Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

Hannah Hjelle Høydal
Caroline Lysebo

Quantifying attention: utilizing Google searches to forecast stock performance of business-to-consumer companies

Master's thesis in Industrial Economics and Technology Management
Supervisor: Peter Molnar
June 2021

NTNU
Norwegian University of Science and Technology

Hannah Hjelle Høydal
Caroline Lysebo

Quantifying attention: utilizing Google searches to forecast stock performance of business-to-consumer companies

Master's thesis in Industrial Economics and Technology Management
Supervisor: Peter Molnar
June 2021

Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

NTNU
Norwegian University of
Science and Technology

Abstract

Google search data has been utilized in various applications, including financial forecasting. For stock return predictions, the common approach is to use Google searches for company tickers as a measure of investor attention. We re-investigate the topic by focusing on consumer-related companies and introduce measures of consumer attention. We analyze the companies of S&P 500 Consumer Discretionary and S&P 500 Consumer Staples with an initial hypothesis that consumer related stocks are driven by expected future earnings, which have potential to be reflected by patterns in Google searches. By utilizing the measures of attention, we are able to improve stock performance predictions, especially for longer time horizons and for companies within the Discretionary sector. We simulate a trading strategy to test for economic significance, and find the inclusion of attention measures could improve the yearly accumulated return by 1%.

Keywords: Google searches; consumer attention; investor attention; stock returns

Sammendrag

Google søkedata har de siste årene blitt anvendt til stadig flere formål, deriblant finansielle prediksjoner. Hva angår prediksjon av aksjekurser, har den vanligste tilnærmingen vært å benytte søkevolum for tickere som mål på investorers interesse. Vi undersøker dette temaet nærmere ved å fokusere på konsumrelaterte selskaper, og foreslår nye mål på forbrukernes interesse og oppmerksomhet. Vi analyser selskapene i S&P 500 Consumer Discretionary og S&P 500 Consumer Staples med en initiell hypotese om at aksjekursen til konsumrelaterte selskaper er drevet av forventinger om fremtidig inntjening, og at svingninger i søkevolum har potensiale til å predikere dette. Vi kombinerer de foreslåtte interessevariablene med standard finansielle variabler, og finner at dette forbedrer nøyaktigheten til prediksjonsmodellen. Dette gjelder særlig når vi predikerer akkumulert avkastning over lenger horisonter, og selskaper kategorisert som "Discretionary". Vi tester den økonomiske signifikansen av funnene våre ved å simulerere en investeringsstrategi, og finner at anvendelsen av de foreslåtte interessevariablene kan forbedre årlig akkumulert avkastning med nærmere 1%.

Nøkkelord: Google søk; forbukerinteresse; investorinteresse; aksjeavkastning

Preface

The thesis concludes our Master of Science in Industrial Economics and Technology Management within Financial Engineering at the Norwegian University of Science and Technology (NTNU) in the spring of 2021.

This thesis should be of interest to scholars and researchers, in addition to financial market practitioners being introduced to new measures for predicting the performance of consumer related stocks. The results can also be valuable for companies gaining insight into the attention patterns of their consumers.

We would like to thank our supervisor Peter Mólnar, Associate Professor at the Norwegian University of Science and Technology and University of Stavanger, for helpful guidance and constructive feedback.

Trondheim, June 10^{th}, 2021

Hannah Hjelle Høydal   Caroline Lysebo

Table of Contents

1 Introduction **8**

2 Literature Review **11**
2.1 Measures of attention 11
2.1.1 Investor attention 12
2.1.2 Consumer attention 13
2.2 Company segmentation and consumer behavior 14
2.3 Delayed financial response to information 14

3 Data **16**
3.1 Google Trends Data 16
3.1.1 Keyword selection 17
3.1.2 Search Volume Index generation 18
3.2 Market Data 20
3.2.1 Return 21
3.2.2 Abnormal Return 21
3.2.3 Volatility 22
3.2.4 Abnormal Trading Volume 22
3.3 Overview of variables 24
3.4 Stationarity 25
3.5 Summary Statistics 25

4 Methodology **27**
4.1 Fama MacBeth regression model 27
4.1.1 Lagged abnormal return 27
4.1.2 Cumulative abnormal return 28
4.2 Individual regression model 29
4.3 Trading strategy 29

5 Results **31**
5.1 Prediction of lagged return 31
5.1.1 Impact of attention on different sectors 32
5.1.2 Directional impact of attention over time 34
5.2 Prediction of cumulative return 37
5.2.1 Prediction of consumer sectors' cumulative return 38
5.2.2 Prediction of consumer industries' cumulative return 42
5.2.3 Prediction of consumer companies' cumulative return 44

5.3 Robustness test: Alternative prediction model 46

6 Trading strategy **48**

6.1 Selecting trading threshold 48

6.2 Comparing trading performance by sector 52

6.3 Trading costs 55

6.4 Robustness test: Trading random 10% of companies 57

7 Conclusion **60**

References **62**

A Appendix **66**

A.1 Indices 66

A.1.1 S&P 500 Consumer Discretionary 66

A.1.2 S&P 500 Consumer Staples 69

A.2 GICS Industry Classification 71

A.3 Industry-related keywords 72

A.4 Brand-related keywords 73

List of Tables

1	Overview of variables	24
2	Correlation matrix	25
3	Fama MacBeth regression results for all companies on lagged abnormal return	32
4	Fama MacBeth regression results for Consumer Staples companies on lagged abnormal return	33
5	Fama MacBeth regression results for Consumer Discretionary companies on lagged abnormal return	34
6	Fama MacBeth regression results for all companies on cumulative abnormal return	38
7	Fama MacBeth regression results for Consumer Staples companies on cumulative abnormal return	39
8	Fama MacBeth regression results for Consumer Discretionary companies on cumulative abnormal return	40
9	Fama MacBeth regression results for all consumer industries on cumulative abnormal returns	43
10	Descriptive statistics from individual linear regressions for all companies	44
11	R^2 overview from individual linear regressions	46
12	Model robustness from choice of regression model	47
13	Trading strategy performance, volatility and Sharpe ratio for selected thresholds	51
14	Trading strategy performance by sector	53
15	Trading strategy performance including trading costs	56
18	Companies filtered by industry	71
19	Industry keywords	72
20	Brand related keywords	73

List of Figures

1 SVI creation procedure 18
2 Overview of sectors and industries 20
3 Impact of brand attention on abnormal return over time 35
4 Impact of industry attention on abnormal return over time 36
5 Impact of investor attention on abnormal return over time 36
6 Trading strategy performance for varying threshold 49
7 Increase in trading strategy performance by including attention variables for varying thresholds 50
8 Trading strategy performance by sector over time 54
9 Model robustness from random company selection when constructing portfolio for trading strategy 58

1 Introduction

Stock performance prediction is a subject of large interest in finance. A reasonably accurate prediction has the prospect to yield financial profits and be a valuable tool in financial applications. Stock returns are affected by a complex set of factors causing uncertainties and challenges in predicting the movements. Researchers, hedge fund managers, asset managers, brokers and other market participants invest considerable resources to investigate new and better methods and additions to this field (Demirer, Pierdzioch, & Zhang, 2017; Cooper, Gutierrez, & Marcum, 2005; Pesaran & Timmermann, 1995; Aiolfi & Favero, 2005). With advances of the digital era, the amount of data has increased and the ways of utilizing it as well. Statistics, technical analysis, fundamental analysis, regressions and more novel methods, such as applying neural networks, are all used to predict and benefit from the market's direction. None of these techniques have proven to be the consistently correct prediction tool so far.

The use of search engine data for financial performance analysis has seen increased attention after Google in 2006 made their search data publicly available through Google Trends. This allows researchers to gain insight into search volume time series for all words and phrases (keywords). The time series represent how search volumes fluctuate and can function as a proxy for attention (Da, Engelberg, & Gao, 2011). The use of Google search volume data, henceforth referred to as Search Volume Index (SVI), within research can broadly be separated into two categories: (1) SVI as a proxy for investor attention, where financial forecasting has been the main objective, and (2) SVI as a proxy for consumer attention, which has focused on aspects of consumer behavior, such as consumption and demand. Most research into forecasting financial markets using Google Trends has been directed at considering SVI as a capture of investor attention. Nevertheless, significant results in forecasting sales through the use of consumer attention has been presented, indicating that this proxy can also add value to forecasting other financial data. Sales is a considerable factor to earnings, and earnings fluctuations are significantly associated with share price changes (Ariff, Loh, & Chew, 1997). This indicates that analyzing consumer attention can provide valuable information in forecasting returns. Further review of the extent and results from current research is presented in Section 2.

In this thesis, we analyze the predictive power of investor attention and consumer attention. Investor attention is measured as an SVI created from company tickers, which is the standard approach within the literature. We quantify consumer attention from two separate SVIs, that have the potential to measure two distinct behaviors; attention towards a company and attention towards an industry. The measure reflecting attention towards a specific company is constructed as an SVI corresponding to the company's brands. To exemplify, we use the search volume fluctuations of "Fanta", "Sprite" and "Coca-Cola" to measure the consumer attention towards *Coca-Cola Company*, since they are all brands under the corporation. This quantifies consumer interest towards a company and the demand for its products. Consumer demand materializes in future earnings, which is reflected in stock prices, meaning that brand attention can have an influence on the company's financial performance. The other consumer attention SVI reflects the trends of an industry. It is constructed from search volume fluctuations of industry-related keywords, which is set to measure the consumer interest of an industry and the demand of its products. To exemplify, the search volumes of "tea", "cigarettes" and "smoothie" are used to measure the consumer interest of the Food, Beverage & Tobacco industry. Again, we want to analyze whether this can predict the financial performance of the industry through its reflection of demand and future earnings. All ticker, brand and industry keywords can be viewed in Appendix.

The goal of this research is to investigate the relationship between attention measures and company performance and how this provides value in financial forecasts. Consumer attention has the potential to contribute with new information exposing trends, sentiment and other factors not incorporated by the financial variables. We limit the scope to companies where consumer attention is likely to play an important role, and therefore study the consumer related companies of S&P 500, which consists of companies within the S&P 500 Consumer Discretionary and S&P 500 Consumer Staples indices. Discretionary goods and services, also called cyclical products, are considered to be non-essential by consumers, but desirable if their income is sufficient to purchase them. Examples of such products are durable goods, high-end apparel and entertainment. Consumer Staples are essential products that people are unable or unwilling to cut out of their budget, regardless of their financial situation. Examples are food and beverages, hygiene products and household goods.

Attention measured by Google searches has previously been used to forecast company performance, but mostly as investor attention captured by Google searches for company tickers. We propose two new measures of consumer attention, and investigate how these along with investor attention are adding predictive power. We find that adding attention variables to the regressions is improving the forecasting accuracy, especially when predicting over longer horizons and companies categorized as Discretionary.

The rest of the paper is organized as follows: Section 2 reviews related literature. Section 3 describes the data sources and collection methods we have used. Section 4 presents the methodology. Section 5 describes the results. Section 6 evaluates a trading strategy building on the prediction model. Section 7 tests the model's robustness and Section 8 concludes.

2 Literature Review

In this section we discuss relevant existing literature. This will put our study into perspective and show how it contributes to current research. First, we will look at how attention is measured and how these measures are utilized. We will especially focus on the use of search volume data in forecasting, as a proxy for both investor and consumer attention. Then, we will review how different types of consumer companies typically react and behave in the market. Finally, we will explore how the financial performance of these companies can be expected to respond to events, and specifically review literature that analyzes the time delay of this response.

2.1 Measures of attention

The process of adequately quantifying attention has been discussed in various research, and several approaches have been tested. The use of attention in forecasting builds on the assumption that attention is an indication of gained information. This information is used in decision-making, and attention can therefore have predictive power on these decisions. Today, online sources stand as the main provider of information, driving researchers to utilize internet activity as a measure of attention (Subrahmanyam, 2019). Several trading strategies have been based on news article counts under the assumption that stock prices are determined by the human behavior of investors, and investors determine stock prices by using publicly available information (Gidófalvi, 2004; Shynkevich, Coleman, Mcginnity, & Belatreche, 2015). Others use text mining technology to quantify the unstructured data of social media (Nikfarjam, Emadzadeh, & Muthaiyah, 2010). Coyne, Madiraju, and Coelho (2017) analyze tweets, likes, follower counts and more, but don't find any strong correlation to stock prices. Audrino, Sigrist, and Ballinari (2020) analyze the impact of sentiment and attention variables on stock market volatility combining social media, news articles, information consumption and search engine data. They are able to improve volatility forecasts significantly, but the magnitude of the improvements is relatively small from an economic point of view. When Google in 2006 made search volume time series publicly available through Google Trends, it let researchers use these time series as a proxy for attention (Da et al., 2011). Utilizing this measure of attention has resulted in various research. Challet and Bel Hadj Ayed (2013) observe that the choice of keywords is crucial, and when applied to suitable assets yield robustly profitable strategies. Resom, Pierre, Klimkiewicz, and Kalampalikis (2018) develop a profitable Dow Jones Index trading strategy based on search volume indices and hypothesizes that similar methodologies are likely to be profitable to numerous other assets as well. Based on this previous research it seems to be valuable information to extract from Google search data, with several immature areas to be further researched.

In regards to financial forecasting, research utilizing Google search volumes as a measure of attention is mainly directed towards investor attention. However, it is no longer sufficient to only consider the attention of investors, as everyday people's opinions and attentiveness can be a key driver of financial performance (Juan Piñeiro-Chousa & Ribeiro-Soriano, 2020). Attention measures should therefore aim at quantifying various types of attention when forecasting stock returns. The following subsections review the two areas of attention measures used in research; investor attention, where financial forecasting has been the main objective, and consumer attention, which has focused on aspects of consumer behavior, such as consumption and demand.

2.1.1 Investor attention

Researchers have aimed at utilizing Google Trends in stock market predictions by using the time series as a proxy for investor attention. This assumes that Google is used by a high amount of investors in relation to the buy or sell of a stock. As professional investors likely use paid data sources for information, Google Trends is set to measure the attention mainly of uninformed investors, often described as retail investors (Da et al., 2011). This limits the share of investors, whose attention, search volume data is set to measure. Nevertheless, Da et al. (2011) find that even if SVI likely measures the attention only of retail investors, it has predictive power on stock prices. Other research on using SVIs as proxies for investors have been conducted into different areas of financial markets and achieved varying results. Kim, Lučivjanská, Molnár, and Villa (2019) find that Google searches neither correlate nor predict abnormal returns, while Swamy and Dharani (2019) find that SVI can predict stock price movements. Hamid and Heiden (2015) use SVI to create a model that significantly outperforms conventional time series models in forecasting volatility, and Challet and Bel Hadj Ayed (2013) conclude that there is consistently some predictive power in Google Trends data, but that it is mostly valid on average. Heyman, Lescrauwaet, and Stieperaere (2019) find that the best performing stocks are likely to revert after a surge in Google search volumes, and Bijl, Kringhaug, and Sandvik (2015) present various significance dependent on the forecast horizon. Ding and Hou (2015) focus fully on retail investor attention, and use SVI to show how the attention of retail investors affects different factors in financial markets, such as shareholder base and stock liquidity.

Da et al. (2011) argue that searches for company name is a bad proxy of investor attention, and that it is better to use the company ticker. There are several concerns with using ticker as a measure of investor attention, but it is still frequently used. Ding and Hou (2015) state three main reasons why stock tickers should be used as search keywords over company names. 1) By applying tickers we avoid the issues with multiple reference names since it works as an unique identifier. 2) Searches for tickers are primarily undertaken by individuals interested in financial information. 3) The ticker is easy to obtain from news or search engines. Based on this research company ticker can be used to measure investor attention.

2.1.2 Consumer attention

By using SVI as a proxy for consumer attention, researchers build on the assumption that a high percentage of consumers use Google as a tool in relation to an event, such as buying a product or traveling to a new place. Research show that a high share of consumers use the internet as a source of information before buying (Ratchford, Talukdar, & Lee, 2001), and this share is especially high in regards to discretionary goods, where 70% of purchases are estimated to be influenced by online interactions (von Helversen, Abramczuk, Kopeć, & Nielek, 2018). We can expect Google to be the main source of this information due to their market share of 92% worldwide (Statcounter, 2021). In the use of SVI as a proxy for consumer attention, most research focus on forecasting other features than financial market data. Roy, Mittal, Basu, and Abraham (2015) state its usefulness in forecasting consumer behavior in the fashion industry, D. H. Park (2017) uses SVIs to predict tourism demand, Zhang (2017) presents its value in predicting consumer confidence, and Paturohman et al. (2018) find that Google Trends empowers the estimation of bank deposits. Other research state the value of using SVIs as a proxy for consumer attention to forecast sales of specific products. Significant results have been presented for the automobile industry (Wijnhoven & Plant, 2017), the food industry (Boone et al., 2017) and the housing market (Wu & Brynjolfsson, 2013).

In the creation of SVI as a measure of consumer attention, researchers vary between focusing on specific companies, by using company-specific keywords such as "Burberry" (Silva, Hassani, Madsen, & Gee, 2019), and focusing on an industry as a whole, by using industry-specific keywords such as "TV online" (Perju-Mitran, 2018). Promising results by both approaches indicate that combining them, by using both company-specific and industry-specific SVIs, has the potential to provide additional value.

2.2 Company segmentation and consumer behavior

The consumer sectors Consumer Discretionary and Consumer Staples have varying features and responses to market events. The economic cycle, which reflects the fluctuations of activity in an economy ("The asymmetric behavior and procyclical impact of asset correlations", 2011), can be a critical determinant of sector performance over the intermediate term. Consumer Discretionary tends to outperform in the early-cycle phase, characterized by lower interest rates and a sharp economic recovery. The first signs of economic recovery are associated with increased consumer confidence and increased borrowing (*OECD, Statistics and Data Directorate*, 2020), which benefits the Discretionary companies. On the contrary, Consumer Staples is more tied to basic needs, is less economically sensitive, and has a record of outperforming the broader market throughout the recession phases (Hoofwijk, 2020).

Another noticeable distinction is the relevance of brand equity. Comparing the two sectors, intangible assets account for a greater proportion of total assets to the Discretionary companies (Mizik, 2014). Mizik (2014) states that there are primarily three forms of intangible assets; intellectual, contracts and brands, where brands matter the most to consumer-facing industries. Brands are valuable because of their ability to maintain and create earnings for the firm over and above the earnings generated by tangible assets. As such, the consumers' perception of a brand, or brand equity, should manifest itself in the market value of the firm and thus have an impact on shareholder value. This impact is most substantial for Consumer Discretionary companies, as brand equity can account for 20-35% of market capitalization in this sector (""How much of intangible value does brand represent?"", n.d.).

