation below, for the sample of 235 U.S mutual funds in the period 2006 - 2015:

$$Flow_t = \beta_0 + \beta_1 MF + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \beta_9 [log(TNA_{t-1}) * ASVI_{t-1}] + \varepsilon_t$$

where $Flow_t$ is the fund's flow in month t

MF is the Market Flow (average of all funds' flows) in month t

r_{t-1} is the fund's return in month t-1 (or lagged returns)

TNA_{t-1} is the total net asset of the fund in month t-1

Age is the number of years the fund has existed

$ExpenseRatio_{t-1}$ is the expense ratio of the fund in month t-1

σ_{t-1} is the volatility of the funds' returns in month t-1

$ASVI_{t-1}$ is the abnormal SVI of the fund in month t-1

Flow	Coef.	Std. Err.	t	P>t	95% Confidence interval	
Average market flow	0.0883	0.0195	4.5200	0.0000	0.0500	0.1266
Lagged flow	0.0000	0.0013	0.0100	0.9940	-0.0025	0.0025
Lagged return	4.7439	1.6589	2.8600	0.0040	1.4922	7.9955
Lagged TNA (log)	0.0217	0.0346	0.6300	0.5300	-0.0461	0.0895
Age	-0.0031	0.0053	-0.5800	0.5610	-0.0134	0.0073
Lagged expense ratio	1.0611	10.5404	0.1000	0.9200	-19.5988	21.7209
Volatility	-0.3106	29.7284	-0.0100	0.9920	-58.5801	57.9589
Lagged ASVI	0.4178	0.4162	1.0000	0.3150	-0.3980	1.2337
Lagged TNA (log) * Lagged ASVI	-0.0613	0.0641	-0.9600	0.3380	-0.1869	0.0643
_cons	-0.1350	0.2876	-0.47	0.639	-0.6986	0.4287
R-squared	0.0014					

***significant at the 1% level **significant at the 5% level *significant at the 1% level

Table 10: SVI-flow regression results with interaction between lagged TNA and lagged ASVI, on a yearly basis.

The table shows the statistical relationship between the dependent variable, fund flow, with the independent variables average market flow, lagged flow, lagged return, the logarithm of lagged Total Net Assets (TNA), fund's age, lagged expense ratio, volatility, lagged Abnormal Search Volume Index (ASVI) and the interaction term between the logarithm of lagged TNA with lagged ASVI. The regression is done on a yearly basis, based on the equation below, for the sample of 235 U.S mutual funds in the period 2006 - 2015:

$$Flow_t = \beta_0 + \beta_1 MF + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \beta_9 [log(TNA_{t-1}) * ASVI_{t-1}] + \varepsilon_t$$

where $Flow_t$ is the fund's flow in month t
MF is the Market Flow (average of all funds' flows) in month t
r_{t-1} is the fund's return in month t-1 (or lagged returns)
TNA_{t-1} is the total net asset of the fund in month t-1
Age is the number of years the fund has existed
$ExpenseRatio_{t-1}$ is the expense ratio of the fund in month t-1
σ_{t-1} is the volatility of the funds' returns in month t-1
$ASVI_{t-1}$ is the abnormal SVI of the fund in month t-1

Flow	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10
Obs	1814	2899	2306	2650	2564	2498	2436	2360	2380	2362
Average market flow	0.0597	0.7086*	1.0405***	0.0634	0.5812**	0.0003	-0.00072	-0.00119	-0.02968	-0.00592
Lagged flow	0.0001	0.0689***	0.0106	0.0084	-0.0007	0.0010***	0.0000	-0.0001	0.003758	0.0001
Lagged return	0.9713	0.2181	-4.5427	0.4482	-3.9464	0.0995	0.0374	0.1750***	0.0806	0.1433***
Lagged TNA (log)	-0.6465***	-0.0060***	1.1094***	0.0389***	-0.4195**	-0.0008	-0.0009	-0.0008	0.0017*	0.0006
Age	0.0189	-0.0005*	-0.0698	-0.0029*	0.0177	-0.0004***	-0.0003**	0.0000	0.0000	0.0001*
Lagged expense ratio	-1494.4360	0.5572	-8.1495	12.1976	-51.3146	0.4794	-1.0348	-0.0820	0.2513	-0.7109
Volatility	0.9512	-1.0350	51.6001	4.7831	-251.0487	-1.7099**	-0.8306	-0.9448	-1.1322	0.0003
Lagged ASVI	1.3875	0.0092	0.9788	0.0523	0.1349	0.0058	-0.0020	0.0115	0.0004	0.0051
Lagged TNA (log) x Lagged ASVI	-0.2144	-0.0012	-0.1811	-0.0076	-0.0109	-0.0007	0.0009	-0.0008	-0.0001	-0.0008
_cons	43.0832	0.0395***	-6.15606**	-0.4984	4.6824**	0.0082	0.0134	-0.0005	-0.0253***	-0.0135**
R-squared	0.0139	0.0217	0.0174	0.0079	0.005	0.0674	0.0043	0.0065	0.0044	0.0148

***significant at the 1% level **significant at the 5% level *significant at the 1% level

Table 11: SVI-flow regression results with interaction between lagged return and lagged ASVI.

