Big Data Strategies – Worth the Hype?

Cross-regional Study on the Ability of Google SVIs in Predicting Abnormal Returns, Liquidity and Volatility

Ksenia Parviainen*

Stockholm School of Economics

Bachelor Thesis
Stockholm School of Economics
Department of Finance
Tutor: Nikita Koptyug
May 14, 2018

Abstract

This paper combines studies around investor attention, equity home bias puzzle and gradual diffusion of local information hypothesis. I study whether Google's Search Volume Index (SVI) has the ability to predict abnormal returns, volatility and liquidity on an industry level. To add to previous research, I study US and UK investors and their behavioural differences in investing within S&P500 and FTSE100 indices. I find US investors to be more home biased when measured in investor attention proxied by SVIs. Further, I find FTSE100 industries to have the ability to predict S&P500 abnormal returns, liquidity and volatility while there isn't enough evidence for the opposite to hold. When looking at searches between local and non-local investors, the non-local searches are mostly affected by same week's local searches. This suggests that information asymmetries between local and non-local investors are diffused within a week.

Keywords: Google, Big Data, Stock Market Prediction, SVI

I would like to thank my supervisor Nikita Koptyug for all the helpful discussions and information regarding how to build my paper. All mistakes found in this work are my own.

*23710@student.hhs.se

Table of Contents

1 Introduction

The investors of today have access to large amounts of financial data with just a single click. However, processing such massive amounts of financial data is a problem in itself and requires the use of computational power and can prove to be difficult to process. 'Big data'[1] is a term that has been coined to describe this scenario. 'Big data strategies' are thus, the methods of processing such large amounts of data and thereby recognising patterns which can be further developed into strategies and algorithms that aim to produce profit for an investor and finding alpha. Are big data strategies worth the hype in finding alpha or have we reached the point of truly efficient financial markets?

The efficient market hypothesis assumes that all public information is processed by investors as soon as it is made public. This means that each investor is devoting an equal amount of time and attention into each asset. Thus, no systematic alpha[2] can be made. However, the reality is a whole lot different. As Kahneman (1973) put it, “attention is a scarce cognitive resource”. In the context of investors, this means that an investors attention is selective and biased based upon their environment and interests at that moment in time. Further, due to the fact there is merely too much data for an investor to process without the help of computational power, an investor would be forced into being selective and biased in their thoughts and actions. Therefore, studying how investors allocate their attention can be a helpful tool in predicting the movements of the financial market and finding the, oh so coveted, alpha.

One may ask, how would you measure investor attention? Historically, this has proven to be difficult since there hasn't been any direct measure that can capture investor attention within financial markets as a whole. Because of this, empiricists have been constrained to more indirect measures such as extreme returns, trading volume, news and headlines (Barber and Odean, 2008), as well as, advertising expense (Chemmanur and Yan, 2009) and price limits (Seasholes and Wu, 2007). All of the aforementioned proxies assume that if a news piece, extreme return or turnover has been generated, investor attention has been triggered. The aforementioned assumption doesn’t take into account any other factors moving these measures. Further, would it viable to assume

[1] Big data can be defined as extremely large datasets that may only be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions.

[2] If alpha is found, it is instantly exploited by arbitrageurs and thus one shouldn’t be able to find a strategy that generates alpha on a continuous basis.

that we actually read every single news article generated, or even, merely click on every attention-grabbing headline of said news articles? Certainly not in 2018.

In order to tackle the shortcomings of the aforementioned indirect investor attention measures, Da et. Al (2011) propose a direct measure of investor attention. The authors argue search engine data to be a better proxy for measuring investor attention and thus being a more helpful tool in predicting market movements. Specifically, Da et. Al (2011) define Google's aggregate search frequency as the 'holy grail' of investor attention and market prediction research. There are three main reasons why this is the case. Firstly, Google is the number one search engine used across the world. More importantly, Google accounts for an 80% market share in developed countries, such as the U.S. and UK[3]. Thus, when studying developed countries the search volumes on Google can be seen as representative of internet searches of the population. Secondly, as Da et Al. (2011) put it, Google "search is a revealed attention measure". If one takes their time to actively search information on a stock on Google they are definitely paying attention to it. Therefore, search engine data can be seen as an active proxy for investor attention, which is both direct and unambiguous in its nature. Thirdly, Google's search engine data is accessible to anyone via the Google Trends service[4], which makes the use of this data possible and easily attainable. The service gathers the amount of searches generated for each search term and publicises it in form of Search Volume Index (SVI)[5]. This data is published up to a minute basis, depending on the time period chosen. Furthermore, Google Trends enables the study of the amount of searches on a regional basis as well as presenting related topics and queries as a means for providing a deeper understanding behind the purpose for the searches made. The other aspect Da et. Al (2011) touch upon is defining whose attention Google's SVIs are capturing: the authors find SVIs to be a suitable proxy for the attention of less informed retail investors. This is because institutional investors have more sophisticated tools, such as Bloomberg and Eikon terminals, at their disposal in aiding their investment decisions.