2.3 Delayed financial response to information

Stock analysts forecast revenues and growth to project how future earnings will develop. Forecasts are important components of security analysis, often leading to a stock's future worth (McDonald, 2013). However, how many periods of growth are needed for the stock price to progress in one way or another is uncertain. The lag from the time information is presented, or an event occurs, until its effect materializes in stock returns, is a focus of research. Yoshinaga and Rocco (2020) find that lagged Google search volume is followed by changes in abnormal returns looking at

57 large Brazilian companies. "10.2307/2491062" (n.d.) find that a portion of the price response to information is delayed, and Lim, Hooy, et al. (2010) state that the size of this delay varies between companies. In essence, lagged financial response to information is frequent in the financial market, but it differs with regards to a number of factors. The potential for information to materialize in return after time, encourages researchers to include lagged models when investigating various variables' effect on financial performance. Focusing on consumer companies, we can expect a change in demand or attention towards a company's products to have a delayed reflection in the stock price, as it takes time for this information to be available to the public. Attention towards the stock itself can be expected to be reflected with a shorter time lag, as this attention often materializes in the trade of the stock. Therefore, including lagged models can be especially important when considering consumer attention, and Chen (2015) finds positive and statistically significant (at the 1% level) predictive ability of consumer confidence on stock returns lagged for 1 and 2 months, indicating a potential for further utilization.

This thesis contributes to the field of research by introducing consumer attention variables to the prediction of financial performance, particularly the performance of consumer companies, and employ distinct lagged models to evaluate when and how these variables influence.

3 Data

In this section we elaborate on the data sources and processing methods applied.

We use six explanatory variables in our analyses; three variables constructed from Google search volumes (one representing investor attention and two representing consumer attention) and three variables of market data (return, volatility and trading volume). Table 1 presents an overview and explanation of all variables.

Our data consists of weekly time series from the period Jan 2015 - Dec 2019 (5 years). We split the time series into training and test sets by the split 80%, 20%. As a result, data from 2015 to 2018 is used as training set and 2019 data as test set.

3.1 Google Trends Data

We collect search volume from the Google Trends webpage. Google Trends is a service by Google that offers users the ability to visualize the relative popularity of a keyword over time, as well as the opportunity to compare the popularity of one keyword with another (*FAQ about Google Trends data*, 2020). The data is not presented in absolute numbers; rather, it is scaled from 0 to 100, where 100 represents the maximum popularity during the chosen time period. Each data point is divided by the total searches of the geography and time range it represents to remove time effects. Worldwide search volumes are used, as most of the companies under research are international. The output from Google Trends is called the search volume index (SVI).

The three search volume indices (SVIs) created from Google Trends are:

1. **SVI^T (investor attention):** This SVI consists of searches for the company ticker, which is the keyword used to measure investor attention in previous research (e.g Da et al. (2011)). To exemplify, the company *Porsche*'s SVI^T is created from the keyword *"PSHG"*.

2. **SVI^B (brand consumer attention):** Brand-related keywords. The SVI is created from keywords that are brand names. To *Porsche*, the SVI^B variable is constructed from the keywords *"Porsche", "Volkswagen", "Audi", "SEAT", "SKODA", "Bentley", "Bugatti",* etc, which are all brands under the *Porsche* corporation.

3. **SVI^I (industry consumer attention):** Industry-related keywords. The industries and companies are classified according to the Global Industry Classification Standard (GICS). Each industry has its respective set of keywords. A car company, such as *Porsche*, is classified as a "Automobiles" company and will thus have keywords related to this industry as SVI^I. To companies within "Automobiles" the SVI^I variable is created from time series of keywords such as *"car", "auto parts", "tires"* etc.

3.1.1 Keyword selection

When measuring consumer and investor attention using Google search volume, the choice of keywords is important as it can be a key determinant for the results. It can also create unwanted bias, as it allows for subjective considerations. To reduce this bias, we follow the standard within literature for the investor attention SVI and use company ticker as keyword. For the consumer attention SVIs, there is no clear standard within literature to follow, and therefore, the SVIs are exposed to more subjective choices. To minimize this bias, we follow specific collection procedures in the selection of all keywords. These procedures are chosen to limit our influence. The following points summarize and describe the collection procedures for all SVIs.

1. **SVI^T:** Company ticker is collected from Refinitive's Company Data catalogue (*Refinitive Eikon Student License*, 2020). The tickers are adjusted by removing endings related to country or stock exchange. This is done to make the tickers more "search friendly". For instance, "PSHG.DE" is replaced with "PSHG".

2. **SVI^B:** For the brand-specific keywords we have manually researched each company and gathered their portfolio of brands. These are extracted from investor presentations, annual reports and company webpages. For instance, to *Prada* the company specific keywords are *"Prada"*, *"Miu Miu"*, *"Coach"* and more.

3. **SVI^I:** The industry-specific keywords are found using Google Trends. Google Trends offers a library of keyword categories where we find the categories corresponding to our industries. We extract the top ten keywords worldwide in each category over the training period and use these to construct the industry consumer attention SVI.

An overview of all industries (GICS), tickers (SVI^T), brand-related keywords (SVI^B) and industry-related keywords (SVI^I) is provided in the Appendix.

3.1.2 Search Volume Index generation

The time series for each keyword, k, are accessed through a Python Google Trends API (*pytrends*), and keywords with too little search volume are removed. We use global search volume data.

Figure 1 visualizes the procedure for generating SVIs. The SVI for investor attention (SVI^T) is created by only conducting step 1.A, 2 and 5, and the SVIs for consumer attention (SVI^B and SVI^I) are created by following all steps.

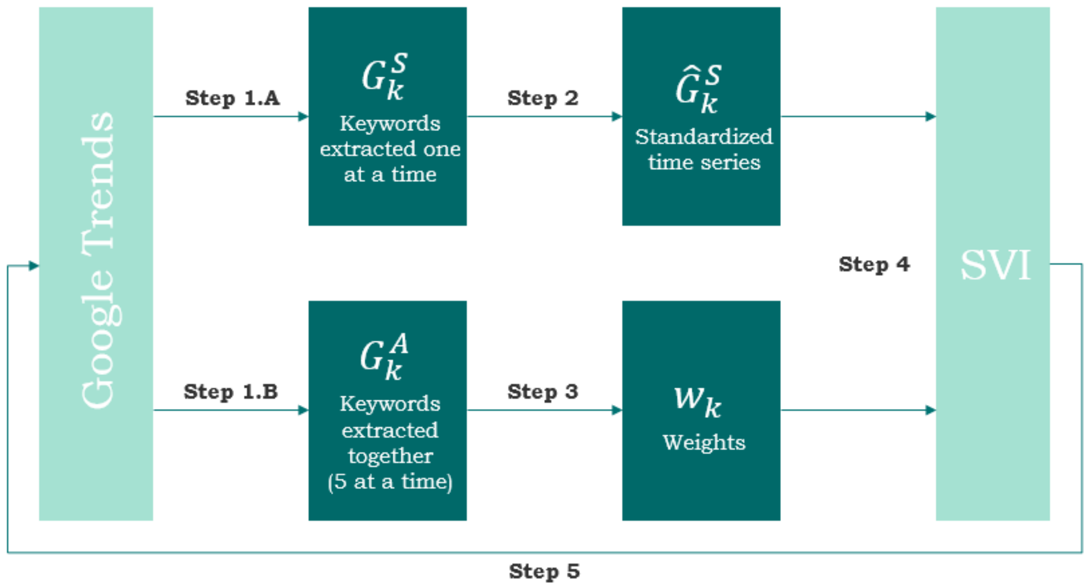


Figure 1: Visualization of steps conducted in the creation of each SVI

Step 1: Extract search volume time series of the keywords selected as specified in 3.1.1 from Google Trends. The keywords are extracted via pytrends in two ways:

1.A One at a time, to get search volumes normalized based on the single keyword - G^S

1.B Together, to get search volumes normalized related to the other keywords in the index - G^A

Step 2: Standardize the keyword time series that are extracted one at a time (G^S) by subtracting the mean and dividing by the standard deviation for each time series. Mean and standard deviation are calculated by using a rolling 1-year average, meaning we use time series data for the year up until a data point in the calculations.

Step 3: Calculate impact/weight of each keyword to their respective SVI. Because of limitations in the API, not allowing extraction of all keywords at the same time, the keywords are extracted in bulks of 5, and then normalized based on a reference keyword added to every extraction. This means that the time series for each keyword k is multiplied by the relative difference between the associated reference time series, A_k (extracted together with keyword k), and a "global" reference time series, A_0. This is done to obtain normalized time series where all volume measures are comparable across extractions. The following formula is used to calculate the normalized value of each keyword:

$$\hat{G^A_{kt}} = G^A_{kt} * \frac{A_{0t}}{A_{kt}} \tag{1}$$

Where
$\hat{G^A_{kt}}$ = normalized value of keyword k at time t
G^A_{kt} = value of keyword k at time t
A_{kt} = value of reference keyword extracted with keyword k at time t. Here we use the value of the first extraction (A_{0t}) as reference for all other.

The keywords within the SVI are weighted by the mean of their normalized search volume ($\hat{G^A_{kt}}$). The weight of keyword k is calculated as:

$$w_k = \frac{\overline{G^A_k}}{\sum_{k=0}^{K} \overline{G^A_k}} \tag{2}$$

where $\overline{G^A_k}$ is the mean of the normalized search volumes for keyword k, set to:

$$\overline{G^A_k} = \frac{1}{T} \sum_{t=0}^{T} \hat{G^A_{kt}} \tag{3}$$

Step 4: The final SVI is calculated by the equation:

$$SVI^X = \sum_{k=0}^{|X|} w_k G^S_k \tag{4}$$

Where
SVI^X = SVI being created. $X \in \{B, I\}$

$|X|$ = number of keywords k in X

Step 5: Repeat step 1-5 for all SVIs for all companies and industries

Each stock is connected to one SVI^B, one SVI^I and one SVI^T.

3.2 Market Data

Our set of companies is based on the constituents of two indices; S&P 500 Consumer Discretionary and S&P 500 Consumer Staples. Together, these indices represents all companies within S&P 500 that are consumer directed. The two sectors are again divided into more specific industries. Figure 2 presents an overview of the indices/ sectors and their respective industries.

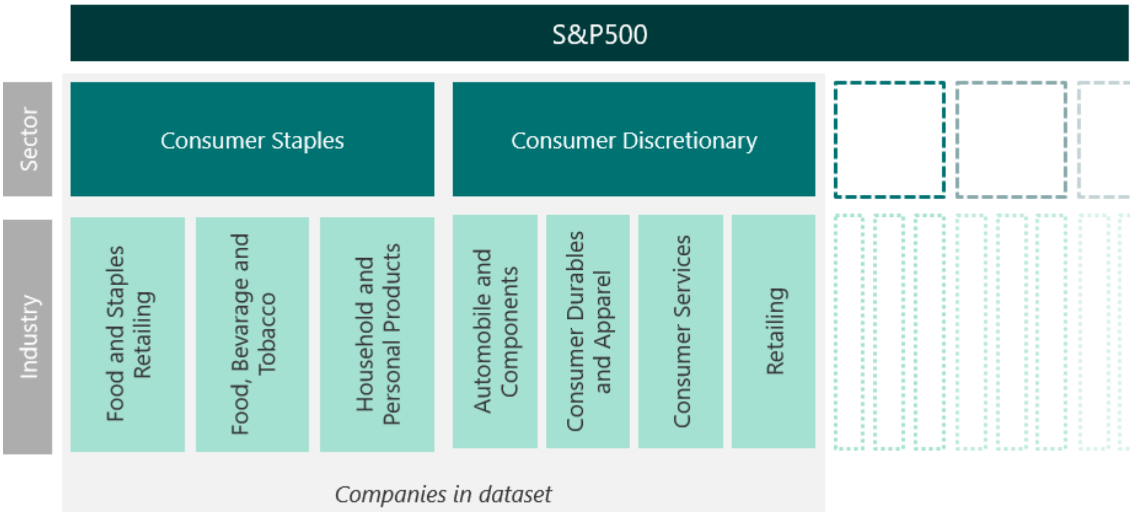


Figure 2: Overview of relevant sectors and industries in S&P 500

Companies that constitute one of the two indices and that meet the following criteria are selected for the dataset:

- The company must have been a constituent of the index for the whole sample period
- The company must have time series for the whole sample period without long periods with missing data

A complete list of the 131 remaining companies is given in Appendix A.2.

We have gathered market data for all companies included in the two indices. For each company, we have collected daily closing, open, high and low prices, in addition to daily trading volumes. This data is obtained from Refinitive's Company Data catalogue (*Refinitive Eikon Student License*, 2020).

3.2.1 Return

We use the below Equation 5 to calculate weekly returns:

$$r_t = 100 * log(\frac{close_t}{close_{t-1}}) \tag{5}$$

Where $close_t$ is the closing stock price of Monday in week t.

The use of Monday closing price is due to the way Google Trends is providing weekly data, where a week is defined as the query average from Monday to Sunday. We therefore apply the closing price of the upcoming trading day; Monday.

3.2.2 Abnormal Return

We want to examine the impact from attention measures on the otherwise unexplained returns. Therefore, the abnormal return calculated using the Fama-French model, is used as dependent variable. The choice of the Fama-French model for calculating the abnormal return is due to its ability to explain the expected returns of portfolios (Blanco, 2017; Kapur, 2007). The abnormal return is set to the α plus the error(ϵ_t) in Equation 6. This represents the part of the actual return that is not explained by market, size or value risk. By using abnormal return as dependent variable in our model, we test whether our Google and market variables can forecast stock price fluctuations, which are not explained otherwise. The Fama-French model expands on the capital asset pricing model (CAPM) by adding size risk and value risk factors to the market risk factor in CAPM. The model takes into account that value and small-cap stocks outperform markets on a regular basis (Fama & French, 1992). The Fama-French 3 Factors are extracted from Kenneth R. French's data library (French, 2020). The firm specific Fama-French beta coefficients are estimated from a rolling 1-year regression:

$$r_t - r_{ft} = \alpha + \beta_1 SMB_t + \beta_2 HML_t + \beta_3 (r_m - r_f)_t + \epsilon_t \tag{6}$$

And the abnormal return is set to be:

$$r_t^{abn} = r_t - r_{ft} - \beta_1(r_m - r_f)_t - \beta_2 SMB_t - \beta_3 HML_t \tag{7}$$

Where r is the return, r^{abn} is the abnormal return, Greek letters are regression coefficients, t is the week and SMB, HML, r_m and r_f are the Fama-French variables.

3.2.3 Volatility

Volatility is included as a parameter in the model, as a result of the relationship between volatility and future returns (Banerjee, Doran, & Peterson, 2007). We use the volatility estimator for daily volatility stated by Garman and Klass (1980) and suggested by Molnár (2010). We use the opening, close, high and low prices during a trading day to calculate the realized volatility for that day:

$$\sigma_d^2 = \frac{1}{2}(h_d - l_d)^2 - (2log(2) - 1)c_d^2 - j_d^2 \tag{8}$$

with:

$$c_d = log(close_d) - log(open_d), \tag{9}$$

$$l_d = log(low_d) - log(open_d) \tag{10}$$

$$h_d = log(high_d) - log(open_d) \tag{11}$$

$$j_d = log(open_d) - log(close_{d-1}) \tag{12}$$

Weekly variance is calculated as:

$$\sigma_t^2 = \sum_{d \in t} \sigma_d^2 \tag{13}$$

And weekly volatility is calculated as:

$$\sigma_t = \sqrt{\sigma_t^2} \tag{14}$$

Where t is the week, d is the day, $high_d$ and low_d are the highest and lowest realized price on a given day, and $open_d$ and $close_d$ are the opening and closing price on a given day.

3.2.4 Abnormal Trading Volume

Previous research finds evidence of a high-volume return premium (Barber & Odean, 2008), indicating that trading volume has an effect on future stock price fluctuations.

Therefore, weekly abnormal trading volume is included as a variable in the model. We calculate the abnormal trading volume as:

$$V_t = \frac{v_t - \frac{1}{52}\sum_{i=0}^{52} v_{t-i}}{SD_{v,t}} \tag{15}$$

Where V_t is abnormal trading volume, v_t is absolute trading volume and $SD_{v,t}$ is standard deviation of volume v for the year preceding week t.

3.3 Overview of variables

Variable	Symbol	Definition	Source
Brand Search Volume	SVI^B	Measure of consumer attention towards a company. The Search Volume Index is constructed from keywords that are brand names related to the company	Google Trends
Industry Search Volume	SVI^I	Measure of consumer attention towards an industry. The Search Volume Index is constructed from keywords related to the industry the company operates within	Google Trends
Ticker Search Volume	SVI^T	Measure of investor attention. The Search Volume Index is created by using the company's ticker as keyword	Google Trends
Volatility	σ	Volatility estimated using the estimator for daily volatility stated by Garman and Klass (1980)	
Return	r	Weekly actual stock return	Refinitive Datastream
Abnormal Return	r^{abn}	Weekly actual stock return minus the expected return from Fama French 3-factor model	Refinitive Datastream
Abnormal Trading Volume	V	Weekly actual trading volume subtracted by previous year's average, divided by previous year's standard deviation	Refinitive Datastream

Table 1: All the variables used in this paper. The first three are constructed from Google Search Data and the last four are financial variables.

3.4 Stationarity

To make the variables comparable we standardize all financial and attention variables, except returns. This is to simplify the interpretations of results. The time series are standardized by subtracting the mean and dividing by the standard deviation. Mean and standard deviation are calculated by using a rolling 1-year average, meaning we use time series data for the year up until a data point in the calculations. After standardization, we test for stationarity using a Fisher unit root test. The Fisher type test is using the Augmented Dickey-Fuller test and rejection of the null hypothesis indicates stationarity. The tests confirm stationarity for all variables.