The table shows the statistical relationship between the dependent variable, fund flow, with the independent variables average market flow, lagged flow, lagged return, the logarithm of lagged Total Net Assets (TNA), fund's age, lagged expense ratio, volatility, lagged Abnormal Search Volume Index (ASVI) and the interaction term between lagged return with lagged ASVI. The regression is done based on the equation below, for the sample of 235 U.S mutual funds in the period 2006 - 2015:

$$Flow_t = \beta_0 + \beta_1 MF + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \beta_9 [r_{t-1} * ASVI_{t-1}] + \varepsilon_t$$

where $Flow_t$ is the fund's flow in month t
MF is the Market Flow (average of all funds' flows) in month t
r_{t-1} is the fund's return in month t-1 (or lagged returns)
TNA_{t-1} is the total net asset of the fund in month t-1
Age is the number of years the fund has existed
$ExpenseRatio_{t-1}$ is the expense ratio of the fund in month t-1
σ_{t-1} is the volatility of the funds' returns in month t-1
$ASVI_{t-1}$ is the abnormal SVI of the fund in month t-1

Flow	Coef.	Std. Err.	t	P>t	95% Confidence Interval	
Average market flow	0.0881372	0.0195363	4.51	0	0.0498449	0.12643
Lagged flow	0.0000	0.0013	0.0100	0.9950	-0.0025	0.0025
Lagged return	4.7371	1.6736	2.8300	0.0050	1.4567	8.0175
Lagged TNA (log)	0.0250	0.0344	0.7300	0.4670	-0.0424	0.0925
Age	-0.0032	0.0053	-0.6100	0.5430	-0.0136	0.0072
Lagged expense ratio	1.0826	10.5410	0.1000	0.9180	-19.5784	21.7436
Volatility	-0.4829	29.7288	-0.0200	0.9870	-58.7532	57.7874
Lagged ASVI	0.0492	0.1557	0.3200	0.7520	-0.2560	0.3544
Lagged return * Lagged ASVI	-0.1986	3.0711	-0.0600	0.9480	-6.2182	5.8210
_cons	-0.1542	0.2869	-0.5400	0.5910	-0.7165	0.4080
R-squared	0.0014					

***significant at the 1% level **significant at the 5% level *significant at the 1% level

A look into the same regression conducted on a yearly basis (Table 12) reveals more interesting insights about this interaction between fund performance and search interests. Contrary to the general finding above, most of the years see a positive coefficient for the interaction term. The most prominent numbers are from 2006 and 2008, where the coefficients reach 8.94 and 2.99 respectively. These highly positive figures suggest that on a yearly basis before 2009, better performing funds have their flow more affected by a change in search interests. If we assume that investors are prone to good news in the good market condition (before the crisis), this finding makes sense. The prevalence of positive news about a specific well performing funds would attract and encourage prospective investors to make their investing decisions, resulting in higher flows. On the other hand, during the bad times of 2009 – 2010, the negative coefficients of the interaction terms suggest that worse performing funds' flow are more strongly affected by the fluctuation in search volumes. This is an expected result, since in the financial distress, negative news about bad performing funds was omnipresent on the internet, and the market condition then made them really relevant to an investor's interests. As a consequence, investors became more informed mostly about bad performing funds, so these funds' flow would be strongly impacted by the bad news. Again, the nature of search interests (positive or negative) is of utmost importance here, but unfortunately the data from Google Trends could not avoid this limitation. Despite all of the possible interpretations from the results, the coefficients are not significantly different from zero on any confidence interval, so the findings unfortunately cannot be backed by empirical evidence. Thus, the hypothesis H4 cannot be accepted.

Table 12: SVI-flow regression results with interaction between lagged return and lagged ASVI, on a yearly basis.

The table shows the statistical relationship between the dependent variable, fund flow, with the independent variables average market flow, lagged flow, lagged return, the logarithm of lagged Total Net Assets (TNA), fund's age, lagged expense ratio, volatility, lagged Abnormal Search Volume Index (ASVI) and the interaction term between lagged return with lagged ASVI. The regression is done on a yearly basis, based on the equation below, for the sample of 235 U.S mutual funds in the period 2006 - 2015:

$$Flow_t = \beta_0 + \beta_1 MF + \beta_2 Flow_{t-1} + \beta_3 r_{t-1} + \beta_4 log(TNA_{t-1}) + \beta_5 * Age + \beta_6 ExpenseRatio_{t-1} + \beta_7 \sigma_{t-1} + \beta_8 ASVI_{t-1} + \beta_9 [r_{t-1} * ASVI_{t-1}] + \varepsilon_t$$

where $Flow_t$ is the fund's flow in month t
MF is the Market Flow (average of all funds' flows) in month t
r_{t-1} is the fund's return in month t-1 (or lagged returns)
TNA_{t-1} is the total net asset of the fund in month t-1
Age is the number of years the fund has existed
$ExpenseRatio_{t-1}$ is the expense ratio of the fund in month t-1
σ_{t-1} is the volatility of the funds' returns in month t-1
$ASVI_{t-1}$ is the abnormal SVI of the fund in month t-1