The purpose of this paper is, to use the direct investor attention measure proposed by Da et. Al (2011) and broaden upon their findings of the helpfulness of SVIs in predicting

[3] Google market share by countries, https://www.statista.com/statistics/220534/googles-share-of-search-market-in-selected-countries/

[4] Google Trends website, https://trends.google.com/trends/

[5] An index number denoting amount of searches for a specific search term at a given time period. Calculated as the amount of searches proportionate to the time and location of a query by dividing the total searches of the geography and time range it represents to compare relative popularity. Resulting numbers scaled on a range of 0 to 100 based on a topic's proportion to all searches on all topics.

stock market movements. To succeed in this, it is important to define search terms that best capture investor attention thereby filtering out noisy data. For instance, the goal is to capture the amount of searches generated by individuals wanting to invest into the Tesla stock, as opposed to, those interested in merely buying Tesla cars. Da et. Al (2011) suggest the use of ticker symbols to be the best solution to this: this is due to the fact that searches bearing the company name are more likely to capture searches unrelated to investing. Furthermore, various variations of company names can be used whilst there is only a single ticker for a company on a specific stock market index. I aim to test the assumed superiority of tickers as search terms by taking advantage of newly added features of Google Trends: related queries and topics. Instead of assuming tickers being the best search term by default, I will test various different search terms and choose those that output related queries and topics most relevant to investing. In this way, I will hope to reduce the noisiness of data and thereby resulting in more precise results. Finally, I differentiate from previous SVI studies on company or market capitalisation level and analyse the effects on industry level. In addition to this, I add to previous research by incorporating Google Trends' geographical filters by studying differences in searches and investment behaviour of US and UK investors on S&P500 and FTSE100 indices.

More precisely, I study four different hypotheses that combine previous findings of SVI, equity home bias puzzle and information asymmetry research[6]. My main hypothesis is that S&P500 industry SVIs are helpful in predicting movements on, not only S&P500, but also on FTSE100 and vice versa. Therefore, I expect the changes in SVIs within an industry on S&P500 to predict market movements for that industry both on S&P500 and FTSE100. Similarly, the changes in SVIs within an industry on FTSE100 would predict movements of that industry on FTSE100 as well as S&P500. I find statistically significant evidence that SVIs for FTSE100 industries are, in fact, able to predict abnormal returns, volatility and liquidity for the same industries on S&P500 but there isn't enough evidence for the opposite to hold. The effect is especially strong when studying abnormal volatility and liquidity. Looking at abnormal returns, only FTSE100 Financials seem to predict movements on S&P500, both when studying effects of abnormal SVIs of one-and-two-week lags. In general, the results with one-week lags indicate positive effects on the dependent variables, while the results for two-

[6] Some examples of good papers: concerning SVI research Da et. Al (2011); concerning equity home bias puzzle French & Poterba (1991); on information asymmetry i.e. Hong & Stein (1999).

and three-week lags indicate negative effects, which suggests the effect of investor attention on the S&P500 and FTSE100 indices to be temporary.

Since SVIs are assumed to capture the attention of retail investors, I expect the predictive ability to be more significant for industries that are more familiar to retail investors. In this paper, familiarity is defined based on average SVIs, where more familiar industries are denoted as those generating average SVIs within 75th percentile. The results show evidence for this to hold with most familiar industries generating more statistically significant results within most of the regressions. Out of these industries Financials generate most results of interest, while also noteworthy, are results for Consumer Staples, Health Care and Materials.

I also study whether gradual diffusion of local information, discussed by Hong and Stein (1999) holds when SVIs are used as a measure of investor attention. If local information does diffuse gradually, I expect this week's SVIs of non-local investors to be affected by, not only last week's SVIs of non-locals, but also by SVIs of locals both for this and last week. I find evidence that the local SVIs do, in fact, affect the non-local volumes. However, the effect is mainly found within same week searches, as opposed to, last week's searches. Thus, it is hard to define whether the gradual diffusion of information actually exists or whether my data isn't frequent enough to capture it. Since, the regressions do generate some results indicating effect of last week's local searches, I believe the latter to be true. Further, I test whether the information asymmetry is stronger for smaller and more remote areas, as well as, for smaller companies measured in market capitalisation. I find that within FTSE100 the local searches (UK searches) have more significant effect on non-local searches (US searches) when compared to results within S&P500. This is an indication of information asymmetry to be stronger for smaller areas since the UK economy is smaller than that of U.S. The results from regressions on market capitalisation, on the other hand, are quite ambiguous. While the results from S&P500 are in line with the hypothesis of smaller market capitalisation generating more information asymmetry, I find the opposite to hold for FTSE100.

Finally, I study whether the home bias is stronger for US than UK investors, as discussed by French and Poterba (1991), Tesar and Werner (1995) and Mishra (2015). I expect for the average SVIs of US investors to be significantly lower than that of UK investors within FTSE100. I also expect the SVIs to be similar for both regions within S&P500. Results from differences in means tests prove this assumption to hold. Interestingly, I also find evidence of UK and US investors to be home biased especially within Technology; since for this particular industry the average local SVIs tend to be high, whilst, average non-local SVI's are significantly lower. A similar effect can be found for Consumer Discretionary for US investors and Utilities for UK investors.

The remainder of this paper is structured as follows. In Section 2 I discuss some of the previous research relevant to the analysis in this study, in Section 3 I present the theoretical framework and the hypotheses studied. In Section 4 I describe the data collection process, while, Section 5 introduces the analytical framework (the variables and regressions used to derive the results). In Section 6 I discuss the main findings of interest, while, in Section 7 I provide conclusions and present the limitations of this paper, as well as, putting forward suggestions for future research.