3.5 Summary Statistics

Correlation coefficients between the variables can be seen in Table 2. We follow the same procedure as Da et al. (2011) when calculating the correlation. First, we calculate correlations individually for each company, and then we average the results across all companies. We do this for the time period from 2015 to 2019.

	r	σ	V	SVI^B	SVI^I	SVI^T
r	1	0.0333	-0.0194	0.1024	-0.1375	0.1231
σ		1	-0.2016	-0.0421	0.0228	0.0003
V			1	0.0810	0.0382	-0.1521
SVI^B				1	-0.1102	0.0081
SVI^I					1	0.2880
SVI^T						1

Table 2: Correlation matrix for the variables included in the dataset

The correlation matrix reported in Table 2 reveals small degrees of correlation among the proposed attention measures and the other variables. This indicates that there is no clear relationship and that the measures of attention can potentially provide additional information. The correlation matrix also shows that we have no issue with highly correlated variables.

The low correlation can indicate that the measures of consumer attention expose information reflecting trends, media pressure, sentiment or other factors related to consumer attention, which are not incorporated in the other variables. While the events causing changes in r, σ, V and SVI^T could be events such as earnings releases, company and industry news, recommendations from analysts, central bank announcements, interest rate changes or heard mentalities (Brooks, 2008), spikes in

SVI^B and SVI^I are perchance caused by social media hypes, campaigns, change in advertising exposures, and more. Consequently, the measures of attention could reflect indicators not captured by the other variables. The correlations between consumer attention variables (SVI^B, SVI^I) and the investor attention (SVI^T) are low, only 0.0081 and 0.2880. The low correlations indicate that people search for ticker and brands/industry related keywords with different motivations. We also note that SVI^B and SVI^I are negatively correlated

4 Methodology

In this section we present the methodology used. We explain the models and assumptions applied. For all models, we use the 2015-2018 training data to estimate the model and the 2019 test data to test and evaluate the performance.

4.1 Fama MacBeth regression model

Fama and MacBeth (1973) cross-sectional regressions are performed to evaluate the relationships between abnormal return and the independent variables. This is a two-step procedure. The first step involves estimation of one cross-sectional regression for each time period, and the second step involves calculating the average of the coefficients from the T cross-sectional regressions. The specific equations given in the following subsections are the cross-sectional regression specifications, while Equation (16) shows how the time-average is calculated for each regression to get the Fama-Macbeth coefficient estimates.

$$\hat{C^j} = \frac{1}{T}\sum_{t=1}^{T}\hat{C_t^j} \tag{16}$$

for j = #independent variables + 1

The choice of regression model follows that of Da et al. (2011), who find a significant relationship between investor attention (measured by Google search volumes) and stock returns using Fama MacBeth regressions.

4.1.1 Lagged abnormal return

To evaluate how the relationship between dependent and independent variables varies over time, lagged regressions are performed. Equation 17 represents the simple models only including search volume variables, Equation 18 includes only financial variables and Equation 19 include all variables. We include varying independent variables in the regressions to evaluate how and when attention measures affect abnormal return compared to financial variables. The models are conducted for all lags from one week to 52 weeks.

$$r_t^{abn} = C_t^0 + C_t^1 SVI_{t-u}^B + C_t^2 SVI_{t-u}^I + C_t^3 SVI_{t-u}^T \tag{17}$$

$$r_t^{abn} = C_t^0 + C_t^1 r_{t-u} + C_t^2 \sigma_{t-u} + C_t^3 V_{t-u} \tag{18}$$

$$r_t^{abn} = C_t^0 + C_t^1 r_{t-u} + C_t^2 \sigma_{t-u} + C_t^3 V_{t-u} + C_t^4 SVI_{t-u}^B + C_t^5 SVI_{t-u}^I + C_t^6 SVI_{t-u}^T \tag{19}$$

where t is the week and u is the time lag between dependent and independent variable.

4.1.2 Cumulative abnormal return

Regressions using cumulative abnormal return with varying time horizons are performed to evaluate whether the attention measures increase in value and significance when predicting over a longer time horizon than one week. We use weekly, monthly, quarterly and half year cumulative abnormal returns. Equation 20 represents the simple models only including search volume variables, Equation 21 includes only financial variables and Equation 22 include all variables.

$$r_t^T = C_t^0 + C_t^1 SVI_{t-1}^B + C_t^2 SVI_{t-1}^I + C_t^3 SVI_{t-1}^T \tag{20}$$

$$r_t^T = C_t^0 + C_t^1 r_{t-1} + C_t^2 \sigma_{t-1} + C_t^3 V_{t-1} \tag{21}$$

$$r_t^T = C_t^0 + C_t^1 r_{t-1} + C_t^2 \sigma_{t-1} + C_t^3 V_{t-1} + C_t^4 SVI_{t-1}^B + C_t^5 SVI_{t-1}^I + C_t^6 SVI_{t-1}^T \tag{22}$$

where t is the week and r^T the cumulative abnormal return for the upcoming week, month, quarter and half year ($T = W$ (weekly), M (monthly), Q (quarterly) and HY (half year)), given by:

$$r_t^W = \frac{close_{t+1} - close_t}{close_t} \tag{23}$$

$$r_t^M = \frac{close_{t+4} - close_t}{close_t} \tag{24}$$

$$r_t^Q = \frac{close_{t+12} - close_t}{close_t} \tag{25}$$

$$r_t^{HY} = \frac{close_{t+26} - close_t}{close_t} \tag{26}$$

4.2 Individual regression model

The Fama MacBeth regressions calculate one set of regression beta coefficients across time for all companies, before it then regress all stock returns for each T time periods against the previously estimated betas. This allows the regressions to take advantage of common relationships between the companies and it results in more data points being available in the estimation. On the other hand, the regressions can be unreliable when there are big differences in the relationship between independent and dependent variables, determined by the specific company. In addition, the Fama MacBeth model focuses on the cross-sectional relationships, while individual regressions focuses on an individual company over time. Therefore, we conduct individual regressions to evaluate how the possibility for individualization affects the predictive power of attention variables.

The individual regressions are conducted as ordinary least square regressions, and weekly, monthly, quarterly and half year cumulative abnormal return are used as dependent variables. Equation 27 represents the simple models only including the search volume variables, Equation 28 includes only financial variables and Equation 29 include all variables.

$$r_t^T = \beta_{0,t} + \beta_{1,t} SVI_{t-1}^B + \beta_{2,t} SVI_{t-1}^I + \beta_{3,t} SVI_{t-1}^T \tag{27}$$

$$r_t^T = \beta_{0,t} + \beta_{1,t} r_{t-1} + \beta_{2,t} \sigma_{t-1} + \beta_{3,t} V_{t-1} \tag{28}$$

$$r_t^T = \beta_{0,t} + \beta_{1,t} r_{t-1} + \beta_{2,t} \sigma_{t-1} + \beta_{3,t} V_{t-1} + \beta_{4,t} SVI_{t-1}^B + \beta_{5,t} SVI_{t-1}^I + \beta_{6,t} SVI_{t-1}^T \tag{29}$$

where t is the week, β's are regression coefficients and r^T the cumulative abnormal return for the upcoming week, month, quarter and half year ($T = W$ (weekly), M (monthly), Q (quarterly) and HY (half year)), given by Equations 23, 24, 25 and 26.

4.3 Trading strategy

To evaluate the potential for financial gains of using attention variables in forecasts, we create a simplified trading strategy using our selected companies from the S&P 500 Consumer Staples and S&P 500 Consumer Discretionary. The trading period is set to the test data (Jan. 2019 - Dec. 2019). We do a re-balancing each week, where the portfolio constituents are selected based on the predicted abnormal returns for the upcoming week. The predictions are made using rolling 1-year regressions, meaning we use data from the year up until the week in the regression model. Both Fama MacBeth regressions and individual regressions are used as prediction models, and the results are compared. The portfolio is created by buying stocks with predicted abnormal return above a threshold, X%, and shorting stocks with predicted abnormal return below -X%. Varying the thresholds also lets us evaluate the model's capability to predict normal versus extreme returns, and how the volatility develops accordingly. The Fama MacBeth and individual regression models are given by Equation 30 and 31, respectively.

$$r_t^{abn} = C_t^0 + C_t^1 r_{t-1} + C_t^2 \sigma_{t-1} + C_t^3 V_{t-1} + C_t^4 SVI_{t-1}^B + C_t^5 SVI_{t-1}^I + C_t^6 SVI_{t-1}^T \tag{30}$$

$$r_t^{abn} = \beta_{0,t} + \beta_{1,t} r_{t-1} + \beta_{2,t} \sigma_{t-1} + \beta_{3,t} V_{t-1} + \beta_{4,t} SVI_{t-1}^B + \beta_{5,t} SVI_{t-1}^I + \beta_{6,t} SVI_{t-1}^T \tag{31}$$

where t is the week, r^{abn} the weekly abnormal return, C^j's are the Fama MacBeth coefficient estimates given by Equation 16, and β's are regression coefficients for the individual regression model.

To measure the added value of our proposed attention measures, we compare the portfolios to portfolios following the same investment strategy, but which excludes the attention variables when making predictions. Thus, the benchmark portfolios only utilize the financial variables (r, σ and V). The Fama MacBeth and individual regression models used for comparison are given by Equation 32 and 33, respectively.

$$r_t^{abn} = C_t^0 + C_t^1 r_{t-1} + C_t^2 \sigma_{t-1} + C_t^3 V_{t-1} \tag{32}$$

$$r_t^{abn} = \beta_{0,t} + \beta_{1,t} r_{t-1} + \beta_{2,t} \sigma_{t-1} + \beta_{3,t} V_{t-1} \tag{33}$$

where t is the week, r^{abn} the weekly abnormal return, C^j's are the Fama MacBeth coefficient estimates given by Equation 16, and β's are regression coefficients for the individual regression model.

We evaluate the trading strategies both including and excluding trading costs. We assume equal weight between buys and shorts each week, and weight each constituent in the long and short part equally. The yearly accumulated abnormal return of the portfolios are used to evaluate whether the attention variables are improving the trading strategy and increasing economic benefit.

5 Results

In this section we present and discuss the results. We evaluate the relationship between abnormal return and attention. The model based on financial variables (past return, volatility and trading volume) is used as benchmark. We conduct regressions for all companies, which includes all stocks in the S&P 500 Consumer Discretionary and S&P 500 Consumer Staples indices, before the dataset is segmented into sectors and separate regressions are performed. First, we conduct regressions for varying time lags between dependent and independent variables to investigate time effects. We then study the relationship between weekly independent variables and longer-horizon cumulative abnormal returns (weekly, monthly, quarterly and half year abnormal return)

5.1 Prediction of lagged return

To measure the impact consumer and investor attention have on abnormal return compared to financial variables (r, σ and V), we conduct two separate regressions. The first regression (columns (1), Table 3) includes all variables (r, σ, V, SVI^B, SVI^I and SVI^T) and the second (columns (2)) includes only the attention variables (SVI^B, SVI^I and SVI^T). To evaluate when attention variables have the most impact, we conduct the two regressions over several time lags. This means varying the time distance between the independent and dependent variables. The results are presented in Table 3.

Companies: All Consumer Companies								
	r^{abn}_{t+1}		r^{abn}_{t+4}		r^{abn}_{t+12}		r^{abn}_{t+26}	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
r_t	0.1664***		-0.0030		-0.0019		0.0010	
	(0.0033)		(0.0024)		(0.0027)		(0.0028)	
σ_t	0.0104		-0.0110		0.0040		0.0064	
	(0.0170)		(0.0143)		(0.0145)		(0.0145)	
V_t	-0.0608**		-0.1183***		-0.0850**		-0.0893**	
	(0.0257)		(0.0344)		(0.0363)		(0.0409)	
SVI^B_t	0.0001	0.0000	0.0002	0.0002	0.0001	0.0001	0.0001	0.0002
	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
SVI^I_t	-0.0001	-0.0001	-0.0004	-0.0002	-0.0002	-0.0001	-0.0002	-0.0001
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
SVI^T_t	-0.0061	-0.0094	-0.0011	-0.0008	0.0017	0.0015	0.0001	0.0005
	(0.0080)	(0.0113)	(0.0110)	(0.0109)	(0.0135)	(0.0138)	(0.0115)	(0.0114)
R^2	0.1754	-0.0019	-0.0315	-0.0009	-0.0126	-0.0002	-0.0118	0.0000
#companies	131	131	131	131	131	131	131	131

Table 3: Fama MacBeth regression results for lagged abnormal return. Columns (1) display the results using weekly r, σ, V, SVI^B, SVI^I and SVI^T as independent variables, and columns (2) use only SVI^B, SVI^I and SVI^T. r^{abn}_{t+1}, r^{abn}_{t+4}, r^{abn}_{t+12} and r^{abn}_{t+26} represent the dependent variable. Standard errors are reported in parentheses. The symbols ***, ** and * denote significance at the 1%, 5%, and 10% levels, respectively.

The results show that none of the attention coefficients have values below a 10% significance level, and they are small compared to the coefficients of past return, volatility and trading volume. The total model (columns (1)) is able to predict the largest proportion of variance when using a 1 week lag between the dependent and independent variables. The search variables are, as expected, not able to predict the time lagged returns by themselves.

5.1.1 Impact of attention on different sectors

The foundation of the consumer attention variables is search volumes related to a specific company or industry. These are variables that have potential to gain significance when regressed on companies with similar features. On the basis of this assumption, the dataset is split into the two sectors Consumer Staples and Consumer Discretionary, and separate Fama MacBeth regressions are performed. This allows us to investigate how attention measures influence distinct types of companies. Table 4 and 5 present the regression results for Consumer Staples and Consumer Discretionary companies, respectively.

Companies: Consumer Staples								
	r^{abn}_{t+1}		r^{abn}_{t+4}		r^{abn}_{t+12}		r^{abn}_{t+26}	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
r_t	0.2287***		-0.0117**		-0.0040		-0.0020	
	(0.0048)		(0.0057)		(0.0062)		(0.0064)	
σ_t	0.0447*		-0.0254		-0.0303		0.0293	
	(0.0263)		(0.0299)		(0.0316)		(0.0325)	
V_t	-0.0642		-0.0742		-0.0383		-0.0897	
	(0.0454)		(0.0496)		(0.0585)		(0.0736)	
SVI^B_t	0.0009***	0.0002	0.0002	0.0001	-0.0003	-0.0002	-0.0001	0.0001
	(0.0003)	(0.0005)	(0.0005)	(0.0005)	(0.0006)	(0.0005)	(0.0006)	(0.0005)
SVI^I_t	0.0007	-0.0002	-0.0004	-0.0003	-0.0000	0.0001	-0.0001	-0.0002
	(0.0006)	(0.0007)	(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
SVI^T_t	-0.0123	-0.0173	0.0273	0.0216	0.0264	0.0252	0.0023	-0.0096
	(0.0135)	(0.0239)	(0.0219)	(0.0211)	(0.0232)	(0.0231)	(0.0232)	(0.0218)
R^2	0.1572	-0.0027	-0.0126	-0.0001	-0.0074	-0.0001	-0.0094	0.0001
#companies	38	38	38	38	38	38	38	38

Table 4: Fama MacBeth regression results for Consumer Staples companies on lagged abnormal return. Columns (1) display the results using weekly r, σ, V, SVI^B, SVI^I and SVI^T as independent variables, and columns (2) use only SVI^B, SVI^I and SVI^T. r^{abn}_{t+1}, r^{abn}_{t+4}, r^{abn}_{t+12} and r^{abn}_{t+26} represent the dependent variable. Standard errors are reported in parentheses. The symbols ***, ** and * denote significance at the 1%, 5%, and 10% levels, respectively.

Companies: Consumer Discretionary								
	r^{abn}_{t+1}		r^{abn}_{t+4}		r^{abn}_{t+12}		r^{abn}_{t+26}	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
r_t	0.1533***		-0.0026		-0.0001		0.0008	
	(0.0032)		(0.0025)		(0.0028)		(0.0029)	
σ_t	0.0088		-0.0090		0.0281		-0.0031	
	(0.0198)		(0.0199)		(0.0182)		(0.0192)	
V_t	-0.0667**		-0.1150***		-0.0910**		-0.0717	
	(0.0285)		(0.0412)		(0.0441)		(0.0465)	
SVI^B_t	0.0003	0.0000	0.0002	0.0002	0.0001	0.0001	0.0001	0.0002
	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
SVI^I_t	0.0001	0.0000	-0.0004	-0.0002	-0.0002	-0.0001	-0.0003	-0.0001
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
SVI^T_t	-0.0078	-0.0075	-0.0043	-0.0027	-0.0032	-0.0015	-0.0046	-0.0022
	(0.0096)	(0.014)	(0.0138)	(0.0137)	(0.016)	(0.0163)	(0.0148)	(0.0146)
R^2	0.1783	-0.0019	-0.0341	-0.0006	-0.0116	0.0001	-0.0090	0.0002
#companies	93	93	93	93	93	93	93	93

Table 5: Fama MacBeth regression results for Consumer Discretionary companies on lagged abnormal return. Columns (1) display the results using weekly r, σ, V, SVI^B, SVI^I and SVI^T as independent variables, and columns (2) use only SVI^B, SVI^I and SVI^T. r^{abn}_{t+1}, r^{abn}_{t+4}, r^{abn}_{t+12} and r^{abn}_{t+26} represent the dependent variable. Standard errors are reported in parentheses. The symbols ***, ** and * denote significance at the 1%, 5%, and 10% levels, respectively.

Neither Consumer Staples nor Consumer Discretionary present any significant coefficients for the attention variables. However, compared to using the full dataset of consumer companies (Table 3), separating into sectors slightly increases the attention variables' effect on abnormal return, as can be seen from the small improvements in R^2 values in columns (2). This indicates that separating the companies into smaller sets, where constituents have more similar features, can potentially increase the value of using attention variables. It is most evident for longer time lags.

5.1.2 Directional impact of attention over time

To analyze the directional impact of attention on abnormal return we plot the coefficients from the Fama MacBeth regressions for all lags from one week to one year. To smooth out the trend and reduce noise we apply a four week moving average. Three separate regressions are conducted, with varying companies included: all companies, Consumer Discretionary companies and Consumer Staples companies. Figure 3, 4 and 5 present the impact of brand attention, industry attention and investor attention, respectively.

Figure 3: Visualization of the SVI^B coefficient from the Fama MacBeth regression using r, σ, V, SVI^B, SVI^I and SVI^T as independent variables and abnormal return as dependent variable. The horizontal axis represent time lag from week of independent variables to week of dependent variable.