Flow	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**
Obs	1814	2899	2306	2650	2564	2498	2436	2360	2380	2362
Average market flow	0.0593	0.7202	1.045***	0.0646	0.5815**	0.0003	-0.0007	-0.0013	-0.0282	-0.0059*
Lagged flow	0.0001	0.0689***	0.010613	0.0084	-0.0007	0.001***	0.0000	-0.0001	0.0039	0.0001
Lagged return	1.9810	0.2186*	-4.2915	0.4405	-4.0329	0.1027**	0.0386	0.1746***	0.0761	0.1434***
Lagged TNA (log)	-0.6296***	-0.0059***	1.1191***	0.0393***	-0.4191**	-0.0008	-0.0010	-0.0008	0.0017	0.0006
Age	0.0178	-0.0005*	-0.0703*	-0.0029*	0.0176	-0.0004***	-0.0003**	0.0000	0.0000	0.0001*
Lagged expense ratio	-1520.9790	0.5402	-7.4590	12.3808	-51.3634	0.4789	-1.0325	-0.0793	0.2500	-0.7080
Volatility	-20.8360	-1.0586	50.9295	4.7396	-251.2552	-1.7209**	-0.8239	-0.9484	-1.0768	0.0066
Lagged ASVI	0.0777	0.0010	-0.0083	0.0107	0.0879	0.0010	0.0030	0.0057	0.0014	0.0001
Lagged return x Lagged ASVI	8.9363	0.0893	2.9893	-0.0771	-1.1864	0.0644	0.0136	0.0225	-0.1771	0.0032
_cons	43.6965	0.0390***	-6.2133**	-0.5033	4.6834**	0.0080	0.0137	-0.0007	-0.0252***	-0.0138**
R-squared	0.0135	0.0217	0.0173	0.0079	0.005	0.0676	0.0042	0.0065	0.0047	0.0146

***significant at the 1% level **significant at the 5% level *significant at the 1% level

6. Conclusion

This section presents the key results from the paper. The first subsection reviews the important findings from the regression analyses while the second subsection suggests some new directions for future researches related to SVI and investor attention.

6.1. Summary of general findings

This paper studies the effects of Google search queries, measured by search volume index (SVI), on mutual fund flow and in addition fund performance. The study uses a sample of 235 U.S mutual funds in the period 2006 – 2015 as the dataset for the regression analyses. A specific focus of the study is on whether a change in abnormal search interests can predict a change in a fund's performance and flow. The paper aims to continue the trend of using Google search queries as a proxy for investor attention, pioneered by Da, et al. (2011). Additionally, the paper also contributes the literature surrounding the different determinants of fund flow, and is the first paper, as far as I am concerned, to make a connection between Google searches and fund flow.

The study first discovers a significant negative link between a change in SVI and the sample fund's performance. This is an unexpected result, since previous literature has mostly concluded a positive correlation between search interests and fund's returns. After that, using fund flow as the dependent variable in the regression, the study finds that flow is positively connected to investors' search volumes, but the result is not statistically significant. Thus, there is no resolute conclusion on whether search volume can act as a reliable predictor of future fund flow.

When looking at the analysis on a yearly basis, I identify an interesting trend: during positive market conditions, flow remains positively aligned with fluctuation in search interests, but during bad times (like the financial distress), they experience an inverse relationship. It can be explained by a surge in search interests for *bad* news, leading to investors' withdrawal of money. Hence, negative flows occur due to a rise in search interests.

In another analysis, I also interact fund size and performance with abnormal SVI to study their influence on SVI's impact on fund flow. The key results show that investors' search interests would generate a bigger impact on flows into smaller and lesser known funds. Additionally, well-performing funds usually have their flows more affected by SVI than bad-performing funds, but this effect is reverse during the years of the financial crisis. Unfortunately, these interpretations cannot be confirmed on an empirical basis, since the final results are not statistically different from zero. As a result, it is not possible to determine whether fund size and performance truly play a role in modifying SVI's impact on mutual fund flow.

6.2. Suggestions for future researches

There are many opportunities for future researches. First, as mentioned many times in this paper, current search data collected from Google cannot depict the nature of the search, being a positive or negative search. A future study with a proper way to differentiate between types of searches would offer phenomenal value in expanding the knowledge on internet searches as a proxy for investor attention. Second, it is interesting to identify different ways to compose search strings to input into Google Trends. There are various studies focusing on how different wording and more loosely related terms can also generate an impact on market movement (Challet & Ayed, 2014). Combining such studies with mutual funds would shine a new light into the many applications of SVI in finance. Third, social media is hugely popular with the investor community in recent years (for example, StockTwits), and these would be an invaluable source of information for everyone. Thus, it is relevant to use insights from social media as a proxy for investor attention. Although these alternative data sources are more likely to be subjected to noises, if handled right, they would still help expanding the many ways to capture attention in the modern internet-connected world and offer fresh perspectives in finance researches.