2 Previous Research

In order to interpret and understand the results and implications of this paper, it is important to have a grasp of the previous research that this study is built upon. This research can be divided into three categories: research on using search engine data as a proxy for investor attention, research concerning home bias, as well as, information asymmetries between local and non-local investors. I also discuss my contributions to the aforementioned categories of research.

2.1 Search Engine Data as a Proxy for Investor Attention

Firstly, it is crucial to understand why active investor attention proxies are a better tool in predicting stock market movements, in comparison to, passive investor attention proxies. Da et Al. (2011) were the first to propose a direct measure for investor attention instead of the indirect measures popular at the time such as: extreme returns, trading volume, news, headlines or advertising expenses. The authors argue that these measures can be noisy and driven by factors other than investor attention. A more suitable alternative for such passive proxies of investor attention are seen to be active proxies, such as, Google’s Search Volume Index (SVI). This measure reports the amount of Google searches for a particular term, relative to all Google searches over some time period; this output is then normalised in order to generate a search index number between 0-100 as a means of comparison compared to other search terms. Further, Da et Al. (2011) study whether SVIs can predict stock market movements and compare these results to the aforementioned indirect measures, in order to, test whether SVIs are, de facto, superior. The authors find increased SVIs to lead to higher stock prices in regards to one and-two-week lags, however, this increase reverses with an increase in the lag time. In addition, Da et. Al (2011) see evidence of the SVI measure to be correlated with the other indirect attention measures studied, but, conclude SVIs to capture investor attention “in a more timely fashion”. The authors also determine the SVI measure to capture the attention of less sophisticated retail investors. This is due to the fact that, institutional investors tend to do their investment research on platforms such as, Bloomberg and Eikon terminals. Further, Da et. Al (2011) propose that retail investor attention leads to a positive price pressure based on discussion by Barber and Odean (2008). Barber and Odean (2008) argue that due to constraints in short-selling, retail investors perform active information search geared towards buying into assets. By selling assets, retail investors are mostly constrained to the assets they already own, thereby, making active information search redundant.

While Da et. Al (2011) argue for the use of ticker symbols for minimal noisy data in regards to studying investor attention measured in SVIs, Lundström and Nestius (2012) propose a similar study in smaller markets with the use of company names instead of tickers. The authors focus on the Nordic indices and find SVIs to help in the prediction of abnormal returns, liquidity and volatility with this effect showing more prominence in smaller firms

In more recent papers, Ding and Hou (2015) find the passive investor attention measures to explain little of variation in the SVI measure. Further, the authors find that the change in SVIs, and thus in investor attention, improves stock liquidity. Contrary to Da et Al’s (2011) results, Chen (2017) finds increased investor attention measured in SVIs to have a negative effect on global index returns. Further, the author finds this result to be more prominent for developed countries and especially for US investors.

Since Google has added new features after the release of the aforementioned papers, my main contribution to this research topic is to incorporate and study said new features. Firstly, I take advantage of the 'related topics and queries' feature as a means of correcting for noisy data, which is prevalent in previous studies. Secondly, I use regional searches instead of global ones to study whether there are behavioural differences in US and UK investors, but also, their attention to the stock markets measured by SVIs. Finally, I study whether the SVIs for a particular regional index have any predictive ability over other regional indices.

2.2 Equity Home Bias Puzzle

Secondly, it is important to understand why investors differ in their investment behaviour when studied across regions. The sc. equity home bias puzzle offers one explanation. French and Poterba (1991) study the amount of home bias in Japan, US and UK. They also propose a few reasons as to why home bias exists. The authors find that the US investors hold 94% of their equity investment in U.S., while UK investors hold 86% of their portfolio in domestic equities. This higher rate of foreign investment in UK to that of U.S., can partially be explained by the higher standard deviation of domestic returns in UK, thus leading to a need of higher international diversification than in U.S. Further, French and Poterba (1991) find that the 18% of UK investors' foreign holdings are equally divided into investments in the U.S., Japan and continental Europe (leading to a holding of around 6% each) The authors define expectation of systematically higher domestic returns, to that of foreign returns, as one explanation to the home bias. This can be explained by both overconfidence of the performance of domestic markets, as well as, being wary of foreign markets leading to a risk premium due to a 'fear of the unknown'. Tesar and Werner (1995) come to similar conclusions: despite gains of international diversification being a well-known fact, investors do not use this strategy to its full potential leading to home bias in investment portfolios. Further, the authors find investors being more willing to invest in foreign equities of geographical proximity. This further suggests that, investors are home biased due to being insecure of investing into markets they are less familiar with.

A more recent paper by Mishra (2015) develops measures of home bias for 42 countries over a ten-year period, 2001-2011, using different models such as: international CAPM (ICAMP), mean-variance, minimum-variance, Bayes-Stein, Bayesian and multi-prior correction to Bayesian. For home bias volumes in UK and U.S. the author finds US investors to be more biased than UK investors, although, the differences are less prominent after robustness checks. Mishra (2015) finds that foreign listings in the domestic market reduce the equity bias toward that foreign country. Further, an increase in natural resources rents and better corporate governance in a country leads to a decrease in home bias. Size, on the other hand, has an opposite effect: the bigger the economy, the higher the home bias.

My input to this line of research lies in, measuring home bias as the differentiating factor in relation to SVIs in local and non-local markets. In addition, I will place a particular emphasis on studying whether US and UK investors differ in their home biasness when investing into both domestic and foreign markets.