Figure 3 shows the brand attention coefficient to be stably positive for the Discretionary index until a time lag of 40 weeks. This indicates that an increase in search activity for brand related keywords will have a positive impact on the abnormal return for the upcoming 40 weeks for the Consumer Discretionary stocks. Regarding the Staples index, there is no clear pattern, but we note the negative spike between 40 and 50 weeks.

Figure 4: Visualization of the SVI^I coefficient from the Fama MacBeth regression using r, σ, V, SVI^B, SVI^I and SVI^T as independent variables and abnormal return as dependent variable. The horizontal axis represent time lag from week of independent variables to week of dependent variable.

Figure 4 shows that the industry attention variable coefficient tends to be negative for both indices over most time lags. This means that an increase in search volume for industry related keywords has negative impact on abnormal return, and that this applies to all types of consumer companies.

Figure 5: Visualization of the SVI^T coefficient from the Fama MacBeth regression using r, σ, V, SVI^B, SVI^I and SVI^T as independent variables and abnormal return as dependent variable. The horizontal axis represent time lag from week of independent variables to week of dependent variable.

Figure 5 presents the time lag effects of investor attention. The graphs representing the Discretionary and Staples indices are moving in opposite directions. This indicates that search activity for company tickers impacts Discretionary and Staples companies in opposite ways. The direction is dependent on time lag, and there is a tendency of Discretionary companies being negatively influenced in the short term (up to 20 weeks), before the effect turns positive in the long term. For Staples companies, the influence is opposite.

By varying time lag between attention variables and abnormal return we are able to analyze how the variables' impact varies over time, both in magnitude and direction. However, after running 52 regressions, only some of the coefficients were significant, and therefore the results are only able to show the general impact of attention on abnormal return, and the coefficients for specific companies may vary from those reported.

We note that brand attention usually has positive impact, industry attention negative impact and investor attention pulls in both directions, depending on time and sector. The observed behaviour is in line with financial theories. Da et al. (2011) discuss how retail investors are new buyers of stocks that receive attention, no matter if the attention is negative or positive. According to Da et al. (2011), retail investors only hold a small selection of stocks and do not short. Consequently, an attention shock will be followed by a net buying of stocks, creating an upward price pressure. This can explain how the brand attention variable is primarily positive for all lags. Behavioural financial theory states that Consumer Staples and Consumer Discretionary companies tend to do well over different parts of the economic cycle (*Investopedia*, 2021), and this could explain the behaviour in Figure 5 where attention variables have opposite effects for the two sectors. In turbulent times and signs of recession, investors tend to move towards slow and steady Consumer Staples stocks. Contrarily, when signs of economic recovery is presenting, investors are expecting a stock market recovery and higher growth in the Discretionary sector.

5.2 Prediction of cumulative return

The results show that attention variables have potential to add value in forecasting, especially over longer time lags. To further evaluate this importance, we extend our research to investigate the relationship between the independent variables and the upcoming weekly, monthly, quarterly and half year cumulative abnormal return.

By forecasting the cumulative return instead of lagged return, like in the previous section, we predict the upcoming trend instead of a specific point in time.

Companies: All consumer companies								
	Weekly return		Monthly return		Quarterly return		Half year return	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
r	0.1664***		0.0448***		0.0155***		0.0074***	
	(0.0033)		(0.0013)		(0.0007)		(0.0005)	
σ	0.0104		0.0013		0.0087**		0.0092***	
	(0.0170)		(0.0063)		(0.004)		(0.0028)	
V	-0.0608**		-0.1043***		-0.0985***		-0.0942***	
	(0.0257)		(0.0179)		(0.0159)		(0.0143)	
SVI^B	0.0001	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001**	0.0001**
	(0.0002)	(0.0003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SVI^I	-0.0001	-0.0001	-0.0003	-0.0002	-0.0002***	-0.0002**	-0.0002***	-0.0002***
	(0.0004)	(0.0004)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
SVI^T	-0.0061	-0.0094	-0.0083	-0.0119**	-0.0042	-0.0045	-0.0021	-0.0027
	(0.0080)	(0.0113)	(0.0052)	(0.0057)	(0.0030)	(0.0031)	(0.0022)	(0.0022)
R^2	0.1754	-0.0019	0.0799	-0.0008	0.0014	0.0001	0.0008	0.0001
#observations	27248	27248	26724	26724	25938	25938	24497	24497

Table 6: Fama MacBeth regression results for all companies on cumulative abnormal return. Columns (1) display the results using weekly r, σ, V, SVI^B, SVI^I and SVI^T as independent variables, and columns (2) use only SVI^B, SVI^I and SVI^T. Weekly, Monthly, Quarterly and Half year return represent the forecasting period of the cumulative abnormal return. Standard errors are reported in parentheses. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

First, this approach is applied to the set containing all consumer companies (Table 6). Compared to forecasting lagged returns (Table 3) there are several more significant values. The attention variables are now able to explain a larger portion of the variance, especially for the longer forecasting horizons, and their significance increases with the length of the horizon. The industry and investor attention coefficients are consistently negative, indicating that increased search volume will affect the forthcoming cumulative abnormal return in a negative direction.

5.2.1 Prediction of consumer sectors' cumulative return

Again, the dataset is split into two sectors (Consumer Staples and Consumer Discretionary), and separate Fama MacBeth regressions are performed. Table 7 and Table 8 display the coefficients and R^2 values.

Companies: Consumer Staples								
	Weekly return		Monthly return		Quarterly return		Half year return	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
r	0.2287***		0.0564***		0.0195***		0.0098***	
	(0.0048)		(0.0025)		(0.0015)		(0.0010)	
σ	0.0447*		-0.0099		-0.0106		-0.0133**	
	(0.0263)		(0.0144)		(0.0084)		(0.0057)	
V	-0.0642		-0.0763***		-0.0632***		-0.0483***	
	(0.0454)		(0.0222)		(0.0152)		(0.0111)	
SVI^B	0.0009***	0.0002	-0.0002	0.0002	-0.0000	0.0001	-0.0000	0.0001
	(0.0003)	(0.0005)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SVI^I	0.0007	-0.0002	-0.0002	-0.0002	-0.0002	-0.0001	-0.0002*	-0.0001
	(0.0006)	(0.0007)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0001)	(0.0001)
SVI^T	-0.0123	-0.0173	0.0050	0.0006	0.0081	0.0026	0.0041	0.0024
	(0.0135)	(0.0239)	(0.0095)	(0.0105)	(0.0060)	(0.0058)	(0.0039)	(0.0037)
R^2	0.1572	-0.0027	0.0949	-0.0006	0.0053	0.0001	0.0040	0.0002
#companies	38	38	38	38	38	38	38	38
#observations	7904	7904	7752	7752	7524	7524	7106	7106

Table 7: Fama MacBeth regression results for Consumer Staples companies on cumulative abnormal return. Columns (1) display the results using weekly r, σ, V, SVI^B, SVI^I and SVI^T as independent variables, and columns (2) use only SVI^B, SVI^I and SVI^T. Weekly, Monthly, Quarterly and Half year return represent the forecasting period of the cumulative abnormal return. Standard errors are reported in parentheses. The symbols ***, ** and * denote significance at the 1%, 5%, and 10% levels, respectively.

To Consumer Staples companies there are few significant values. Thus, the actual coefficient for specific companies can deviate from the reported coefficients. However, the proportion of variance explained when using only attention variables has increased, even if it is still very small. This implies that the attention measures alone can not explain the upcoming cumulative abnormal returns, but it increases the accuracy of the total model. For the half year horizon, the attention measures alone are able to explain 0.02% of the variance. This is 5% of what the application of all variables are capable of (0.4%).

Companies: Consumer Discretionary								
	Weekly return		Monthly return		Quarterly return		Half year return	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
r	0.1533***		0.0420***		0.0149***		0.0068***	
	(0.0032)		(0.0013)		(0.0008)		(0.0006)	
σ	0.0088		0.0033		0.0164***		0.0148***	
	(0.0198)		(0.0082)		(0.0056)		(0.0039)	
V	-0.0667**		-0.1115***		-0.1076***		-0.1006***	
	(0.0285)		(0.0217)		(0.0209)		(0.0173)	
SVI^B	0.0003	0.0000	0.0002	0.0001	0.0002*	0.0001	0.0002***	0.0002***
	(0.0002)	(0.0003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SVI^I	0.0001	0.0000	-0.0003*	-0.0001	-0.0002***	-0.0002**	-0.0003***	-0.0002***
	(0.0004)	(0.0004)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
SVI^T	0.0078	-0.0075	-0.0147**	-0.016**	-0.0082**	-0.0074*	-0.0045	-0.0032
	(0.0096)	(0.0140)	(0.0067)	(0.0071)	(0.0041)	(0.0040)	(0.0039)	(0.0029)
R^2	0.1783	-0.0019	0.0923	0.0001	0.0071	0.0002	0.0045	0.0003
#companies	93	93	93	93	93	93	93	93
#observations	19344	19344	18972	18972	18414	18414	17391	17391

Table 8: Fama MacBeth regression results for Consumer Discretionary companies on cumulative abnormal return. Columns (1) display the results using weekly r, σ, V, SVI^B, SVI^I and SVI^T as independent variables. and columns (2) use only SVI^B, SVI^I and SVI^T. Weekly, Monthly, Quarterly and Half year return represent the forecasting period of the cumulative abnormal return. Standard errors are reported in parentheses. The symbols ***, ** and * denote significance at the 1%, 5%, and 10% levels, respectively.

For Consumer Discretionary companies the attention measures' coefficients are significant for many of the reported time horizons. The statistically significant values demonstrate that the independent variable consistently affects the dependent variable. We note that brand and industry attention is significant for quarterly and half year return, while investor attention is significant for monthly and quarterly return. This indicates that the attention variables reflect the characteristics of the companies, and that this reflection has a shorter time horizon for investor attention than consumer attention (brand and industry). We note that the enhanced results compared to Consumer Staples could be a factor of more data points, as the number of observations are more than double of what we have for the Consumer Staples companies.

One would expect that attention measures have greater impact to non-essential products, and the results seem to support this preconception. This follows current research into how consumer behaviour differs in relation to the different sectors.

Consumer Staples companies sell products that people must, or think they must, buy, while Consumer Discretionary companies manufacture or sell products that aren't essential for living and are dependent on the disposable income of households. These unnecessary products, such as hotels, leisure activities, high-end apparel, entertainment and automobiles, are more often researched and/or purchased online (Okonkwo, 2009; Jackman & Naitram, 2015; S. Park, Lee, & Song, 2017), and discretionary companies are more often a subject of e-commerce than staples companies (Achille & Zipser, 2021). Therefore, sales of these companies' products seem better reflected by Google search volumes. The results from our research support these conclusions, amplifying the significant relationship between Consumer Discretionary companies and attention through online searches.

Investor attention has the closest relationship to returns for the shorter time lags, which is compatible with the nature of the variable. SVI^T primarily captures investor attention, more specifically the attention of retail investors (Da et al., 2011). We can expect the attention of investors to materialize in the buy or sell of a stock in short time, and therefore the relationship between investor attention and abnormal return should be strongest for short time horizons. In regards to the consumer attention variables, these could reveal a change in consumer patterns and hence a shift in a company's cash flow and returns. An increase of Google searches for a company name, such as Hilton Hotels, could be linked to increased web traffic on the company's website or connected to news traffic (Yi & Hwang, 2009). Website traffic would be generating increased revenues and potentially higher expectations of future earnings. For increased revenues to materialize in stock returns, the information needs to become available to the public. We could therefore expect a delay between consumer attention and returns, depending on how long it takes for company performance to be publicly available through financial results, company announcements or other sources of information.

In general, it seems apparent for both indices that attention variables gain significance when the forecast horizon lengthens. In comparison to forecasting all consumer companies as a whole, the enhanced results when splitting the dataset indicate that regressing on sets of companies where constituents have more similar features, increases the value of using attention variables.

5.2.2 Prediction of consumer industries' cumulative return

The results improve when directing the regression model towards sectors. Therefore, the dataset is narrowed further into industries. The dataset is split into seven industries according to Global Industry Classification Standard, that all have their own set of industry specific keywords. Separate Fama MacBeth regressions are conducted for each industry. The results are displayed in Table 9.

It is apparent that the results differ among industries. In general, investor attention has the highest coefficients, but few significant values except for the Food, Beverage & Tobacco industry. Industry attention is more significant overall, and compared to brand attention, it is more influential also in the absolute size of coefficients. The features of the respective SVI^Is seem to be more in line with the cumulative return of an industry, and this is to be expected as the industry attention variable is constructed from the same set of keywords for each regression. In addition, since the industry attention variable is constructed to capture relatively large fields of attention, which have high absolute search volumes, the variable appears more robust than the variable meant to capture the attention of specific companies. This means reduced noise and sampling bias. Compared to sector regressions, the R^2 values from industry regressions are generally higher. This substantiates that attention variables become more valuable when regressing on companies with similar features.

	Consumer Discretionary															
	Automobiles & Components				**Consumer Durables & Apparel**				**Consumer Services**				**Retailing**			
	Weekly	Monthly	Quarterly	Half Year	Weekly	Monthly	Quarterly	Half Year	Weekly	Monthly	Quarterly	Half Year	Weekly	Monthly	Quarterly	Half Year
r	0.1440***	0.0968***	0.0336***	0.0077***	0.1526***	0.0469***	0.0163***	0.0074***	0.1752***	0.0485***	0.0197***	0.0081***	0.1956***	0.0510***	0.0166***	0.0083***
	(0.0043)	(0.0044)	(0.0046)	(0.0009)	(0.0036)	(0.0019)	(0.0012)	(0.0008)	(0.0048)	(0.003)	(0.0018)	(0.0012)	(0.0045)	(0.0028)	(0.0018)	(0.0014)
σ	0.0692	-0.0199	-0.0218	-0.0242	0.0153	-0.0162	0.0124	0.0134**	0.0696*	-0.0045	-0.0070	-0.0078	0.0506	0.0278	0.0430***	0.0339***
	(0.0738)	(0.0300)	(0.0147)	(0.0169)	(0.0278)	(0.0152)	(0.0088)	(0.0064)	(0.0361)	(0.0222)	(0.0142)	(0.0100)	(0.0339)	(0.0245)	(0.0156)	(0.0105)
V	-0.0198	-0.0073	0.0272	0.0602*	-0.0400	-0.1145***	-0.1286***	-0.1264***	-0.0260	-0.1092***	-0.0987***	-0.0710***	-0.0840**	-0.1335***	-0.1193***	-0.1074***
	(0.0165)	(0.0057)	(0.0276)	(0.0163)	(0.0466)	(0.0361)	(0.0255)	(0.0195)	(0.0519)	(0.0341)	(0.0290)	(0.0240)	(0.0415)	(0.0365)	(0.0237)	(0.0193)
SVI^B	-0.0010	0.0003	0.0015	0.0016*	0.0004	0.0002	0.0002	0.0002**	-0.0008	-0.0003	-0.0003	-0.0002	0.0003	0.0002	0.0002	0.0002*
	(0.0041)	(0.0018)	(0.0009)	(0.0004)	(0.0003)	(0.0002)	(0.0001)	(0.0001)	(0.0005)	(0.0004)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0001)
SVI^I	0.0067	0.0018	-0.0015	-0.0023	0.0002	-0.0003	-0.0003**	-0.0003***	0.0008	0.0000	-0.0001	-0.0001	-0.0000	-0.0002	-0.0002	-0.0002
	(0.0077)	(0.0026)	(0.0011)	(0.0017)	(0.0004)	(0.0002)	(0.0001)	(0.0001)	(0.0005)	(0.0003)	(0.0002)	(0.0001)	(0.0006)	(0.0003)	(0.0002)	(0.0001)
SVI^T	0.0514	0.0285	-0.0596	-0.0013	0.0219	0.0024	-0.0041	-0.0126***	-0.0127	-0.0119	-0.0084	0.0062	-0.0222	-0.0102	-0.0091	-0.0034
	(0.0406)	(0.0135)	(0.0565)	(0.0910)	(0.0139)	(0.0100)	(0.0063)	(0.0038)	(0.0222)	(0.0157)	(0.0105)	(0.0060)	(0.0179)	(0.0145)	(0.0098)	(0.0055)
R^2	0.1695	0.1059	0.0029	0.0035	0.1879	0.0105	0.0053	0.0032	0.1471	0.0796	0.0253	0.0115	0.1781	0.0149	0.0043	0.0009
#companies	8	8	8	8	38	38	38	38	32	32	32	32	29	29	29	29

	Consumer Staples											
	Food & Staples Retailing				**Food, Beverage & Tobacco**				**Household & Personal Products**			
	Weekly	Monthly	Quarterly	Half Year	Weekly	Monthly	Quarterly	Half Year	Weekly	Monthly	Quarterly	Half Year
r	0.2074***	0.0469***	0.0129***	0.0071***	0.2387***	0.0606***	0.0196***	0.0095***	0.2036***	0.0576***	0.0104	0.0146***
	(0.0085)	(0.0050)	(0.0027)	(0.0017)	(0.0062)	(0.0032)	(0.0021)	(0.0013)	(0.0147)	(0.0115)	(0.0076)	(0.0050)
σ	-0.0623**	-0.0069	0.0033	0.0053	-0.0014	-0.0313	-0.0271**	-0.0202**	-0.1294	0.0290	-0.1236	-0.1015
	(0.0294)	(0.0171)	(0.0093)	(0.0059)	(0.0324)	(0.0190)	(0.0117)	(0.0080)	(0.1916)	(0.1173)	(0.1092)	(0.0735)
V	-0.0058	0.0202	0.0037	0.0179***	-0.0587	-0.0862***	-0.0847***	-0.0777***	-0.1176	-0.4318***	-0.3001***	-0.2155***
	(0.0289)	(0.0166)	(0.0088)	(0.0056)	(0.0467)	(0.0314)	(0.0189)	(0.0136)	(0.1225)	(0.1472)	(0.0916)	(0.0638)
SVI^B	-0.0002	0.0004	0.0005	0.0006***	-0.0012***	-0.0003	-0.0001	-0.0001	-0.0001	0.0015	0.0016	0.0010
	(0.0011)	(0.0006)	(0.0003)	(0.0002)	(0.0004)	(0.0003)	(0.0002)	(0.0001)	(0.0023)	(0.0014)	(0.0011)	(0.0008)
SVI^I	-0.0001	-0.0006	-0.0007**	-0.0007***	0.0006	-0.0002	-0.0003	-0.0003**	-0.0021	-0.0020	-0.0022*	-0.0014
	(0.0011)	(0.0006)	(0.0003)	(0.0002)	(0.0006)	(0.0003)	(0.0002)	(0.0001)	(0.0025)	(0.0016)	(0.0013)	(0.0009)
SVI^T	0.0421	-0.0461	-0.0123	0.0163	-0.0014	0.0234**	0.0174***	0.0123***	-0.1183	-0.1037	-0.0297	-0.0135
	(0.0513)	(0.0292)	(0.0154)	(0.0100)	(0.0168)	(0.0118)	(0.0065)	(0.0045)	(0.1137)	(0.0843)	(0.0648)	(0.0472)
R^2	0.1512	0.0132	0.0020	0.0003	0.0937	0.0122	0.0020	0.0012	0.1574	0.0143	0.0058	0.0020
#companies	5	5	5	5	27	27	27	27	5	5	5	5

Table 9: Fama MacBeth regression results for consumer industries on cumulative abnormal return. Columns (1) display the results using weekly r, σ, V, SVI^B, SVI^I and SVI^T as independent variables. and columns (2) use only SVI^B, SVI^I and SVI^T. Weekly, Monthly, Quarterly and Half year represent the forecasting period of the cumulative abnormal return. Standard errors are reported in parentheses. The symbols ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

5.2.3 Prediction of consumer companies' cumulative return

The regression improves considerably when the dataset is narrowed and separated into sets of companies with more similar features. This is evident in the transition from predicting all consumer companies to predicting Consumer Staples and Consumer Discretionary separately, and when the datasets are further decomposed into seven industries. Therefore, we continue to narrow the regressions, by considering each company individually. The Fama MacBeth regression model is replaced with separate linear regression models, conducted for each company. This allows us to explore how the relationship between the variables change when narrowing the regression. A company's specific response to attention fluctuations gains importance, and the prediction of it's stock performance becomes independent on the group average. Descriptive statistics from the regressions for weekly, monthly, quarterly and half year abnormal return are presented in Table 10.