References

Askitas, N. & Zimmermann, K. F., 2009. Google econometrics and unemployment forecasting. *The Institute for the Study of Labor.*

Baker, S. & Fradkin, A., 2011. What Drives Job Search? Evidence from Google Search data. *Stanford Institute for Economic Policy Research.*

Barber, B. M. & Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies,* 21(2), pp. 785-818.

Barber, B. M., Odean, T. & Zheng, L., 2005. Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows. *The Journal of Business,* 78(6), pp. 2095-2120.

Bergstresser, D. & Poterba, J., 2002. Do after-tax returns affect mutual fund inflows?. *Journal of Financial Economics,* Volume 63, pp. 381-414.

Bijl, L., Kringhaug, G. & Sandvik, E., 2015. Predictive Power of Google Search Volume on Stock Returns.

Bloom, R., Amber, K. T., Hu, S. & Kirsner, R., 2015. Google Search Trends and Skin Cancer: Evaluating the US Population's Interest in Skin Cancer and Its Association With Melanoma Outcomes. *JAMA Dermatol,* 151(8), pp. 903-905.

Cao, C., Chang, E. C. & Wang, Y., 2008. An empirical analysis of the dynamic relationship between mutual fund flow and market return volatility. *Journal of Banking & Finance,* Volume 32, pp. 2111-2123.

Carhart, M. M., 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance,* 52(1), pp. 57-82.

Challet, D. & Ayed, A. B. H., 2014. Do Google Trend data contain more predictability than price returns?.

Challet, D. & Ayed, A. B. H., 2014. Predicting financial markets with Google Trends and not so random keywords.

Chen, J., Hong, H. G., Huang, M. & Kubik, J., 2004. Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization. *American Economic Review,* 94(5), pp. 1276-2302.

Choi, H. & Varian, H., 2011. Predicting the Present with Google Trends. *Google Inc.*

Da, Z., Engelberg, J. & Gao, P., 2011. In search of attention. *The Journal of Finance,* 66(5), pp. 1461-1499.

Dietzel, M. A., Braun, N. & Schaefers, W., 2014. Sentiment-Based Commercial Real Estate Forecasting with Google. *Journal of Property Invesment & Finance,* 32(6), pp. 540-569.

Drake, M. S., Roulstone, D. T. & Thornock, J. R., 2010. Googling for Information around Earnings Announcements. *Ohio State University.*

Dzielinski, M., 2012. Measuring economic uncertainty and its impact on the stock market. *Finance Research Letters,* 9(3), pp. 167-175.

Edelen, R. M. & Warner, J. B., 2001. Aggregate price e!ects of institutional trading: a study of mutual fund flow and market returns. *Journal of Financial Economics,* Volume 59, pp. 195-220.

Fricke, E., Fung, S. & Goktan, M. S., 2014. Google Search, Information Uncertainty, and Post-Earnings. *Journal of Accounting and Finance,* 14(2), pp. 11-27.

Ginsberg, J. et al., 2009. Detecting influenza epidemics using search engine query data. *Nature,* Volume 457, pp. 1012-1014.

Gruber, M. J., 1996. Another puzzle: The growth in actively managed mutual funds. *The Journal of Finance,* 51(3), pp. 783-810.

Guercio, D. D. & Tkac, P. A., 2008. Star Power: The Effect of Morningstar Ratings on Mutual Fund Flows. *Journal of Financial and Quantitative Analysis,* 43(4), pp. 907-936.

Guzman, G., 2011. Internet Search Behavior as an Economic Forecasting Tool: The case of Inflation Expectations. *Journal of economic and social measurement,* 36(3), pp. 119-167.

Huang, J., Kelsey, D. W. & Yan, H., 2007. Participation Costs and the Sensitivity of Fund Flows to Past Performance. *The Journal of Finance,* 63(3), pp. 1273-1311.

Huberman, G. & Tomer, R., 2001. Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *The Journal of Finance,* 56(1), pp. 387-396.

Indro, D., Jiang, C. X., Hu, M. Y. & Lee, W. Y., 1999. Mutual Fund Performance: Does Fund Size Matter?. *Financial Analysts Journal,* 55(3).

Investment Company Institute, 2015. *Recent Mutual Fund Trends.* [Online]
Available at: http://www.icifactbook.org/fb_ch2.html
[Accessed 01 05 2016].

Ippolito, R. A., 1992. Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. *The Journal of Law & Economics,* 35(1), pp. 45-70.

Johnson, H. A. et al., 2004. Analysis of Web Access Logs for Surveillance of Influenza. *Stud Health Technol Inform,* 107(2), pp. 1202-1206.