2.3 Information Asymmetries between Local and Non-local Investors

Finally, my dataset of two markets provides the ability to study the differences in behaviour of local and non-local investors. Thereby, offering a complimentary perspective to the home bias research. Hong and Stein (1999) develop a hypothesis of gradual diffusion of local information, which is based on a theoretical world of two heterogeneous investor groups – newswatchers and momentum traders. Newswatchers observe private signals about future price movements but are unable to incorporate information obtained by other newswatchers', thus leading to underreaction to signals in the short run. Momentum traders, on the other hand, base their investment decisions on past prices and will arbitrage the underreaction of newswatchers but are limited to simple trading strategies. Therefore, their attempts at arbitrage lead to overreaction in the long run. Although this arbitrage exploitation does profit on average, the difficulty lies in determining which of the two factors is driving the price increase. If the price is increasing due to newswatchers underreacting to a signal and momentum traders starting to exploit the arbitrage, these momentum traders will earn money. If on the other hand, the price increase is due to previous momentum trades rather than the arbitrage exploitation, the momentum traders will be 'late to the party' and therefore will lose money.

Cziraki et Al. (2017) test the Hong and Stein's (1999) gradual diffusion of local information as well as Da et Al.'s (2011) positive price pressure hypotheses. Using S&P500 and Google's SVI data filtered by U.S. states they define a measure of abnormal asymmetric attention capturing unusual patterns in attention allocation of locals compared to non-locals. If abnormal asymmetric attention increases, this means local investors are paying an unusual amount of attention to local stocks. However, similar unusual behaviour isn't observed amongst non-locals. Further Cziraki et Al. (2017) test whether the changes in prices are temporary as the positive price-pressure suggests, or, whether the changes are more permanent as the gradual diffusion of local information suggests. Contrary to Da et Al.'s (2011) findings, the authors find no such reversal in prices, thus, concluding the gradual diffusion of local information to hold for their data. The general findings of the study are that firms that generate abnormally high asymmetric attention from local, compared to, non-local investors earn higher returns. These higher returns are generated due to information frictions and are especially prominent for stocks of firms located in remote areas, for stocks with high bid-ask spreads, as well as, stocks with a higher degree of analyst forecast dispersion[7].

One limitation of the Cziraki et. Al (2017) paper lies in using monthly SVIs instead of weekly SVIs as used in Da et. Al (2011). One of my contributions is therefore correcting for this, by using weekly data and generating a comparison between local and non-local SVIs that is more comparable with results of Da et. Al (2011). In addition, I take a slightly different perspective of studying whether the gradual diffusion of local information holds. I also study what drives non-local searches: is it just the past non-local searches, or, do local searches also play a role?

3 Theory & Hypotheses

As previously mentioned, this paper focuses on studying whether investor attention, measured by SVIs, can be used to predict stock market movements within S&P500 and FTSE100 indices on an industry level. Another aspect of this paper is to study whether there are any differences in investment behaviour of US and UK investors in their level of home biasness. In addition, to study differences in regards to the effect of their SVIs on local markets relative to non-local markets. The hypotheses chosen for this paper and their theoretical framework are explored in the successive text.

3.1 Hypotheses

The previous research (Da et. Al, 2011; Lundström and Nestius, 2012; Cziraki et. Al, 2017) has concluded SVIs to be a suitable proxy for retail investor attention. This is due to the fact that retail investors do not generally have access to more sophisticated information outlets, such as Bloomberg and Eikon terminals. Further, Barber and Odean (2008) discuss the limitations facing retail investors when making investment decisions. Since most retail investors are unable to short sell their investment, decisions regarding the selling of assets is constrained to the assets they already own; thereby, making information search mostly redundant. Therefore, the active information search regarding investment opportunities is focused on buying decisions. But even here an

[7] The higher the degree of analyst forecast dispersion the more variation can be found in investment analysis and recommendation for a specific asset across analysts.

individual investor faces some constraints. Since time is a scarce resource and the amount of information available is too much for an individual to process manually, I believe retail investors to be rather selective and biased in the companies they pay attention to. Further, I believe each individuals' environment, and (social) media in particular, play a significant role in which companies a retail investor is exposed to and pays attention to the most. Therefore, my first hypothesis is the following:

1. *Predictive ability of SVIs is more prominent for industries, which include more of companies with strong brands or generating a lot of media attention.*

This means that, industries such as Financials, Industrials and Technology could generate results with most significance. This is due to the fact that, investors would be quite familiar with companies that belong to said industries, via, the cars and electronic devices they own or the financial institution they have their assets with. Studying both the US and UK markets and investors, also allows for the study of whether there are differences in which industries grab the most attention amongst each of the groups. For example, one could assume Technology to be an especially interesting industry within S&P500 since companies, such as Apple, Facebook, Google, PayPal and eBay tend to generate a lot of media attention. Similarly, Financials might be a popular industry within the FTSE100 due to the industry's significant impact on UK economy (London in particular). In addition, the study regarding this hypothesis would reveal whether US and UK investors pay attention to the same industries, whether the popularity of an industry among investors is defined by the popularity of that industry on the domestic index, or, whether there are significant differences in which industries grab attention when investing domestically compared to internationally.