		Fama MacBeth regression	Individual regression				
			Mean	Median	SD	Q 0.25	Q 0.75
Weekly abnormal return	r	0.1664	0.1906	0.1877	0.0520	0.1550	0.2148
	σ	0.0104	-0.0200	-0.0049	0.1128	-0.0836	0.0538
	V	-0.0608	0.0018	-0.0003	0.0794	-0.0482	0.0591
	SVI^B	0.0001	-0.0005	0.0004	0.0151	-0.0037	0.0041
	SVI^I	-0.0001	-0.0005	-0.0004	0.0051	-0.0025	0.0011
	SVI^T	-0.0061	0.0067	-0.0034	0.0914	-0.0612	0.0602
Monthly abnormal return	r	0.0448	0.0470	0.0460	0.0130	0.0378	0.0552
	σ	0.0013	0.0115	0.0175	0.0448	-0.0156	0.0463
	V	-0.1043	-0.0006	-0.0080	0.0504	-0.0323	0.0274
	SVI^B	0.0001	-0.0004	-0.0004	0.0153	-0.0052	0.0044
	SVI^I	-0.0003	-0.0003	-0.0002	0.0047	-0.0022	0.0016
	SVI^T	-0.0083	-0.0038	-0.0011	0.0882	-0.0681	0.0411
Quarterly abnormal return	r	0.0155	0.0141	0.0137	0.0055	0.0097	0.0170
	σ	0.0087	0.0036	0.0032	0.0310	-0.0128	0.0183
	V	-0.0985	0.0007	0.0067	0.0367	-0.0198	0.0236
	SVI^B	0.0001	0.0019	0.0013	0.0127	-0.0031	0.0056
	SVI^I	-0.0002	-0.0007	-0.0004	0.0042	-0.0027	0.0012
	SVI^T	-0.0045	-0.0033	-0.0075	0.0789	-0.0462	0.0402
Half year abnormal return	r	0.0074	0.0063	0.0062	0.0030	0.0042	0.0079
	σ	0.0092	0.0075	0.0057	0.0258	-0.0081	0.0193
	V	-0.0942	-0.0034	-0.0028	0.0290	-0.0184	0.0157
	SVI^B	0.0001	0.0020	0.0015	0.0180	-0.0022	0.0048
	SVI^I	-0.0002	-0.0010	-0.0008	0.0040	-0.0019	0.0007
	SVI^T	-0.0021	-0.0026	-0.0046	0.0681	-0.0457	0.0302

Table 10: Descriptive statistics of coefficients from conducting individual linear regressions for each company in the dataset. Coefficients from Fama MacBeth regression are included for comparison (Table 3).

The results show that both brand attention and industry attention have increased their impact compared to the Fama MacBeth regressions. This is not the case for investor attention. From the standard deviation we see that the investor attention coefficient have the highest variation, followed by brand attention, and then industry attention. This indicates that how brand and investor attention affects a company's performance is variable and dependent on individual differences between companies. Industry attention's effect, on the other hand, differs less between companies. This could be expected, as industry attention, by definition, is tailored to a larger group of companies, while brand and investor are company-specific.

Considering how the attention coefficients varies over time horizons, we note that while the impact of industry and investor attention remain steady/decrease in absolute size with the length of forecasting horizon, the impact of brand attention tends to increase. It is also apparent that industry attention has a negative effect on performance and brand attention yields a negative effect in the short term before it turns positive. Investor attention has positive influence in the short term, but turns negative as time horizon expands. This is in line with Da et al. (2011) which find an increase in searches for ticker keywords to predict higher stock prices in the next two weeks followed by an eventual price reversal within the year. It is also in line with established consumer journey frameworks, such as AIDA (awareness, interest, desire, action) (Barry & Howard, 1990). We can draw a parallel between increased search volumes and step two; interest. Two more steps, desire and action, needs to be undertaken for interest to materialize in earnings, which could be the reasoning why the impact of brand attention grows with forecast horizon.

Table 11 summarizes the average coefficient of determination (R^2) from conducting separate linear regressions for each company. Compared to previous R^2-values, adjusting the regressions to consider companies separately improves the predictive power. Individual regressions for each company tailors the coefficients to the specific company's stock performance movements, and allow for differences between companies. At the same time, it restricts the amount of data used in each regression, making the model subject to more noise. The results show that the possibility for individualization outweighs the decrease in performance from noise, as the individual regressions outperforms the Fama MacBeth regressions.

	Weekly return	Monthly return	Quarterly return	Half year return
all variables	0.1958	0.0999	0.0077	0.0049
r, σ, V	0.1956	0.0997	0.0071	0.0033
SVI^B, SVI^I, SVI^T	-0.0042	-0.0008	0.0011	0.0019

Table 11: Overview of mean R^2 from conducting individual linear regressions for each company with abnormal return as dependent variable and selected independent variables (stated in column 1).

Comparing the regressions including and excluding SVIs, it is clear that the attention variables increase the predictive power of the model for all tested cumulative returns. The added value increases as the time horizon of the forecast cumulative abnormal return increases. This indicates that it takes time for changes in attention towards a company to affect the company's financial performance. Past return, volatility and trading volume, on the other hand, affect performance in the short term, but quickly looses predictive power as the time horizon lengthens.

5.3 Robustness test: Alternative prediction model

It is of interest to test the Fama MacBeth regression model against another regression model to ensure that the results are due to the attention variables and not to the choice of model. We only test the Fama MacBeth regression model, and not the individual regression model, as this model has obvious substitutes. We conduct a panel data regression with fixed effects and compare the results to that of Fama MacBeth. We apply the robustness test to Table 6, meaning that we conduct panel regressions for the dataset containing all companies.

Companies: All Consumer Companies								
	Weekly return		Monthly return		Quarterly return		Half year return	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
r_t	0.1832***		-0.0023		-0.0011		0.0010	
	(0.0022)		(0.0021)		(0.0017)		(0.0028)	
σ_t	0.0112		-0.0010		0.0043		0.0059	
	(0.0152)		(0.0341)		(0.0124)		(0.0145)	
V_t	-0.0712**		-0.1238***		-0.0760**		-0.0872**	
	(0.0132)		(0.0255)		(0.0263)		(0.0509)	
SVI_t^B	0.0001	0.0000	0.0001	0.0002	0.0001	0.0001	0.0001	0.0001
	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
SVI_t^I	-0.0001	-0.0001	-0.0003	-0.0002	-0.0003	-0.0001	-0.0002	-0.0001
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
SVI_t^T	-0.0051	-0.0087	-0.0012	-0.0005	0.0009	0.0013	0.0001	0.0005
	(0.0060)	(0.0103)	(0.0120)	(0.0120)	(0.0135)	(0.0138)	(0.0115)	(0.0114)
R^2	0.1554	-0.0011	-0.0335	-0.0008	-0.0132	-0.0002	-0.0107	0.0000
#companies	131	131	131	131	131	131	131	131

Table 12: Panel regression results for all companies on cumulative abnormal return, using panel data regressions. Columns (1) display the results using weekly r, σ, V, SVI^B, SVI^I and SVI^T as independent variables, and columns (2) use only SVI^B, SVI^I and SVI^T. Weekly, Monthly, Quarterly and Half year return represent the forecasting period of the cumulative abnormal return. Standard errors are reported in parentheses. The symbols ***, ** and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 12 presents the results from the panel regression. The numbers are only slightly changed, meaning that our conclusions hold.

6 Trading strategy

In this chapter, we evaluate trading strategies to test the economic significance of our results. The test data from 2019 is used as trading period. We use rolling 1-year regressions to utilize data from one year to predict abnormal returns for the upcoming week. We test two regression models: Fama MacBeth regressions and individual linear regressions. The predictions of abnormal returns are used to construct an equally-weighted portfolio where we buy stocks with predicted return above a certain threshold, X%, and short stocks with predicted return below -X%. The portfolio is held for one week before re-balancing, meaning we trade at a weekly frequency.

The trading strategy return is calculated by Equation 34

$$Return_{portfolio,t} = \frac{Return_{long,t} - Return_{short,t}}{2} \tag{34}$$

6.1 Selecting trading threshold

We conduct the trading strategy several times where the long/short threshold of buying/shorting stocks is modified. The strategy is tested for thresholds from 0.1% to 3.0%, for Fama MacBeth regressions and individual regressions. A threshold of 0.1% means only trading companies with predicted returns above/below +/-0.1%.

Figure 6 shows how both regression models behave when the threshold limit is adjusted. For both regressions the return premium takes a concave pattern where the return reaches its apex at a threshold of around 1.2% for individual regression and 1.7% for Fama MacBeth regression, before it declines. When the threshold exceeds 1.2%, the Fama MacBeth regression model outperforms the individual regression model. At these higher thresholds, a smaller share of the companies are traded each week, making the portfolio less diverse and thus increasing volatility. This indicates predictions with variable reliability as the models become more unstable.

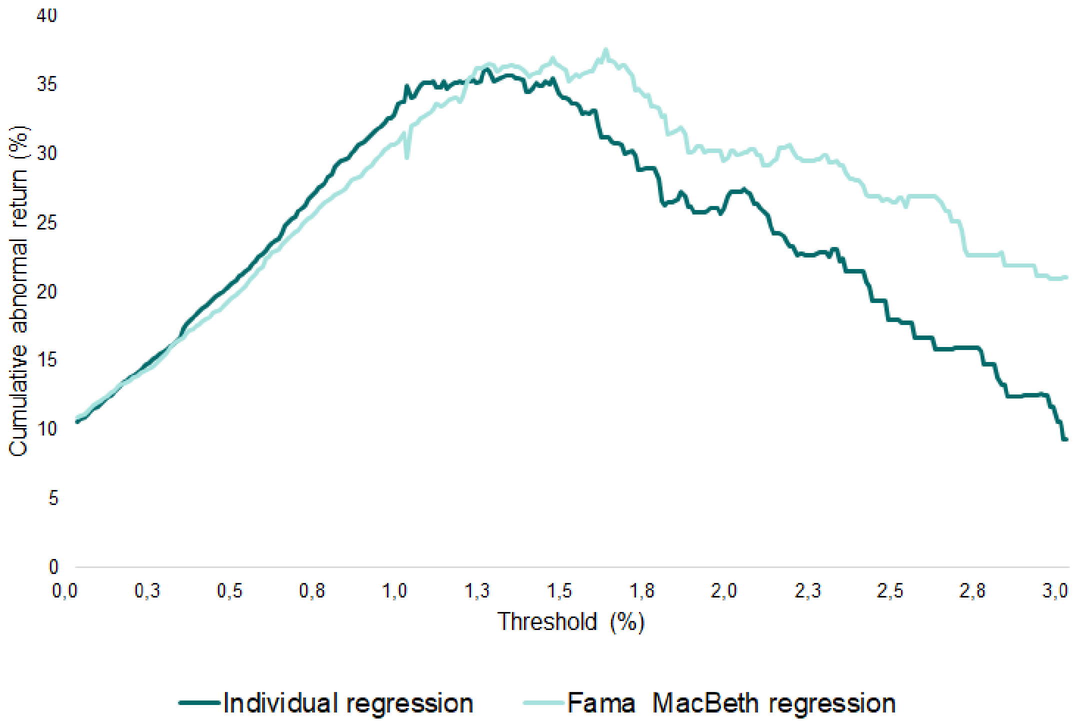


Figure 6: Yearly cumulative abnormal return for thresholds in the range from 0.0% to 3.0%. Return calculated for intervals of 0.01%. All variables (r, σ, V, SVI^B, SVI^I and SVI^T) are used as regressors.

The concave pattern shows how the choice of threshold is essential for the trading strategy's performance. At too low thresholds, the strategy encourages too many trades each week. This lowers the performance since the predicted returns have too high errors to be reliable so close to zero. For too high thresholds, there are not enough stocks with predicted returns above the given threshold, leading to multiple weeks without trade of stocks. Consequently, at high thresholds, the model does not make a substantial profit. It is though evident that at all thresholds in range, both models provide profitable strategies (accumulated return above 0%), but this is excluding trading costs.

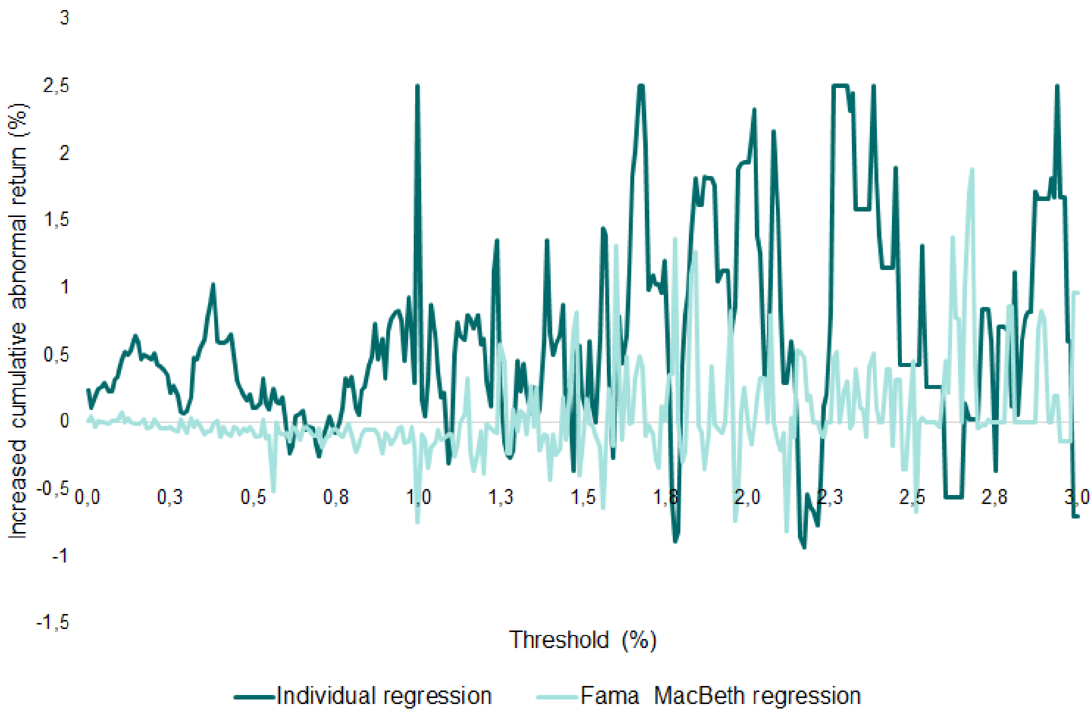


Figure 7: Yearly increase in cumulative abnormal return by including attention variables (SVI^B, SVI^I and SVI^T) in the regression model compared to only using financial variables (r, σ and V). Models conducted for thresholds in the range from 0.0% to 3.0%. Return calculated for intervals of 0.01%.

Figure 7 shows the yearly increase in cumulative abnormal return by including the attention variables over the range of trading thresholds. This is the difference between conducting the model using all variables and the financial variables only. It is evident that the two regression models react substantially different to the inclusion of attention. Individual regression has much higher variability and generally a larger positive effect compared to Fama MacBeth. This corresponds to the results in section 5.2.3, where the attention variables' influence increased when conducting individual regressions compared to Fama MacBeth. All SVI coefficients were higher in absolute size, and consequently, the inclusion/exclusion of companies as the threshold is adjusted results in higher spikes than for Fama MacBeth regressions. The contrast in size of attention coefficients can be seen from the "Weekly abnormal return" row in Table 10.