Kahneman, D., 1973. *Attention and effort.* Englewood Cliffs: NJ: Prentice-Hall.

Kaniel, R., Starks, L. & Vasudevan, V., 2007. Headlines and Bottom Lines: Attention and Learning Effects from Media Coverage of Mutual Funds. *SSRN Electronic Journal.*

Kennickell, A., Starr-McCluer, M. & Surette, B., 2000. Changes in U.S. family finances at the end of the 1990s: results from the 1998 survey of consumer finances. *Federal Reserve Bulletin,* 86(1), pp. 1-29.

Lui, C., Panagiotis, T. M. & Eni, M., 2011. On the predictability of the US elections through search volume activity.

Patel, J., Zeckhauser, R. & Hendricks, D., 1994. Investment flows and performance: evidence from mutual funds, cross-border investments, and new issues. In: *Japan, Europe and international financial markets.* New York: Cambridge University Press, pp. 51-70.

Shimshoni, Y., Efron, N. & Matias, Y., 2009. On the Predictability of Search Trends.

Sirri, E. R. & Peter, T., 1998. Costly Search and Mutual Fund Flows. *The Journal of Finance,* 53(5), pp. 1589-1622.

Solomon, D. H., Soltes, E. F. & Sosyura, D., 2014. Winners in the spotlight: Media Coverage of Fund Holdings as a Driver of Flows. *Journal of Financial Economics,* 113(1), pp. 53-72.

Statista, 2015. *Number of mutual funds in the United States from 1997 to 2014.* [Online]
Available at: http://www.statista.com/statistics/255590/number-of-mutual-fund-companies-in-the-united-states/
[Accessed 01 05 2016].

Takeda, F. & Wakao, T., 2014. Google search intensity and its relationship with returns and trading volume of Japanese stocks. *Pacific-Basin Finance Journal,* Volume 27, pp. 1-18.

Willard, S. D. & Nguyen, M. M., 2013. Internet Search Trends Analysis Tools Can Provide Real-time Data on Kidney Stone Disease in the United States. *Urology,* 81(1), pp. 37-42.

Wuoristo, L., 2012. Can you google the future? A study on the predictive power of Google Trends on company shares in the UK.

Yuan, Y., 2008. Attention and trading. *Unpublished working paper. University of Iowa.*

Appendix

Appendix 1: The code for the Google Trends web crawling program

```
from pytrends.pyGTrends import pyGTrends
import time
from random import randint

def run_search(keyword):
    google_username = '' #Enter google username
    google_password = '' #Enter google password

    #Connect to google
    connector = pyGTrends(google_username, google_password)

    #Request a report based on keyword
    connector.request_report(keyword, geo = 'US', date = '03/2006 118m')
    print(keyword)

    #Wait a bit so Google doesn't block the script
    time.sleep(randint(5,10))

    #Save a csv
    path = './output/'
    connector.save_csv(path, keyword)

with open("fund name.txt", "r") as source:
    lines = [line.rstrip('\n').replace('/', ' ').replace(',','') for line in source]
    for name in lines:
        run_search(name)

print("Finished!")
```

Appendix 2: The list of chosen mutual funds

AARP Funds: AARP Aggressive Fund

Advance Capital I, Inc.: Balanced Fund; Retail Class Shares

Advance Capital I, Inc.: Equity Growth Fund; Retail Class Shares

Advantage Funds, Inc.: Global Alpha Fund; Class T Shares

Advisors' Inner Circle Fund: Westwood Income Opportunity Fund; Institutional Shares

Aegis Funds: Aegis Value Fund; Class I Shares

AIM Sector Funds (Invesco Sector Funds): Invesco Energy Fund; Investor Class Shares

Alpine Series Trust: Alpine Dynamic Dividend Fund; Institutional Class Shares

Amana Mutual Funds Trust: Growth Fund; Investor Class Shares

Amana Mutual Funds Trust: Income Fund; Investor Class Shares

AMCAP Fund; Class A Shares

American Balanced Fund; Class A Shares

American Century Capital Portfolios, Inc.: Mid Cap Value Fund; Class R Shares

American Century Government Income Trust: Government Bond Fund; Investor Class Shares

American Century International Bond Funds: International Bond Fund; Investor Class Shares

American Century Investment Trust: Prime Money Market Fund; Investor Class Shares