Discussing differences in US relative to UK investors offers a perfect segue to the discussion regarding the second hypothesis. As French and Poterba (1991) showed, there are differences in how much is invested domestically when comparing portfolios of US and UK investors. The larger size of the U.S. economy, for example, offers more diversification opportunities domestically leading to US investors being more home biased than UK investors. Whether this home bias holds when home bias is measured in SVIs is also interesting, and, therefore the second hypothesis is as follows:

2. *US investors are more home biased than UK investors measured in SVIs within domestic and foreign indices.*

For this assumption to hold, I expect the higher home bias to be directly linked to lower SVIs for US investors searching for FTSE100 industries, when compared to, the volumes generated by UK investors searching for FTSE100 industries. I also expect the SVIs of both the US and UK investors to be somewhat similar for the industries within S&P500. Further, this hypothesis enables the study of whether home bias has been reduced due to globalisation throughout the years. French and Poterba (1991) find the amount invested into domestic equities for US investors to be 94%, while UK investors have 86% of their equities invested domestically. On the other hand, a more recent paper by Mishra (2015) finds differences in volumes of home bias of US relative to UK investors to be less prominent. Further, the amount of equities invested domestically tends to be around 60% across the different models used. Therefore, I will also study whether my results are more in line with Mishra's (2015) findings than that of French and Poterba's (1991).

While it is interesting to study differences in home bias within two markets, it is important that said markets are somewhat similar, in order to make reliable comparisons. In the case of US and UK investors, comparisons between the regions are easier to make since these countries do not have any language barriers and are culturally similar. As, Tesar and Werner (1995) find, investors tend to invest into countries of geographical proximity. Further, Mishra's (2015) findings show that US investors are most comfortable in investing into UK out of all European countries. Similarly, French and Poterba (1991) find that a third out of all foreign investments of UK investors is invested into US equities. Therefore, one can expect differences between investment behaviours of US and UK investors, but, the similarity of their respective economies reduces the amount of variables to account for; thereby, simplifying the study of causalities.

The third hypothesis combines the two previous hypotheses and studies both the ability of SVIs in predicting market movements as well as aiming to capture differences in the investment behaviour of US and UK investors. The following hypothesis is tested:

3. *SVIs for S&P500 industry have predictive power over the same industry within FTSE100 and vice versa.*

In order to grasp the reasoning behind the third hypothesis, it is important to understand why SVIs are thought to be able to predict stock market movements and why this movement is expected to be positive. Lundström and Nestius (2012) define three reasons why SVIs are a plausible tool for predicting positive stock price movements.

Firstly, search engine data is a good measure of investor attention. Investors will often search financial information on companies they are interested in investing and thus increased search queries for a specific firm can be interpreted as an increased investor attention in that firm. Specifically, as Da et Al. (2011) have argued, the search query data captures the attention of retail investors, since more sophisticated institutional investors tend to use other financial information outlets, such as Bloomberg and Eikon terminals. Further, Lundström and Nestius (2012) expect a time delay in increased search queries and increased investor attention being translated into asset prices. Therefore, the search engine data can be used as a predictive tool in analysing future stock price movements, volatility and liquidity. Secondly, the authors believe Google's search engine to be the best proxy for search engine data due to its wide use. Google's market share has been around 80% for a decade. Moreover, Google especially is a good proxy for search engine data in the context of studying the behaviour of UK and US investors. As of 2017 Google's market share has been around 78% in US and 83% in UK, which is a significant difference from its competitors Yahoo (2.33%) and Bing (2.75%)[8]. Thirdly, Lundström and Nestius (2012) assume increased investor attention to be observed as a positive price pressure. This assumption is based on Barber and Odean's (2008) findings that propose retail investors to be mainly net buyers when it comes to investing. This is due to the restrictions in short selling, which leads to retail investors doing active research on stocks when wanting to buy them. On the other hand, when selling, retail investors will tend to sell the stocks they already own, thus making active information search somewhat redundant. Based on these three reasons I also believe Google's SVIs being a good measure of retail investor attention and can be used in predicting positive price movements, increased trading volume and volatility.

Further, it is interesting to note the contradictory results of Chen's (2017) paper finding increased SVIs to have a negative effect on returns as well as seeing these results to be more prominent for US investors out of all regions studied. My data differs slightly to that of Chen's (2017) since I study regional indices instead of a global one. However, I am still curious of whether my results indicate an expected positive or, surprisingly, negative price pressure. Further, Cziraki et. Al (2017) challenge Da et. Al's (2011) results of temporary price pressure and see evidence of the effects being permanent. The

[8] Market share of search engines, http://gs.statcounter.com/search-engine-market-share

data of this paper is somewhat similar to that of Cziraki et. Al's (2017) as I also combine the study of predictive ability of SVIs with the study of the information processed by local and non-local investors. Though, contrary to Cziraki et. Al's (2017) data being in monthly frequencies, I have followed Da et. Al's (2011) example in choosing weekly frequencies. Therefore, I am able to develop a further understanding of whether the impact of SVIs on the stock market is temporary or permanent.

The other part of the hypothesis lies in testing the interconnectedness of the S&P500 and FTSE100 indices. The following two examples show why the assumption of the markets being connected in their movements can be made. Firstly, we live in a globalised world where no region, industry or even company functions in a vacuum without an effect on others. For example, the data breach[9] that occurred to a S&P500 constituent Equifax didn't only affect Equifax but also Experian, Equifax's FTSE100 comparable. This breach resulted in a similar drop in stock prices for both companies around the incident, as Figure 1 shows.