	Threshold (%)	Alpha	Volatility	Sharpe ratio
Individual regression	0.1	12.50%	0.67%	0.15
	0.3	16.39%	0.80%	0.18
	0.5	21.28%	1.13%	0.17
	0.7	26.69%	1.31%	0.19
	0.9	31.62%	1.54%	0.19
	1.0	34.94%	1.68%	0.20
	1.5	33.66%	3.64%	0.09
	2.0	27.28%	4.32%	0.06
	3.0	9.28%	2.84%	0.03
Fama MacBeth regression	0.1	12.67%	0.68%	0.16
	0.3	16.34%	0.75%	0.19
	0.5	20.18%	0.90%	0.20
	0.7	25.34%	1.34%	0.17
	0.9	29.51%	1.57%	0.17
	1.0	29.72%	1.57%	0.18
	1.5	35.50%	3.19%	0.11
	2.0	29.92%	4.22%	0.07
	3.0	21.08%	4.21%	0.05

Table 13: Alpha (abnormal return), volatility and Sharpe ratio for a trading strategy using r, σ, V, SVI^B, SVI^I and SVI^T as independent variables, buying stocks with predicted return higher than a given threshold (X%), and shorting stocks with predicted return below -X%

Table 13 presents performance results for selected thresholds, using both the individual regression model and the Fama MacBeth regression model. In addition to abnormal return, the table shows the volatility and corresponding Sharpe ratio for the selected thresholds. This illustrates how the risk varies based on threshold, and allows us to evaluate the risk/return compensation. The Sharpe ratio is set to cumulative abnormal return earned in excess of the risk-free rate per unit of volatility. The average 10-year U.S treasury yield for 2019 (2.14%) is used as risk-free rate (Macrotrends, 2021). The table shows that the Sharpe ratio follows a concave pattern, and has its peak at 1.0% for individual regressions and 0.5% for Fama MacBeth regressions. This indicates that even if the Fama MacBeth regressions have a cumulative abnormal return curve with a peak shifted towards higher thresholds than individual regressions (Figure 6), these high alphas bring high risk(volatility).

It is thus evident that the optimal threshold for a trading strategy can vary regarding return and risk preferences. In this thesis, we assume the portfolio to be part of a bigger, more diverse portfolio, where risk/return compensation is optimized through diversity, meaning we regard the alpha as the most central determinant for this purpose. A portfolio based only on Consumer Discretionary and Consumer Staples companies should be regarded as relatively risky independent of threshold choice. In the following sections, we therefore use alpha as determinant for the choice of threshold.

6.2 Comparing trading performance by sector

As seen in section 5, the regression coefficients vary among sectors and industries. We compare the trading strategy for both sectors, using both Fama MacBeth regressions and individual regressions to examine the performance.

The threshold, X, is chosen specifically for each week by running the trading model for the previous one year, testing each threshold in the range from 0.0% to 3.0%. The threshold that corresponds to the highest accumulated abnormal return for that previous year is used to re-balance for the upcoming week. We apply this threshold strategy to make the trading simulation realistic by only using past data. It also allows the model to dynamically change its threshold and thereby its exposure to risk and trading costs, as it changes the share of companies that is traded each week. The procedure for selecting threshold for each week is presented in Algorithm 1.

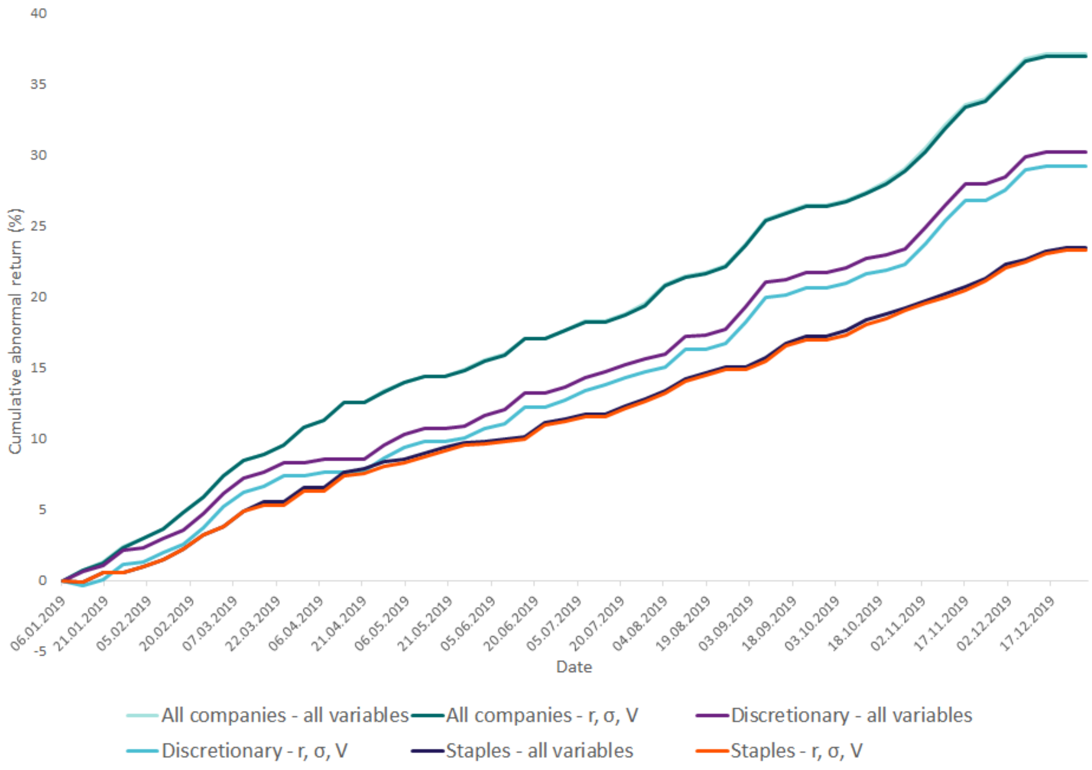
Algorithm 1: Threshold selection procedure

```
Input: week t
Result: optimal threshold X for week t
best_alpha = -inf
best_threshold = None
/* loop through all thresholds in range                                  */
for threshold in range(start : 0.1, end : 3.0, step : 0.1) do
    /* loop through all weeks in previous year                           */
    for week in [t-52,...,t-1] do
        /* predict weekly returns using past year data [week-52, week-1] */
        predictions = predict(week)
        actual = actual_abnormal_returns(week)
        /* trade stocks according to threshold                           */
        short = stocks_predicted_below(predictions, actual, threshold)
        long = stocks_predicted_above(predictions, actual, threshold)
        weekly_abnormal_return = trade(short, long)
        alpha = accumulate(alpha, weekly_abnormal_return)
    /* check if new alpha is higher than the currently best              */
    if best_alpha ≤ alpha then
        best_alpha = alpha
        best_threshold = threshold
return best_threshold
```

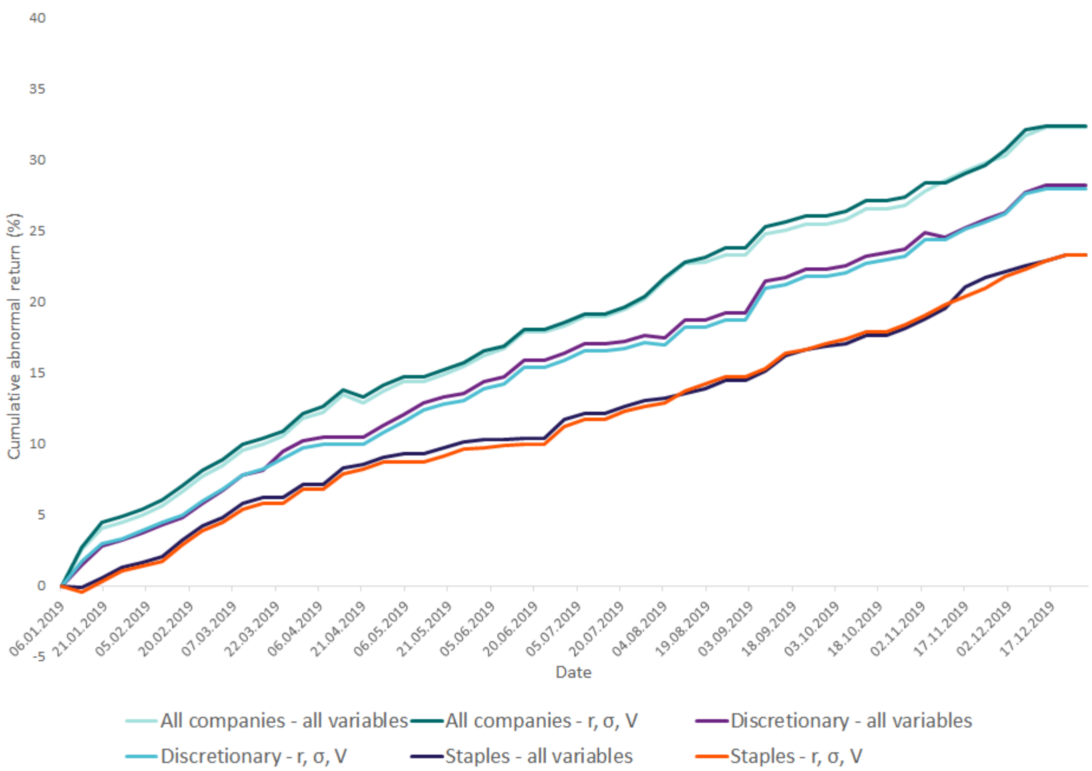
Table 14 presents the results from conducting the trading strategies for various sets of companies. We note that individual regressions generate higher returns than Fama MacBeth regressions. In addition, the inclusion of attention measures improves the performance of individual regressions by up to 1%. The increase is largest for Consumer Discretionary companies. For Fama MacBeth regressions, the inclusion of attention measures only improves the yearly cumulative abnormal return marginally.

	All companies		Consumer Discretionary		Consumer Staples	
	Individual	Fama MacBeth	Individual	Fama MacBeth	Individual	Fama MacBeth
All variables	37.12%	32.81%	30.20%	28.19%	23.43%	23.25%
r, σ, V	36.91%	32.40%	29.24%	27.94%	23.26%	23.25%

Table 14: Yearly cumulative abnormal return for conducting the trading strategy using selected independent variables (stated in column 1).



(a) Individual regressions



(b) Fama MacBeth regressions

Figure 8: Accumulated abnormal return, with and without attention variables, for different segments of companies. The portfolios are updated on a weekly frequency. Trading costs are excluded.

Figure 8a and 8b present the cumulative abnormal return achieved over time by applying the trading strategy using individual regressions and Fama MacBeth regressions, respectively. The strategy is applied to only Consumer Discretionary companies, only Consumer Staples companies, and All companies.

For individual regressions, including attention variables improves the underlying prediction model when considering the final yearly accumulated returns, regardless of segmentation of companies. However, Consumer Discretionary has the largest final gap between yearly return with and without attention measures, meaning this is the set of companies where attention measures contribute the most. This demonstrates that attention could be a relevant indicator of future return in this sector, and that the market has not fully incorporated it into its expectations. The accumulated yearly return of 2019 yields 30.20% including attention variables, and 29.24% excluded. Attention measures thus have the potential to increase the yearly return by around 1%. The same conclusions can not be transferred with same certainty to Consumer Staples and All companies, where the increase by including attention variables is smaller. The portfolio including All companies has the highest accumulated return (37.12%), followed by Consumer Discretionary and Consumer Staples.

The Fama MacBeth regression model also results in high cumulative abnormal returns, but these are slightly lower than for the individual regressions. There is also no clear increase in accumulated returns by including attention variables, indicating that these do not provide value when one set of regression coefficients are used on all companies. For Fama MacBeth regressions, as for individual regressions, the model performs the best when all companies are included, followed by Consumer Discretionary and then Consumer Staples.

6.3 Trading costs

Regarding whether the proposed strategy will yield payoff in the real world, we take trading costs into account. These consist mainly of transaction fee, bid-ask spread and market impact. Market impact is irrelevant assuming the trades to be too small to influence the market, and will hence be ignored. We estimate the transaction fees by observing the fees of online brokers. *Interactive Brokers* (2020) offer a fixed fee of $0.005 per share, while Speed Trader offers a flat fee of $2.95 per trade, independent of number of shares. By assuming a minimum of 100 shares per trade, the average cost will be no more than \$0.0295 per share for both brokers. The average share price for the stocks in our dataset is \$151 over the last 52 weeks, resulting in an average transaction fee of 0.02%. This is in line with similar papers (Bijl et al., 2015; Karlsen & Hesla, 2019)

Since our dataset does not contain efficient bid-ask spread, we rely on estimates from the literature. Ball and Chordia (2001) examine true spreads in large and mid cap companies, and report a quoted spread of 0.2%. Norges Bank Investment Management estimates indirect costs of 0.154%, covering transaction, spread and market impact costs. We build on their results and apply a total trading cost of 0.2%, covering spread and transaction costs. This is a rather conservative estimate. Table 15 shows whether the trading strategies are still profitable after accounting for trading costs.

		Individual regression		Fama MacBeth regression		
Set of companies	Regressors	Without tr.costs	With tr.costs	Without tr.costs	With tr.costs	S&P Index
All companies	All variables	37.12%	33.11%	32.81%	26.30%	
	r, σ, V	36.91%	32.89%	32.40%	24.55%	
Consumer Discretionary	All variables	30.20%	26.35%	28.20%	24.31%	20.7%
	r, σ, V	29.24%	25.39%	27.94%	23.20%	
Consumer Staples	All variables	23.43%	19.43%	23.25%	18.21%	23.1%
	r, σ, V	23.26%	19.30%	23.25%	18.17%	

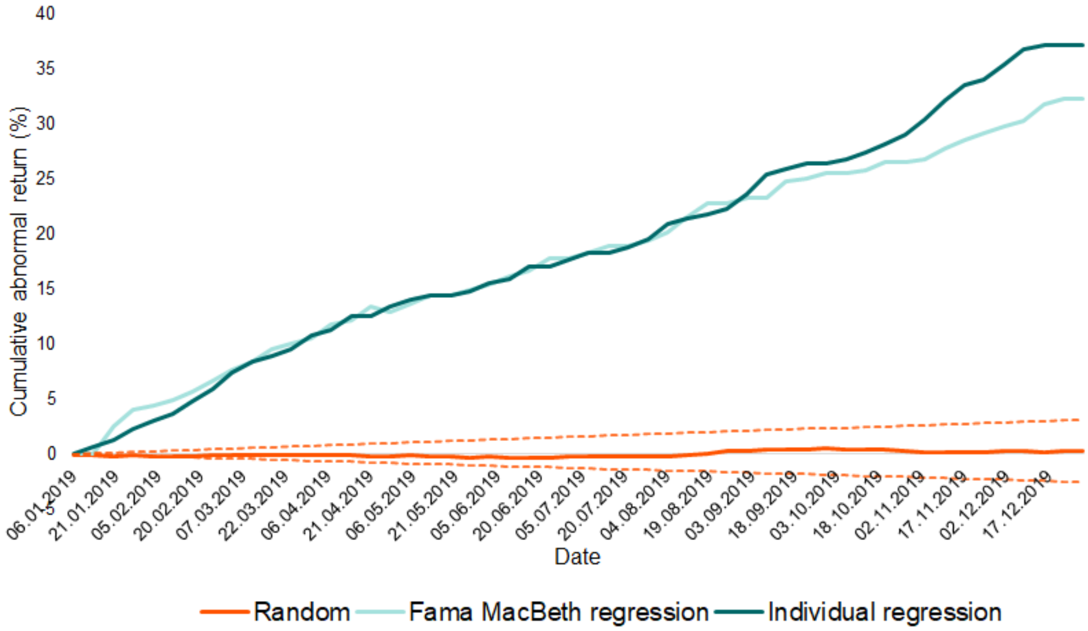
Table 15: Cumulative abnormal return (%) for conducting the trading strategy for different sets of companies (stated in column 1) using selected independent variables (stated in column 2) with and without transaction costs.

As seen in Table 15 the proposed models reach positive accumulated yearly abnormal returns after trading costs are considered. We can observe that trading costs decrease the performance by approximately 4% p.a, but the trading strategies still remain highly profitable.

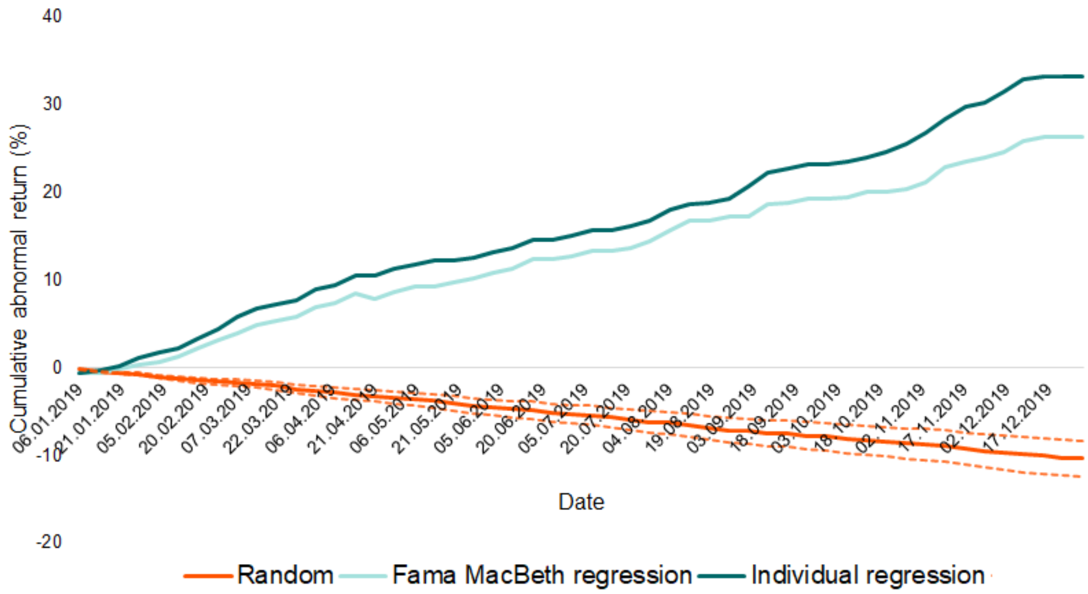
The 2019 actual annual returns of S&P 500 Consumer Discretionary and S&P 500 Consumer Staples are included for comparison. We highlight that the trading strategy results are cumulative abnormal returns and not cumulative actual returns, hence they are not directly comparable. The cumulative abnormal returns are used to evaluate the models' ability to predict the deviation between actual returns and the expected return as calculated in section 4. However, the comparison provides an indication of the models potential for economic profit. It is clear that the models using Consumer Discretionary companies have the best performance compared to the index when including trading costs.

6.4 Robustness test: Trading random 10% of companies

In order to test the robustness of our trading model we compare our results to the results of a randomized trading strategy. 10% of the stocks are randomly selected and purchased, and 10% are randomly selected and shorted. The portfolio return is calculated as in Equation 34. The 10% corresponds to the average percentage of portfolio traded each week when running our proposed model on individual companies with trading costs. 1000 iterations of the strategy is performed, and the mean cumulative abnormal return obtained over time is presented in Figure 9. In addition, the confidence interval for a 99% probability limit (confidence level) is plotted to visualize the variance in the iterations. These intervals measure the degree of uncertainty in the sample. We test for robustness both including and excluding transaction costs.