American Century Mutual Funds, Inc.: Capital Growth Fund; B Class Shares

American Century Mutual Funds, Inc.: Capital Value Fund; Investor Class Shares

American Century Mutual Funds, Inc.: Heritage Fund; Class C Shares

American Century World Mutual Funds, Inc.: Emerging Markets Fund; C Class Shares

American Century World Mutual Funds, Inc.: Global Growth Fund; Class B Shares

American Mutual Fund; Class A Shares

Aquila Funds Trust: Aquila Three Peaks Opportunity Growth Fund; Class A Shares

Arbitrage Funds: Arbitrage Fund; Class R Shares

Ariel Investment Trust: Ariel Appreciation Fund; Investor Class Shares

Ariel Investment Trust: Ariel Fund; Investor Class Shares

Artisan Partners Funds, Inc.: Artisan International Fund; Investor Shares

Artisan Partners Funds, Inc.: Artisan International Value Fund; Investor Shares

Artisan Partners Funds, Inc.: Artisan Value Fund; Investor Shares

Baron Investment Funds Trust: Baron Asset Fund; Retail Shares

Baron Investment Funds Trust: Baron Growth Fund; Retail Shares

Baron Investment Funds Trust: Baron Opportunity Fund; Retail Shares

Baron Investment Funds Trust: Baron Small Cap Fund; Retail Shares

Baron Select Funds: Baron Partners Fund; Retail Shares

Bertolet Capital Trust: Pinnacle Value Fund

Berwyn Funds: Berwyn Fund

Bishop Street Funds: Dividend Value Fund; Class I Shares

Bishop Street Funds: Government Money Market Fund; Class I Shares

Bishop Street Funds: Large Cap Growth Fund; Institutional Class Shares

BlackRock Bond Fund, Inc.: BlackRock Bond Fund; Investor A1 Shares

BlackRock Bond Fund, Inc.: BlackRock Total Return Fund; Class K Shares

BlackRock Equity Dividend Fund; Investor A Shares

BlackRock Funds III: BlackRock S&P 500 Index Fund; Class K Shares

Bond Fund of America; A

Bridgeway Funds, Inc.: Bridgeway Blue Chip 35 Index Fund

Broadview Funds Trust: Broadview Opportunity Fund

Buffalo Funds: Buffalo Small Cap Fund

Calamos Investment Trust: Calamos Growth Fund; Class A Shares

Capital Income Builder; Class A Shares

Cash Management Trust of America; Class A Shares

Cash Reserve Fund, Inc.: Prime Series; Cash Reserve Prime Class Shares

CGM Trust: CGM Focus Fund

CGM Trust: CGM Mutual Fund

CGM Trust: CGM Realty Fund

Columbia Acorn Trust: Columbia Acorn Fund; Class A Shares

Columbia Acorn Trust: Columbia Acorn International; Class Z Shares

Columbia Funds Series Trust I: Columbia Balanced Fund; Class Z Shares

Columbia Funds Series Trust I: Columbia Dividend Income Fund; Class T Shares

Commerce Funds: Bond Fund

Consulting Group Capital Markets Funds: High Yield Investments

Consulting Group Capital Markets Funds: Money Market Investments

Croft Funds Corporation: Croft Value Fund; Class R Shares

Davis New York Venture Fund, Inc.: Davis New York Venture Fund; Class A Shares

Delaware Group Adviser Funds: Delaware Diversified Income Fund; Class R Shares

Dodge & Cox Funds: Dodge & Cox Balanced Fund

Dodge & Cox Funds: Dodge & Cox Income Fund

Dodge & Cox Funds: Dodge & Cox International Stock Fund

Dodge & Cox Funds: Dodge & Cox Stock Fund

Dreyfus LifeTime Portfolios, Inc.: Growth Portfolio; Investors Shares

Dreyfus LifeTime Portfolios, Inc.: Income Portfolio; Investors Shares

Dreyfus Premier Investment Funds, Inc.: Dreyfus Greater China Fund; Class A Shares

Dreyfus State Municipal Bond Fund: Maryland Fund; Class C Shares

Dreyfus State Municipal Bond Fund: Michigan Fund; Class B Shares

Dreyfus State Municipal Bond Fund: Minnesota Fund; Class C Shares

Dreyfus State Municipal Bond Fund: North Carolina Fund; Class B Shares

Dreyfus State Municipal Bond Fund: Ohio Fund; Class C Shares

Dreyfus/Laurel Funds Trust: Dreyfus International Bond Fund; Class I Shares

Dryden Municipal Series Fund: Florida Series; Class A Shares

Dryden Municipal Series Fund: New Jersey Series; Class A Shares

Dryden Municipal Series Fund: Pennsylvania Series; Class A Shares

Eaton Vance Municipals Trust: Eaton Vance National Municipal Income Fund; Class B Shares

Eaton Vance Special Investment Trust: Eaton Vance Greater India Fund; Class B Shares

Endowments Trust: Bond Portfolio

EquiTrust Series Fund, Inc.: Managed Portfolio; Class A Shares

EquiTrust Series Fund, Inc.: Managed Portfolio; Traditional Shares Class B

EuroPacific Growth Fund; Class A Shares

Evergreen Money Market Trust: Evergreen Money Market Fund; Class A Shares

Excelsior Funds, Inc.: Government Money Fund

Fairholme Funds, Inc.: Fairholme Fund

Federated Equity Funds: Federated Kaufmann Fund; Class R Shares

Federated Equity Funds: Federated Strategic Value Dividend Fund; Institutional Shares

Federated High Yield Trust; Service Shares

Federated Income Securities Trust: Federated Stock & California Muni Fund, Inc.; Class A Shares