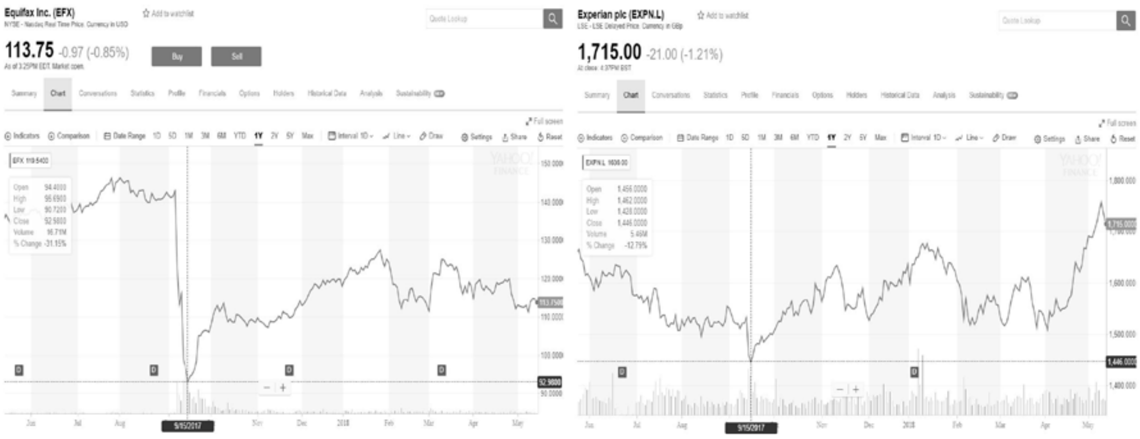

Figure 1: Equifax and Experian Comparison during the Equifax Data Breach Incident in September 2017

Similarly, an industry wide change can affect all of the companies within the industry, as the dot-com bubble of 2001 proved[10], or even the global financial markets as a whole, as in the financial crisis of 2008[11]. Secondly, many of the industries and the underlying companies have significant businesses in both UK and US. For example in Financials, most of the constituents for both S&P500 and FTSE100 have concentrated their American and European businesses within the 'financial hubs' of the regions (New York

---

[9] Information on the Equifax data breach, https://www.ftc.gov/equifax-data-breach
[10] Effects of dot-com bubble on IT stocks, http://money.cnn.com/2000/11/09/technology/overview/
[11] Article on 2008 financial crisis, https://www.theguardian.com/business/2008/dec/28/markets-credit-crunch-banking-2008

and London). Therefore, the performance of these regions has a vital impact on each of the financial institutions as a whole, regardless of whether they are listed of the S&P500 or FTSE100.

Finally, studying both of the aforementioned indices enables the ability to capture differences in behaviour of local investors relative to non-local investors. The fourth hypothesis is as follows:

4. *Local investors are exposed to and thus process new information before non-local investors.*

This hypothesis is based on the papers of Hong and Stein (1999) as well as Cziraki et Al. (2017) discussing how local information is believed to diffuse gradually. The authors of the papers found a delay in non-local investors processing information on a company compared to local investors. I aim to study this sc. gradual diffusion of local information hypothesis further by testing whether there are delays in US investors processing information on FTSE100 industries and/or delays in UK investors processing information on S&P500 industries. More specifically, I believe changes in SVIs of US investors on an S&P500 industry to be translated with a delay to SVIs of UK investors on the same industry. Similarly, I expect changes in searches of UK investors on a FTSE100 industry to translate with a delay to the amount of searches of US investors for that industry. Further, based on the results of Cziraki et Al. (2017), I believe this information asymmetry to be more prominent for smaller companies. This will be tested by including tests on the market capitalisation level, to compliment previous results gained from the study conducted on an industry level. The specifics of how the fourth and the previously mentioned hypotheses are tested, is explored as follows in Section 5.

4 Data

Due to Google Trends providing weekly Search Volume Index dating back five years, I study the effects of search engine data on the US and UK markets over the period of 2013/03/10 – 2018/03/10. A data set with weekly frequency has been chosen to maintain comparability with previous research, such as Da et. Al (2011) which also uses weekly data. I have collected both stock measure and search volume index data for the S&P500 and FTSE100 indices. The following section provides a more in-depth description of the data collection process for each of the datasets – stock market and Google Trends data.

4.1 Stock Market Data

This paper is focused on studying effects found on an industry level. Therefore, I collect historical stock measure data for industries rather than for each constituent within the indices. Most of the data is collected from *Yahoo! Finance* with the exception of FTSE100, UK 3-month gilt and US 3-month T-bill data, which have been collected from *Investing.com*.

For the S&P500 data I find Select Sector SPDR ETFs[12] to be a fitting proxy for this purpose, due to the fact that the ETFs are built using a widely accepted industry division and have a high trading volume. More specifically, the ETFs are built to replicate the movements of the S&P500 index while being divided into ten industries: Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Materials, Technology, Utilities and Real Estate. Each underlying constituent within industries is weighted according to its weight on the S&P500 index. Unfortunately, the ETF for Real Estate doesn't cover the full five-year study period and thus the whole industry is dropped from further analysis reducing the initial S&P500 data from 500 companies to 473. Therefore, the data represents nine of the aforementioned industries after the drop.