(a) Excluding trading costs



(b) Including trading costs

Figure 9: Aggregated abnormal return over time, with and without trading costs, for trading strategies using the Fama MacBeth regression model, the individual regression model and the model choosing random 10% of companies to trade in each re-balancing. Dotted lines represent the 99% confidence intervals of the random trading strategy. The portfolios are updated on a weekly frequency.

As seen in Figure 9, our suggested model outperforms the randomized model. The randomized model should be equally weighted and the accumulated return of 0%

could therefore be expected, as seen in Figure 9a. Similarly, we see a negative return when trading costs are considered (Figure 9b). Further adjustments to the trading model could also be evaluated, but we regard this as one of the most important. It seems clear from this test that our conclusions hold and are a result of the impact investor and consumer attention have on abnormal stock return.

7 Conclusion

The relationship between stock market performance and Google search data has been a subject of interest among researchers, hedge fund managers, brokers and more. Search volumes for company tickers have been used as measurement of investor attention and applied to financial forecasting. However, consumer attention could be an important indicator of the future prospects of a company, and has potential to be useful in forecasting performance. We therefore re-investigate the topic from a new perspective by including a measure of consumer attention utilizing Google searches for carefully selected keywords (such as brands) in addition to the measure of investor attention. Since consumer related stocks are driven by expected future earnings, which have potential to be reflected by consumer attention, we consider companies included in the S&P 500 Consumer Discretionary and S&P 500 Consumer Staples Index.

We study whether attention variables can predict future abnormal returns. The analysis finds that it requires a relatively long forecasting horizon for attention to materialize and influence stock performance. We also find that by segmenting the companies by similar features, our proposed set of attention measures achieves stronger predictive power. Furthermore, the results show that Discretionary companies have a stronger relationship to consumer attention relative to Staples companies.

To test the economic significance of our results we conduct a trading strategy. We construct a portfolio of consumer related stocks, and re-balance it weekly based on predictions from the suggested model. The yearly accumulated abnormal return is improved when our proposed attention measures are included as prediction variables. We also take trading costs into account, and find the portfolio to still yield a positive return.

Further research

This paper illustrates the usefulness of including attention in forecasting financial performance. Some of the choices and assumptions can be evaluated further. Fama French factors still show a strong long-term performance, but during the last decades, various other factor models have received increased interest. To state whether or not this has an impact on the results, other ways of calculating abnormal returns could therefore be assessed. We also note that our dataset only covers five years and is somehow restricted. Further research is needed to evaluate whether the results can be generalized over time. With longer time series the attention variables can also be utilized over longer time horizons, which has the potential to provide additional value as their impact on abnormal return increases with time.

In addition to evaluation of the assumptions used in our thesis, we have several suggestions for further research into this field. Firstly, a similar study should be conducted for other sectors, evaluating the potential for generalization. Secondly, SVI as a measure of consumer attention can potentially be utilized in other financial applications, as an addition to using only investor attention. It might also be possible to create even more relevant measures of attention, for example by applying a more systematic way of selecting the keywords or by combining SVIs with other measures of attention, such as sentiment analysis, news counts, views and likes.

References

(n.d.).

Achille, A., & Zipser, D. (2021). A perspective for the luxury-goods industry during—and after—coronavirus. *McKinsey Digital*.

Aiolfi, M., & Favero, C. A. (2005). Model uncertainty, thick modelling and the predictability of stock returns. *Journal of Forecasting*, *24*(4), 233-254.

Ariff, M., Loh, A., & Chew, P. (1997). The impact of accounting earnings disclosures on stock prices in Singapore. *Asia Pacific Journal of Management*, *14*, 17-27.

The asymmetric behavior and procyclical impact of asset correlations. (2011). *Journal of Banking Finance*, *35*(10), 2559-2568.

Audrino, F., Sigrist, F., & Ballinari, D. (2020). The impact of sentiment and attention measures on stock market volatility. *International Journal of Forecasting*, *36*(2), 334 - 357.

Ball, C., & Chordia, T. (2001, 10). True spreads and equilibrium prices. *The Journal of Finance*, *56*, 1801 - 1835.

Banerjee, P. S., Doran, J., & Peterson, D. R. (2007). Implied volatility and future portfolio returns. *Journal of Banking Finance*, *31*(10), 3183-3199.

Barber, B., & 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), 785-818.

Barry, T. E., & Howard, D. J. (1990). A review and critique of the hierarchy of effects in advertising. *International Journal of Advertising*, *9*(2), 121-135.

Bijl, L. R., Kringhaug, G., & Sandvik, E. (2015). Predictive power of Google search volume on stock returns..

Blanco, B. (2017). The use of capm and fama and french three factor model.

Boone, T., Ganeshan, R., Hicks, R. L., & Sanders, N. R. (2017, 12). Can Google Trends improve your sales forecast?

Brooks, C. (2008). *Introductory econometrics for finance.* Cambridge, UK: Cambridge University Press.

Challet, D., & Bel Hadj Ayed, A. (2013, 07). Predicting financial markets with Google Trends and not so random keywords. *SSRN Electronic Journal*.

Chen, M.-H. (2015). Understanding the impact of changes in consumer confidence on hotel stock performance in taiwan. *International Journal of Hospitality Management*, *50*, 55-65.

Cooper, M., Gutierrez, J., & Marcum, B. (2005, 02). On the predictability of stock returns in real time. *The Journal of Business*, *78*, 469-500.

Coyne, S., Madiraju, P., & Coelho, J. (2017). Forecasting stock prices using social media analysis. In (p. 1031-1038).

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

Demirer, R., Pierdzioch, C., & Zhang, H. (2017, 08). On the short-term predictability of stock returns: A quantile boosting approach. *Finance Research Letters*, *22*, 35-41.

Ding, R., & Hou, W. (2015). Retail investor attention and stock liquidity. *Journal of International Financial Markets, Institutions and Money*, *37*, 12 - 26.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427–465.

Faq about google trends data. (2020). https://support.google.com/trends/answer/4365533?hl=en. (Accessed: 2020-10-15)

French, K. R. (2020). *French database.* https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. (Accessed: 2020-10-30)

Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. In (Vol. 53, p. 67-78).

Gidófalvi, G. (2004, 12). Using news articles to predict stock price movements.

Hamid, A., & Heiden, M. (2015). Forecasting volatility with empirical similarity and Google Trends. *Journal of Economic Behavior Organization*, *117*, 62 - 81.

Heyman, D., Lescrauwaet, M., & Stieperaere, H. (2019). Investor attention and short-term return reversals. *Finance Research Letters*, *29*, 1 - 6.

Hoofwijk, M. (2020). Sector performance during the financial crisis of 2008.

"how much of intangible value does brand represent?". (n.d.).

Interactive brokers. (2020). https://www.interactivebrokers.com/en/index.php?f=1590&p=stocks1. (Accessed: 2021-03-05)

Investopedia. (2021). https://www.investopedia.com/terms/c/consumer-discretionary.asp. (Accessed: 2021-05-29)

Jackman, M., & Naitram, S. (2015). Research note: Nowcasting tourist arrivals in barbados – just google it! *Tourism Economics*, *21*(6), 1309-1313.

Juan Piñeiro-Chousa, M. L.-C., & Ribeiro-Soriano, D. (2020). Does investor attention influence water companies' stock returns? *Technological Forecasting and Social Change*, *158*, 120115.

Kapur, V. (2007). Is the fama and french model a good indicator of market sectoral performance?

Karlsen, M. R., & Hesla, H. S. (2019). *Predicting stock returns with google searches:*

One size does not fit all. NTNU.

Kim, N., Lučivjanská, K., Molnár, P., & Villa, R. (2019). Google searches and stock market activity: Evidence from Norway. *Finance Research Letters*, *28*, 208 - 220.

Lim, K.-p., Hooy, C.-w., et al. (2010). The delay of stock price adjustment to information: A country-level analysis. *Economics Bulletin*, *30*(2), 1609–1616.

Macrotrends. (2021). *10 year treasury rate.* (Accessed: 2021-05-25)

McDonald, R. (2013). *Derivatives markets.* Pearson.

Mizik, N. (2014). Assessing the total financial performance impact of brand equity with limited time-series data. *Journal of Marketing Research*, *51*(6), 691-706.

Molnár, P. (2010, 09). Properties of range-based volatility estimators. *International Review of Financial Analysis*, *23*.

Nikfarjam, A., Emadzadeh, E., & Muthaiyah, S. (2010, 03). Text mining approaches for stock market prediction. *2010 The 2nd International Conference on Computer and Automation Engineering, ICCAE 2010*, *4*, 256 - 260.

Oecd, statistics and data directorate. (2020). `https://www.oecd.org/sdd/consumerconfidenceshowsaslowingdowninpaceofrecoveryforthefirsthalfof2010.htm`. (Accessed: 2020-24-04)

Okonkwo, U. (2009). Sustaining the luxury brand on the internet. *Journal of brand management*, *16*(5-6), 302–310.

Park, D. H. (2017). The development of travel demand nowcasting model based on travelers' attention: Focusing on web search traffic information. *The Journal of Information Systems*, 171 - 185.

Park, S., Lee, J., & Song, W. (2017). Short-term forecasting of japanese tourist inflow to south korea using google trends data. *Journal of Travel & Tourism Marketing*, *34*(3), 357-368.

Paturohman, S., Suhartanto, D., & Muflih, M. (2018). The effect of consumer interest on Islamic bank deposits: An analysis using Google Trends. In *2018 international conference on information technology systems and innovation (icitsi)* (p. 105-109).

Perju-Mitran, A. (2018). Assessing romanians'opinion on over the top video television using google trends. *Journal of Information Systems & Operations Management*, 132–139.

Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance*, *50*(4), 1201–1228.

Ratchford, B., Talukdar, D., & Lee, M.-S. (2001, 03). A model of consumer choice of the internet as an information source. *International Journal of Electronic*

Commerce, *5*, 7-21.

Refinitive eikon student license. (2020). `https://eikon.thomsonreuters.com/index.html`. (Accessed: 2020-09-30)

Resom, A., Pierre, J., Klimkiewicz, M., & Kalampalikis, N. (2018, 01). Trading the stock market using Google search volumes: a long short-term memory approach. *International Journal of Financial Engineering and Risk Management*, *3*, 3.

Roy, S., Mittal, D., Basu, A., & Abraham, A. (2015). *Stock market forecasting using lasso linear regression model* (Vol. 334).

Shynkevich, Y., Coleman, S., Mcginnity, T., & Belatreche, A. (2015, 07). Stock price prediction based on stock-specific and sub-industry-specific news articles..

Silva, E., Hassani, H., Madsen, D., & Gee, L. (2019, 04). Googling fashion: Forecasting fashion consumer behaviour using google trends. *Social Sciences*, *8*, 111.

Statcounter. (2021). *Search engine market share worldwide.* (Accessed: 2021-05-21)

Subrahmanyam, A. (2019). Big data in finance: Evidence and challenges. *Borsa Istanbul Review*, *19*(4), 283 - 287.

Swamy, V., & Dharani, M. (2019). Investor attention using the Google search volume index - impact on stock returns. *Review of Behavioral Finance*, *11*, 55-69.

von Helversen, B., Abramczuk, K., Kopeć, W., & Nielek, R. (2018). Influence of consumer reviews on online purchasing decisions in older and younger adults. *Decision Support Systems*, *113*, 1 - 10.

Wijnhoven, F., & Plant, O. (2017). Sentiment analysis and Google Trends data for predicting car sales.

Wu, L., & Brynjolfsson, E. (2013, 04). The future of prediction: How Google searches foreshadow housing prices and sales. *SSRN Electronic Journal*.

Yi, S.-W., & Hwang, S.-J. (2009). An empirical study on the relevance of web traffic for valuation of internet companies. *Journal of Intelligence and Information Systems*, *15*(4), 79–98.

Yoshinaga, C., & Rocco, F. (2020, 10). Investor attention: Can Google search volumes predict stock returns? *BBR. Brazilian Business Review*, *17*, 523 - 539.

Zhang, H. (2017, 12). Incorporating Google Trends data in predicting consumer confidence in Sri Lanka.

A Appendix

A.1 Indices

A.1.1 S&P 500 Consumer Discretionary

COMPANY	TICKER
ADVANCE AUTO PARTS	AAP
APTIV PLC	APTV
AUTOZONE	AZO
AZON.COM	AMZN
BENETEAU	BEN
BEST BUY CO	BBY
BMW	BMWG.DE
BOOKING HOLDINGS	BKNG
BORGWARNER	BWA
BURBERRY GROUP	BRBY.L
CALLAWAY GOLF	ELY
CARMAX	KMX
CARNIVAL CORP	CCL
CARNIVAL CORPORATION & PLC	CCL
CHIPOTLE MEXICAN GRILL	CMG
CHOW TAI FOOK JEWELLERY	1929.HK
CIE FINANCIERE RICHEMONT	CFR
CROWN RESORTS	CWN
D.R. HORTON	DHI
DAIMLER	DAIGN.DE
DARDEN RESTAURANTS	DRI
DECKERS OUTDOOR CORP	DECK
DOLLAR GENERAL CORP	DG
DOLLAR TREE	DLTR
DOMINO'S PIZZA	DPZ
EBAY	EBAY
ETHAN ALLEN INTERIORS	ETH
ETSY	ETSY
EXPEDIA GROUP	EXPE
FORD MOTOR CO	F

GALAXY ENTERTAINMENT GROUP	0027.HK
GAP	GPS
GARMIN LTD	GRMN
GENERAL MOTORS CO	GM
GENUINE PARTS CO	GPC
HASBRO	HAS
HILTON WORLDWIDE HOLDINGS	HLT
HOME DEPOT	HD
HOTEL SHILLA	8770.KS
HUGO BOSS	BOSSn
INTERCONTINENTAL HOTELS	IHG
KANGWON LAND	35250.KS
KERING	PRTP
L BRANDS	LB
LAS VEGAS SANDS CORP	LVS
LEGGETT & PLATT	LEG
LENNAR CORP	LEN
LKQ CORP	LKQ
LOWE'S COMPANIES	LOW
LUK FOOK	590.HK
LULULEMON ATHLETICA	LULU
MARRIOTT INTERNATIONAL	MAR
MARRIOTT INTL	MAR
MCDONALD'S CORP	MCD
MELCO	200.HK
MGM CHINA HOLDINGS	2282.HK
MGM RESORTS INTERNATIONAL	MGM
MOHAWK INDUSTRIES	MHK
MONCLER	MONC
MOVADO GROUP	MOV
NEWELL BRANDS	NWL
NIKE	NKE
NIKON CORP	7731.T
NORDSTROM	JWN
NORWEGIAN CRUISE LINE	NCLH
NORWEGIAN CRUISE LINE HOLDINGS LTD	NCLH

O'REILLY AUTOMOTIVE	ORLY
PARADISE CO	34230.KQ
POLARIS	PII
POOL CORP	POOL
PORSCHE AUTOMOBIL HOLDING	PSHG
PRADA	1913.HK
PULTEGROUP	PHM
PVH CORP	PVH
RALPH LAUREN CORP	RL
RESORTTRUST	4681.T
RH	RH
ROSS STORES	ROST
ROYAL CARIBBEAN CRUISES LTD	RCL
SALVATORE FERRAGAMO	SFER
SANDS CHINA	1928.HK
SHANGRI-LA ASIA	69.HK
SJM HOLDINGS	880.HK
SLEEP NUMBER CORP	SNBR
STARBUCKS CORP	SBUX
SWATCH GROUP	UHR
TAPESTRY	TPR
TARGET CORP	TGT
TEMPUR SEALY	TPX
TESLA	TSLA
THE STAR ENTERTAINMENT GROUP	SGR
TIFFANY & CO	TIF
TJX COMPANIES	TJX
TOD'S GROUP	TOD
TOLL BROTHERS	TOL
TRACTOR SUPPLY CO	TSCO
ULTA BEAUTY	ULTA
UNDER ARMOUR	UAA
VAIL RESORTS	MTN
VF CORP	VFC
WHIRLPOOL CORP	WHR
WYNN MACAU	1128.HK

WYNN RESORTS	WYNN
WYNN RESORTS LTD	WYNN
YUM! BRANDS	YUM

A.1.2 S&P 500 Consumer Staples

COMPANY	**TICKER**
INTER PARFUMS	IPAR
SHISEIDO	SSDOY
AMOREPACIFIC CORP	90430.KS
WALMART	WMT
PHILIP MORRIS INTERNATIONAL	PM
ALTRIA GROUP	MO
MONSTER BEVERAGE CORP	MNST
HORMEL FOODS CORP	HRL
TYSON FOODS	TSN
BROWN FORMAN CORP	BF-B
DIAGEO	DGE
PERNOD RICARD	PERP
REMY COINTREAU	RCOP
DAVIDE CAMPARI-MILANO	CPRI
TREASURY WINE ESTATES	TWE
PROCTER & GAMBLE CO	PG
COCA-COLA CO	KO
COSTCO WHOLESALE CORP	COST
ESTEE LAUDER COMPANIES	EL
MONDELEZ INTERNATIONAL	MDLZ
COLGATE-PALMOLIVE CO	CL
KIMBERLY-CLARK CORP	KMB
WALGREENS BOOTS ALLIANCE	WBA
SYSCO CORP	SYY
GENERAL MILLS	GIS
HERSHEY CO	HSY
ARCHER-DANIELS-MIDLAND CO	ADM
KROGER CO	KR
MCCORMICK & COMPANY	MKC

CHURCH & DWIGHT CO	CHD
KELLOGG CO	K
CONAGRA BRANDS	CAG
CAMPBELL SOUP CO	CPB
J M SMUCKER CO	SJM
LAMB WESTON HOLDINGS	LW
MOLSON COORS BEVERAGE CO	TAP
HANESBRANDS	HBI
PEPSICO	PEP
CONSTELLATION BRANDS	STZ
KRAFT HEINZ CO	KHC
CLOROX CO	CLX