Fidelity Advisor Series I: Fidelity Advisor Leveraged Company Stock Fund; Class I Shares

Fidelity Capital Trust: Fidelity Capital Appreciation Fund

Fidelity Capital Trust: Fidelity Focused Stock Fund

Fidelity Capital Trust: Fidelity Value Fund

Fidelity Colchester Street Trust: Government Portfolio; Select Class Shares

Fidelity Commonwealth Trust: Fidelity Small Cap Discovery Fund

Fidelity Contrafund

Fidelity Contrafund: Fidelity Advisor New Insights Fund; Class A Shares

Fidelity Financial Trust: Fidelity Independence Fund

Fidelity Hastings Street Trust: Fidelity Fund

Fidelity Hastings Street Trust: Fidelity Growth Discovery Fund

Fidelity Hereford Street Trust: Fidelity Money Market Fund

Fidelity Income Fund: Fidelity Government Income Fund

Fidelity Income Fund: Fidelity Total Bond Fund; Fidelity Total Bond Shares

Fidelity Investment Trust: Fidelity Canada Fund; Canada Shares

Fidelity Investment Trust: Fidelity China Region Fund

Fidelity Investment Trust: Fidelity Diversified International Fund

Fidelity Investment Trust: Fidelity Emerging Markets Fund

Fidelity Investment Trust: Fidelity Europe Fund

Fidelity Investment Trust: Fidelity International Discovery Fund

Fidelity Investment Trust: Fidelity Japan Fund

Fidelity Investment Trust: Fidelity Latin America Fund

Fidelity Magellan Fund

Fidelity Mt Vernon Street Trust: Fidelity Growth Company Fund

Fidelity Phillips Street Trust: Fidelity Government Cash Reserves

Fidelity Puritan Trust: Fidelity Balanced Fund

Fidelity Puritan Trust: Fidelity Low-Priced Stock Fund

Fidelity Puritan Trust: Fidelity Puritan Fund

Fidelity School Street Trust: Fidelity Strategic Income Fund

Fidelity Securities Fund: Fidelity Blue Chip Growth Fund

Fidelity Securities Fund: Fidelity Dividend Growth Fund

Fidelity Securities Fund: Fidelity Leveraged Company Stock Fund

Fidelity Securities Fund: Fidelity OTC Portfolio

Fidelity Select Portfolios: Energy Portfolio

Fidelity Select Portfolios: Insurance Portfolio

Fidelity Union Street Trust II: Fidelity Municipal Money Market Fund

First American Investment Funds, Inc.: International Fund; Class C Shares

First Eagle Funds: First Eagle Global Fund; Class A Shares

First Eagle Funds: First Eagle Gold Fund; Class A Shares

First Eagle Funds: First Eagle Overseas Fund; Class A Shares

First Investors Equity Funds: Global Fund; Class A Shares

First Investors Equity Funds: Opportunity Fund; Class A Shares

First Investors Income Funds: Fund For Income; Class A Shares

First Investors Income Funds: Government Fund; Class A Shares

Forester Funds, Inc.: Forester Value Fund; Class N Shares

Forum Funds: Auxier Focus Fund; Investor Shares

Forum Funds: Merk Hard Currency Fund; Investor Shares

Forum Funds: Polaris Global Value Fund

FPA Funds Trust: FPA Crescent Fund

Frank Funds: Frank Value Fund; Investor Class Shares

Franklin Custodian Funds: Franklin Growth Fund; Class A Shares

Franklin Custodian Funds: Franklin Income Fund; Class A Shares

Franklin Custodian Funds: Franklin Utilities Fund; Class A Shares

Franklin Gold and Precious Metals Fund; Class A Shares

Franklin High Income Trust: Franklin High Income Fund; Class A Shares

Franklin Money Fund; Class A Shares

Franklin Strategic Series: Franklin Biotechnology Discovery Fund; Class A Shares

Gabelli Asset Fund; Class AAA Shares

Gabelli Equity Series Funds, Inc.: Gabelli Focus Five Fund; Class AAA Shares

Gabelli Investor Funds, Inc.: Gabelli ABC Fund; Class AAA Shares

Gabelli Utilities Fund; Class AAA Shares

Gateway Trust: Gateway Fund; Class A Shares

GMO Trust: GMO Quality Fund; Class VI Shares

Goldman Sachs Trust: Goldman Sachs Mid Cap Value Fund; Service Shares

Growth Fund of America; Class A Shares

GuideStone Funds: Equity Index Fund; Investor Class Shares

GuideStone Funds: Growth Allocation Fund; Investor Class Shares

GuideStone Funds: Growth Equity Fund; Investor Class Shares

GuideStone Funds: Money Market Fund; Investor Class Shares

Guinness Atkinson Funds: Alternative Energy Fund

Harbor Funds: Harbor Bond Fund; Institutional Class Shares

Harbor Funds: Harbor Capital Appreciation Fund; Institutional Class Shares

Harbor Funds: Harbor International Fund; Institutional Class Shares

Harris Associates Investment Trust: Oakmark Equity and Income Fund; Class I Shares