The FTSE100 data is collected in a similar fashion. Since I didn't find any ready-made industry ETFs for the purpose of my study I have created my own using the same industry division as used for the SPDR ETFs. The sector for each company has been derived from London Stock Exchange's website and then allocated into an industry according to ICB classification[13], which FTSE Russell follows. Although, I have followed the ICB guidelines for the most part, there are few exceptions, which are more in line with SPDR's division of the ETFs. Even though ICB classification divides Technology and Telecommunications into two separate sectors, the ETFs used have combined these companies into one Technology sector and therefore I have followed the latter division. Further, I found few companies for which industries could not be defined using the ICB classification. For said companies, I have found the industry

[12] Select Sector SPDR, http://www.sectorspdr.com/sectorspdr/tools/sector-tracker

[13] ICB Classification, http://www.ftserussell.com/financial-data/industry-classification-benchmark-icb

classification for a comparable company within S&P500 and allocated the FTSE100 equivalent into the same industry. For example, Experian has been classified into Industrials due to the fact that its comparable, Equifax, is classified within Industrials in the SPDR ETF framework. The weights for each constituent have been derived from the latest UKX Quarterly Data publication[14]. As per S&P500 data, Real Estate has been dropped from analysis reducing the initial data of 100 companies to 96. Table 1 presents the descriptive statistics for the dataset.

Table 1: Descriptive Statistics of the Sample Properties

		Total Sample	Consumer Discretionary	Consumer Staples	Energy	Financials	Health Care	Industrials	Materials	Technology	Utilities
No. of Constituents	Total	563	102	43	35	86	67	84	36	76	33
	S&P500	473	82	34	31	69	62	70	25	72	28
	Large Cap	17		11	6						
	Mid Cap	33		15	18						
	Small Cap	14		8	6						
	Total	64									
	FTSE100	90	20	10	4	17	5	14	11	4	5
	Large Cap	16		3		6	3	1	3		
	Mid Cap	29		4		8	1	10	6		
	Small Cap	12		3		3	1	3	2		
	Total	57									
No. of Observations	Total	4,968	522	522	522	522	522	522	522	522	522
	S&P500	2,349	261	261	261	261	261	261	261	261	261
	FTSE100	2,349	261	261	261	261	261	261	261	261	261

Table 1 is a summary of descriptive statistics for the sample in total, for each of the industries (Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Materials, Technology and Utilities) as well as for each capital size (Large, Small and Mid Cap) for industries this is calculated for (FTSE100: Consumer Staples, Financials, Health Care, Industrials and Materials; S&P500: Consumer Staples and Energy) over the five-year sample period 2013/03/10 -2018/03/10. *No. of Constituents* describes the total number of firms within each of the industries as well as capital sizes as of 2018/04/20 for S&P500 and as of 2017/12/18 for FTSE100. *No. of Observations* describes the total number of weeks observed for each of the industries within an index as well as in total.

4.2 Google Trends Data

The Google search engine data is derived from the Google Trends service. Since I have used ETFs as a proxy for S&P500 industries, the SVI data is collected for each of the nine ETFs used. For FTSE100, data has been collected for each of the underlying companies and transformed into industry SVIs by weighting each company SVI according to its FTSE100 weight. In order to perform my analysis, three types of SVI data are required – global, US and UK searches. To avoid downloading each file manually, I have written a piece of code using the programming language *R*, which automatically scrapes the data for each search term in a predefined list.

Although Da et Al. (2011) find tickers to be the least noisy type of search term; I find this not to hold for my data. Especially when looking at UK searches of S&P500 constituents or US searches of FTSE100 constituents. Instead of following Da et. Al’s (2011) practise, I look at ‘related topics’ and ‘related queries’ generated by Google

[14] FTSE Russell publicises weight calculations on each FTSE100 constituent on a quarterly basis. Such report can be found at, https://www.ftse.com/analytics/factsheets/Home/DownloadConstituentsWeights/?indexdetails=UKX

Trends and choose search terms that return related topics/queries, such as other FTSE100/S&P500 constituents, financial terms etc. (terms that can be interpreted as investment related). Same search terms for global, US and UK searches are chosen to maintain comparability of results. I also find that searching for a term as a 'search term' or as a 'company' outputs different SVIs. For example, searching for 'Whitbread' as a search term or as a company returns different SVI values, as can be seen in Figure 2.

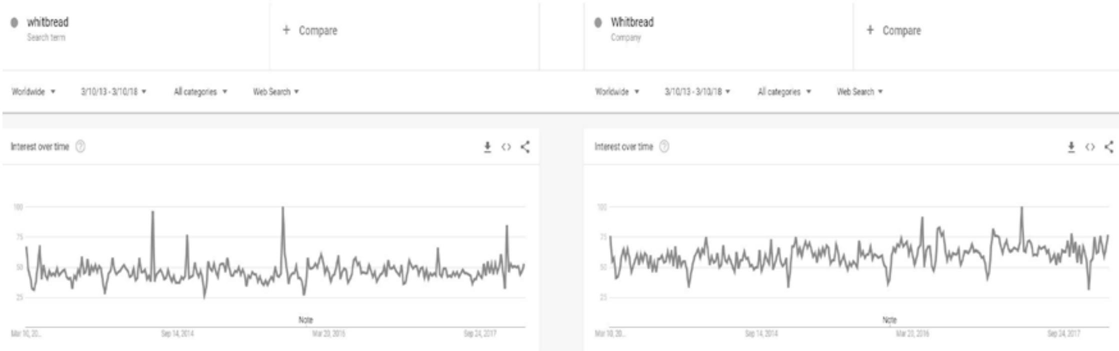

Figure 2: Differences in SVIs Generated for Same Word Seaches as a Search Term vs. as a Company

Therefore, I use manual searches 'as a company' as a last resort for constituents that do not generate SVI data in other ways. After these operations there are a few companies that are missing data and therefore are dropped. These drops result in the final data consisting of 90 FTSE100 companies divided into nine industries, as well as, nine ETFs representing 473 constituents within the S&P500 divided into industries. The full list of search terms can be found in Appendix B.