A.2 GICS Industry Classification

Table 18: Companies filtered by industry

FOOD, BEVERAGE & TOBACCO	CONSUMER SERVICES	RETAILING	CONSUMER DURABLES & APPAREL
ARCHER-DANIELS-MIDLAND CO	CARNIVAL CORPORATION & PLC	ADVANCE AUTO PARTS	BENETEAU
BROWN FORMAN CORP	CARNIVAL CORP	AUTOZONE	BURBERRY GROUP
CAMPBELL SOUP CO	CHIPOTLE MEXICAN GRILL	AZON.COM	CALLAWAY GOLF
CHURCH & DWIGHT CO	CROWN RESORTS	BEST BUY CO	CIE FINANCIERE RICHEMONT
COCA-COLA CO	DARDEN RESTAURANTS	BOOKING HOLDINGS	D.R. HORTON
COLGATE-PALMOLIVE CO	GALAXY ENTERTAINMENT GROUP	CHOW TAI FOOK JEWELLERY	DECKERS OUTDOOR CORP
CONAGRA BRANDS	DOMINO'S PIZZA	CARMAX	ETHAN ALLEN INTERIORS
COSTCO WHOLESALE CORP	HILTON WORLDWIDE HOLDINGS	DOLLAR GENERAL CORP	GAP
DAVIDE CAMPARI-MILANO	HOTEL SHILLA	DOLLAR TREE	GARMIN LTD
DIAGEO	INTERCONTINENTAL HOTELS	EBAY	HASBRO
ESTEE LAUDER COMPANIES	KANGWON LAND	ETSY	HUGO BOSS
GENERAL MILLS	LAS VEGAS SANDS CORP	EXPEDIA GROUP	KERING
HERSHEY CO	MARRIOTT INTL	GENUINE PARTS CO	LULULEMON ATHLETICA
J M SMUCKER CO	MELCO	HOME DEPOT	LEGGETT & PLATT
KELLOGG CO	MGM CHINA HOLDINGS	LUK FOOK	LENNAR CORP
KIMBERLY-CLARK CORP	MARRIOTT INTERNATIONAL	L BRANDS	MONCLER
KROGER CO	MCDONALD'S CORP	LKQ CORP	MOVADO GROUP
LAMB WESTON HOLDINGS	MGM RESORTS INTERNATIONAL	LOWE'S COMPANIES	NIKON CORP
MCCORMICK & COMPANY	NORWEGIAN CRUISE LINE	NORDSTROM	MOHAWK INDUSTRIES
MOLSON COORS BEVERAGE CO	PARADISE CO	O'REILLY AUTOMOTIVE	NEWELL BRANDS
MONDELEZ INTERNATIONAL	RESORTTRUST	POOL CORP	NIKE
PERNOD RICARD	SANDS CHINA	ROSS STORES	POLARIS
PROCTER & GAMBLE CO	ROYAL CARIBBEAN CRUISES LTD	TARGET CORP	NORWEGIAN CRUISE LINE HOLDINGS LTD
REMY COINTREAU	SHANGRI-LA ASIA	TIFFANY & CO	PRADA
SYSCO CORP	SJM HOLDINGS	TJX COMPANIES	PVH CORP
TREASURY WINE ESTATES	STARBUCKS CORP	TRACTOR SUPPLY CO	RALPH LAUREN CORP
WALGREENS BOOTS ALLIANCE	THE STAR ENTERTAINMENT GROUP	ULTA BEAUTY	PULTEGROUP
ALTRIA GROUP	VAIL RESORTS		RH
MONSTER BEVERAGE CORP	WYNN MACAU		SALVATORE FERRAGAMO
PHILIP MORRIS INTERNATIONAL	WYNN RESORTS		SLEEP NUMBER CORP
	WYNN RESORTS LTD		SWATCH GROUP
	YUM! BRANDS		TAPESTRY
			TEMPUR SEALY
			TOD'S GROUP
			TOLL BROTHERS
			UNDER ARMOUR
			VF CORP

HOUSEHOLD & PERSONAL PRODUCT	AUTOMOBILES & COMPONENTS	FOOD & STAPLES RETAILING
AMOREPACIFIC CORP	APTIV PLC	CONSTELLATION BRANDS
HORMEL FOODS CORP	BMW	HANESBRANDS
INTER PARFUMS	BORGWARNER	KRAFT HEINZ CO
SHISEIDO	DAIMLER	PEPSICO
CLOROX CO	FORD MOTOR CO	TYSON FOODS
WHIRLPOOL CORP	GENERAL MOTORS CO	WALMART
	PORSCHE AUTOMOBIL HOLDING	
	TESLA	

A.3 Industry-related keywords

Table 19: Industry keywords

FOOD, BEVERAGE & TOBACCO	CONSUMER SERVICES	RETAILING	CONSUMER DURABLES & APPAREL
coffee	hotel	logistics	curtains
tea	hotels	distribution	decorations
starbucks	barcelona	logistica	quatro
mate	inn	walmart	stickers
bottle	vegas	distribution center	decoration
cola	restaurant	asset	frame
smoothie	restaurants	logistic	stickers
coca	pizza	ambar	mirror
soda	dominos	myhermes	frames
teer	mcdonalds	telefono	decoracao
cigarettes	consumer services	marketing	lottery
cigarettes	consumer service	google adwords	dog
tobacco	clc	ad	birthday
smoke shop	clc sonsumer	adsense	imagenes
tabac	journal of retailing	advertising	christmas
marlbro	consumer industry	seo	google
hookah		amazon	youtube
shisha		netflix	cat
lighter		nike	shoes
		walmart	nike
		bon coin	adidas
		le boin coin	boots
		facebook	dress
		adidas	bijoux
		zara	cartier
			mont blanc
			hublot
			luxury

HOUSEHOLD & PERSONAL PRODUCTS	AUTOMOBILES & COMPONENTS	FOOD & STAPLES RETAILING	
paint	apk	lidl	
doors	car	aldi	
leroy	ford	carrefour	
flooring	honda	asda	
leroy merlin	toyota	grocery	
lowes	bmw	leclerc	
saw	bike	sainsburys	
tiles	olx	supermercado	
carpet	audi	tesco	
nails	mercedes	morrisons	
makeup	tires		
cream	auto parts		
nail	autozone		
perfume	pneu		
face	gps		
sephora	garmin		
ipl	turbo		
avon	advance		
	tires		

A.4 Brand-related keywords

Table 20: Brand related keywords

INTER PARFUMS:	**SHISEIDO:**	**AMOREPACIFIC CORP:**	**WALMART:**	**PHILIP MORRIS INTERNATIONAL:**
Abercrombie & Fitch	Shiseido	Amore Pacific	Walmart	Philip Morris
Karl Lagerfeld	cle de peau	Sulwhasoo	Walmart george	Marlboro
Anna Sui	NARS	Laneige	Terra & Sky	Parliament
Lanvin	bareMinerals	Mamonde	Time and Tru	Virginia S
MCM	Anessa	Innisfree	Wonder Nation	L&M
Mont Blanc	Dolce & Gabbana		Athletic Works	Lark
Graff	Drunk elephant		Brahma	Merit
Guess	d program		EV1	Muratti
Hollister	Elixir		No Boundaries	Bond Street
Jimmy Choo	IPSA		Secret Treasures	Chesterfield
Oscar de la renta	Laura Mercier		And1	Next
Paul Smith	Senka		Avia	Red & White
Repetto	Tory Burch		Sams choice	IQOS
Rochas			Equate	
s.t. dupont			Mainstays	
			Ol Roy	

ALTRIA GROUP:	**MONSTER BEVERAGE CORP:**	**HORMEL FOODS CORP:**	**TYSON FOODS:**	**BROWN FORMAN CORP:**
Altria	monster energy	hormel	tyson	Brown Forman
PhilipMorrisUSA	full throttle	applegate	jimmy dean	Jack daniels
US smokeless	burn drink	austin blues barbeque	hillshire farm	woodford reserve
John Middleton		bacon1	raised & rooted	old forester
Ste Michelle		black label bacon	aidells	coopers craft
Philip morris capital corporation		cafe h	statefair	slane irish whiskey
ABInBev		chi-chi's	saralee	the benriach
Juul labs		hormel chili	wright brand	the glendronach
Cronos		compleats	bosco's	herradura
Helix innovations		cure 81	the bruss company	el jimador
		columbus craft	barberfoods	pepe lopez
		dan's prize	fast fixin	finlandia
		dinty moore	like mom's	chambord
		dona maria	landshire	korbel
		embasa	steak eze	fords gin
		fire braised	true chews	
		fuse burger	top chews	
		happy litle plants	rewben	
		herb ox		
		herdex		
		house of tsang		
		lloyds barbeque		
		mary kitchen		
		natural choice		
		not so sloppy joe		
		old smokehuse		
		skippy		
		stagg chili		

DIAGEO:	PERNOD RICARD:	REMY COINTREAU:	DAVIDE CAMPARI-MILANO:	TREASURY WINE ESTATES:
Diageo	Pernod Ricard	Remy Cointreau	Davide Campari-Milano	Treasury Wine Estates
black & white	royal salute		aperol	19 crimes
buchanans	mumm		campari	acacia vineyard
J&B	martell		skyy vodka	annies lane
johnnie walker	beefeater		wild turkey	beaulieu vineyard
grand old parr	chivas regal		wrey and nephew	belcreme de lys
lagavulin	absolut vodka			beringer
the singleton	havana club			blossom hill
talisker	jameson			coastal estates
windsor	the glenlivet			coldstream hills
bulleit	perrier-jouet			embrazen
crown royal	malibu			etude
ciroc	ricard			fifth leg
ketel one	ballantines			heemskerk
smirnoff	kenwood			hewitt vineyard
bundaberg	campo viejo			ingoldby
captain morgan	brancott estate			jamieson's run
ron zacapa	jacobs creek			killawarra
baileys	royal stag			leo buring
casamigos	imperial blue			lindeman's
don julio	100 pipers			maison de grand esprit
gordons	imperial			matua
tanqueray	passport scotch			meridian
Mcdowells	clan campbell			metala
Shui Jing Fang	seagrams gin			penfolds
Yeni raki	ramazzotti			pepperjack
Ypioca	pastis 51			provenance
Guinness	olmeca			rawson's retreat
	ararat			rosemount estate
	blenders pride			run riot
	kahlua			saltram

Appendix

PROCTER & GAMBLE CO:	COCA-COLA CO:	COSTCO WHOLESALE CORP:	ESTEE LAUDER COMPANIES:	MONDELEZ INTERNATIONAL:
Pampers	Coca Cola	costco	Estee Lauder	5 star
Downy	Sprite	kirkland signature	Aerin	alpen gold
Dreft	Fanta	costco insurance	aramis	barni
rindex 3en1	Dasani	costco wholesale	Aveda	belvita
charmin	smartwater	innovel solutions	Becca	bournvita
puffs	Minute Maid		Bobbi Brown	cadbury
tampax	innocent		Bumble and bumble	cadvury diary milk
this is l	simply juices and drinks		Clinique	chips ahoy!
braun	georgia coffee		Darphin	clorets
gillette	costa coffee		DKNY	freia
always discret	fuze tea		DonnaKaran	honey maid
joy+glee	honest tea		Dr Jart	kinh do
gillette venus	fairlife		Frederic Malle	lacta
head & shoulders	powerade		Ermenegildo Zegna	marabou
herbal essences	ciel		Glamglow	maynards bessett's
pantene	schweppes		Jo Malone	milka
ambi pur	vitaminwater		Kilian	oreo
dawn ultra	gold peak tea		La mer	perfect snacks
microban 24	appletiser		Lab series	philadelphia
mr. clean	topo chico		Le labo	sour patch kids
swiffer	aquarius		MAC	tate's bake shop
oral-b	ades		Michael Kors	toblerone
fixodent	fresca		Origins	trident
clearblue	I lohas		Rodin	triscuit
meta mucil	ayataka		smashbox	wheat thins
pepto bismol	barqs		Tom ford beauty	
prilosec otc	dogadan		Too Faced	
zzzquil	peace tea			

BENETEAU:	BEST BUY CO:	BMW:	BOOKING HOLDINGS:	BORGWARNER:
Beneteau	Best Buy	BMW	Booking.com	Borgwarner
Figaro Beneteau	Insignia	MINI	Kayak	Borgwarner drivetrain
Oceanis	Rocketfish	Rolls Royce	Priceline	Borgwarner engine
Flyer	Dynex	BMW M	Agoda.com	
Barracuda	Platinum		Rentalcars.com	
Antares	Modal		OpenTable	
Gran turismo	Best buy essentials		Booking	
Swift trawler			Agoda	
Grand trawler			Rentalcars	

COLGATE-PALMOLIVE CO:	KIMBERLY-CLARK CORP:	WALGREENS BOOTS ALLIANCE:	SYSCO CORP:	GENERAL MILLS:
Colgate	poise	almus	Sysco	betty crocker
Palmolive	plenitud	no7	Arrezzio	jus-tol
Protex	huggies	liz earle	Bakers source	pillsbury
Sanex	goodnites	soap & glory	Block & barrel	bisquick
Softsoap	drynites	sleek makeup	buckhead pride	immaculate baking
Hills	little swimmers	yourGoodSkin	butchers block	knack & back
Sorriso	kleen bebe	walgreens	casa solana	pillsbury atta
Speed stick	kleenex		citavo	cheerios
lady speed stick	cottonelle		sysco earth plus	kix
Suavitel	scottex		fire river farms	wheaties
Murphy	u by kotex		jade mountain	lhaagen-dazc
meridol	kotex		newport pride	old el paso
Irish spring body wash	kotex whote		wholesale farms	wanchai ferry
Toms of maine	kotex anydays		portico seafood	v.pearl
tahiti	kotex goodfeel		pica y salpica	muir glen
softlan	intimus		riserva	blue buffalo
Ajax	camelia			totino's
fleecy	wypall			nature valley
pinho sol	kimtech			parampara
axion	kleenguard			yoplait
cuddly				
stasoft				
elmex				
fabuloso				
soupline				
hello				
fluffy				
PCA skin				
filorga				
elta md skincare				

Appendix

HERSHEY CO:	ARCHER-DANIELS-MIDLAND CO:	KROGER CO:	MCCORMICK & COMPANY:	CHURCH & DWIGHT CO:
hersheys	archer daniels	baker's	billy bee	arm & hammer
kitkat		city market	cattlemen's	oxi clean
jolly rancher		food less	cholula	rephresh
twizzler		fredmeyer	drogheria	replens
reese's		gorbes	ducros	vitafusion
brookside		kroger	frank's redhot	lilcritters
almond joy & mounds		marianos	kohinoor	pb8
allan candy		metro market	lawry's	wellgate
rolo		pick'n save	margao	clump & seal
breath savers		roundy's	mccormick	felinePine
		vitacost	stubb's	orange glo
			zatarain's	batiste
				orajel
				nair
				spinbrush
				simply saline
				viviscal waterpik
				arrid
				pepsodent

KELLOGG CO:	CONAGRA BRANDS:	CAMPBELL SOUP CO:	J M SMUCKER CO:	LAMB WESTON HOLDINGS:
kellogs	slim jim	campbell	dunkin	colossal crisp
cheezit	gardein	capecod	jif	crispycoat
pringles	duncan hines	goldfish	smucker's	hearty house
pop tarts	hunts	kettle brand	uncrustable's	tavern traditions
all-bran	reddi wip	latejuly	meow mix	lamb weston
nutrigrain	vlasic	pacific foods	cafe bustelo	
frosted flakes	boom chicka pop	spaghettios	milk-bone	
chrunchy nuts	duke's	pop secret	laura scrudder's	
coco pops	orville redenbacker's		r.w knudsen family	
froot lops	udi's		hungry jack	
corn flakes			kibbles n bites	
corn pops			9lives	
			martha white	
			folgers	
			truRoots	
			sahale snacks	

Appendix

MOLSON COORS BEVERAGE CO:	HANESBRANDS:	PEPSICO:	CONSTELLATION BRANDS:	KRAFT HEINZ CO:
carling	Hanes	pepsi	constellation brands	heinz
coors banquet	Champion	lays	corona	greenseas
coors light	Playtex	mountain dew	funky buddha	original juice co.
george killian's irish red	Maidenform	doritos	modelo	kraft heinz
granville island brewing	JMS	gatorade	Kim Crawford	bull's-eye
hamm's	Just my Siza	tropicana	Meiomi	biaglut
hop valley	Nur Der	quaker oats	The prisoner wine company	nipiol
miller high life	Nur Die	lipton	cooper & thief	plasmon
molson	Bras N Things	aquafina	charles smith wines	ore-ida
pilsner urquell	Berlei	ruffles	7 moons	de ruijter
steel reserve	Gear for Sports	cheetos	robert mondavi winery	kravan cevitam
terrapin		brisk	ruffino	roosvicee
vizzy hard seltzer		tostitos	flotidian	wijko
		fritos	chant IPA	cat prefer chef
		diet pepsi	pacifico	food in a minute
		sierra mist	Svedka	nutri+plus
		7up	Copper & kings	purepet
		mirinda	FCBC	lea & perrins
		walkers	Wildish	hp sauce
		pepsi black	Nelsons green brier	bagel bites
			rosatello	capri sun
			auros	cool whip
			mount veeder	cracker barrel
			belle meade bourbon	delimex
			simi	grey poupon
			young & co	heinz 57 sauce
			woodbridge	heinz chili sauce
			two lane	heinz cocktail
			casa noble tequila	jack daniel's sauce
			high west distillery	jell-o
			mi campo	jet-puffed

CLOROX CO:	ADVANCE AUTO PARTS:	APTIV PLC:	AUTOZONE:	AZON.COM:
agua jane	Advance auto parts	Aptiv	AutoZone	Kindle
ayudin	Carquest		Valucraft	Amazon Web Services
blessed herbs	Worldpac		Duralast	Amazon Books
burt's bees	autopart international		Duralast Gold	Amazon Go
clorox	diehard power ahead			Amazon Pop Up
cloroxpro				Whole Foods
formula 409				Prime Video
handi wipes				Prime Music
liquid-plumr				Prime Reading
mistolin				AmazonFresh
mortimer				Alexa
neocell				Echo
pine-sol				Fire tablets
pinoLuz				Fire TV
soy vay				Amazon Digital Game Store
trenet				Amazon Studios
				Audible
				Diapers.com
				Bookpages
				Telebook
				IMDb
				Junglee
				PlanetAll
				LiveBid
				Acccept.com
				Shopbop
				TextPayMe
				Brilliance Audio
				Zappos
				Touchco
				Woot

NTNU

Norwegian University of
Science and Technology