Harris Associates Investment Trust: Oakmark Fund; Class I Shares

Harris Associates Investment Trust: Oakmark Global Select Fund; Class I Shares

Harris Associates Investment Trust: Oakmark International Fund; Class I Shares

Harris Associates Investment Trust: Oakmark Select Fund; Class I Shares

Hartford Mutual Funds, Inc.: Hartford Capital Appreciation Fund; Class A Shares

Hartford Mutual Funds, Inc.: Hartford Floating Rate Fund; Class Y Shares

Hartford Mutual Funds, Inc.: Hartford Healthcare Fund; Class Y Shares

Heartland Group, Inc.: Heartland Value Fund; Investor Class Shares

Heartland Group, Inc.: Heartland Value Plus Fund; Investor Class Shares

Hennessy Funds Trust: Hennessy Cornerstone Growth Fund, Series II; Original Class Shares

Hennessy Funds Trust: Hennessy Focus Fund; Investor Class Shares

Heritage Income Trust: High Yield Bond Fund; Class B Shares

Heritage Series Trust: Mid Cap Stock Fund; Class B Shares

HIMCO Variable Insurance Trust: HIMCO VIT Index Fund; Class IB Shares

Hussman Investment Trust: Hussman Strategic Growth Fund

ICON Funds: ICON Energy Fund; Class S Shares

Income Fund of America; Class A Shares

Investment Company of America; Class A Shares

iShares Gold Trust

iShares Silver Trust

iShares Trust: iShares China Large-Cap ETF

iShares Trust: iShares Core S&P 500 ETF

iShares Trust: iShares iBoxx $ High Yield Corporate Bond ETF

iShares Trust: iShares MSCI EAFE ETF

iShares Trust: iShares Select Dividend ETF

iShares Trust: iShares US Preferred Stock ETF

Ivy Funds: Ivy Asset Strategy Fund; Class Y Shares

Ivy Funds: Ivy High Income Fund; Class Y Shares

Jacob Funds Inc.: Jacob Internet Fund; Investor Class Shares

Janus Aspen Series: Balanced Portfolio; Institutional Shares

Janus Investment Fund: Janus Balanced Fund; Class T Shares

Janus Investment Fund: Janus Contrarian Fund; Class T Shares

Janus Investment Fund: Janus Enterprise Fund; Class T Shares

Janus Investment Fund: Janus Flexible Bond Fund; Class T Shares

Janus Investment Fund: Janus Fund; Class T Shares

Janus Investment Fund: Janus Global Life Sciences Fund; Class T Shares

Janus Investment Fund: Janus Overseas Fund; Class T Shares

Janus Investment Fund: Janus Research Fund; Class T Shares

Janus Investment Fund: Janus Triton Fund; Class T Shares

Janus Investment Fund: Janus Twenty Fund; Class T Shares

Janus Investment Fund: Janus Venture Fund; Class T Shares

Janus Investment Fund: Perkins Mid Cap Value Fund; Class L Shares

Jensen Quality Growth Fund; Class J Shares

John Hancock Capital Series: John Hancock Allocation Growth + Value Portfolio; Class A Shares

John Hancock Funds II: Blue Chip Growth Fund; Class 1 Shares

John Hancock Funds II: Capital Appreciation Fund; Class 1 Shares

John Hancock Funds II: Core Equity Fund; Class 1 Shares

John Hancock Funds II: Global Bond Fund; Class 1 Shares

John Hancock Funds II: High Yield Fund; Class 1 Shares

John Hancock Funds II: Index 500 Fund; Class NAV Shares

John Hancock Funds II: International Small Cap Fund; Class 1 Shares

John Hancock Funds II: Large Cap Fund; Class 1 Shares

John Hancock Funds II: Large Cap Value Fund; Class 1 Shares

John Hancock Funds II: Large Cap Value Fund; Class NAV Shares

John Hancock Funds II: Mid Cap Stock Fund; Class 1 Shares

John Hancock Funds II: Natural Resources Fund; Class 1 Shares

JPMorgan Trust I: JPMorgan Prime Money Market Fund; Capital Shares

Kinetics Mutual Funds, Inc.: Internet Fund; No Load Class

Nationwide Mutual Funds: Nationwide Fund; Institutional Service Class Shares

Nomura Partners Funds, Inc.: Japan Fund; Class S Shares

Northern Lights Fund Trust II: Al Frank Fund; Advisor Class Shares

Northern Lights Fund Trust: Free Enterprise Action Fund

Professionally Managed Portfolios: Winslow Green Growth Fund; Institutional Class Shares

Sanford C Bernstein Fund, Inc.: International Portfolio; AB International Class C Shares

Tocqueville Trust: Delafield Fund

Voya Corporate Leaders Trust Fund

William Blair Funds: Large Cap Growth Fund; Class N Shares

William Blair Funds: Mid Cap Growth Fund; Class N Shares