Lastly, for the second part of testing in relation to the fourth hypothesis, I collect SVIs and weights for the constituents of Consumer Staples and Energy SPDR ETFs. Such volumes and weights represent the market capitalisation of both UK and US searches. This results in a dataset of 64 companies, which are further divided into Large, Mid and Small Cap with top 25% denoted as Large Cap, 50-75% as Mid Cap and bottom 25% as Small Cap. This division is done based on the weights of each of the constituents on S&P500 index. The same procedure is performed on the FTSE100 companies using Consumer Staples, Energy, Financials, Industrials and Materials resulting in a dataset of 57 companies. As per stock market data, the descriptive statistics for the dataset are presented in the Table 2.

Table 2: Descriptive Statistics, Search Volume Index (SVI)

(a) Total

	Mean	Std.Dev.	Min	Max	75th Percentile
S&P500, US Searches	27.699	9.599	9.954	42.750	33.015
S&P500, UK Searches	20.948	7.354	7.797	28.743	25.395

S&P500, Global Searches	44.527	11.616	30.897	67.912	43.969
FTSE100, US Searches	8.660	4.944	1.987	18.419	10.448
FTSE100, UK Searches	31.604	6.982	20.904	43.009	33.530
FTSE100, Global Searches	39.723	4.611	33.370	46.460	46.034

(b) Average SVIs by Industry

Industry	(1) S&P500, US Searches	(2) S&P500, UK Searches	(3) FTSE100, US Searches	(4) FTSE100 UK Searches	(5) S&P500, Global Searches	(6) FTSE100, Global Searches
Consumer Discretionary	37.364	25.337	5.412	20.905	43.521	33.370
Consumer Staples	42.751	25.395	3.904	26.673	54.843	43.969
Energy	33.015	28.743	10.448	27.159	67.912	39.624
Financials	19.755	18.234	12.812	43.009	32.797	46.461
Health Care	25.958	21.161	8.753	33.530	45.540	37.200
Industrials	9.954	26.759	8.992	28.808	44.594	35.121
Materials	25.812	24.471	18.419	30.636	46.035	40.714
Technology	28.751	10.640	1.987	40.966	34.605	44.794
Utilities	25.931	7.797	7.209	32.752	30.897	36.250

Table 2 is a summary of descriptive statistics for the Search Volume Index (SVI) data. Table (a) Total presents the mean, standard deviation, min, max and 75th percentile values for each SVI data subset (S&P500 US, UK and global searches; FTSE100 US, UK and global searches) in total. Table (b) Average SVIs by Industry presents each of the aforementioned descriptive statistics broken down by industry for each of the SVI data subset.

5 Analytical Framework

As previously discussed, I aim to study the predictive power of SVIs on stock market returns, volatility and liquidity as well as recognise differences in investment behaviour of US respective UK investors. This is achieved by analysing stock measure and search engine data for S&P500 and FTSE100 constituents on an industry level. Studying the power of SVIs in 2018 gives a few advantages compared to the earlier studies. Google has updated and added features to its Google Trends service, which creates an opportunity to have a deeper understanding of the effect of search engine data on the financial markets. Da et. Al (2011) defined ticker symbols as the least noisy measure capturing investor attention, but, the current feature of seeing related topics and queries shows that this isn't necessarily always the case. Further, Lundström and Nestius (2012) study the topic using company names but acknowledge the method's possibility of capturing attention of non-investors since there aren't any ways to filter objectives behind the searches. This is also no longer the case with the aforementioned features added to Google Trends. Therefore, it is possible to reduce the noise of search terms using several tactics, such as ticker symbols, looking at related topics etc.

The analysis is built on using abnormality in variables rather than studying absolute changes. Specifically, I will analyse whether changes in the amount of search queries for a specific asset lead to changes in abnormal returns, volatility and liquidity.

I follow Da et Al.'s (2011) example in defining abnormality. Thus, abnormal value for a specific variable is defined as the difference between the actual and expected values. To ensure robustness of results I implement different ways of calculating expected values. Expected value for stock returns is calculated with the help of CAPM. For datasets representing the US market the market portfolio used is the S&P500 index and the risk-free rate is defined as 3-month US T-bills. Similarly, for the UK market the FTSE100 index is chosen as a proxy for the market portfolio, and, 3-month UK gilts are used as a proxy for the risk-free rate. Further, the expected SVIs are defined using two different approaches. The first approach is to calculate the expected values using the median of the prior eight weeks, whilst the second approach, is to calculate the mean of the prior eight weeks. Thus, two datasets will be obtained when studying abnormal returns and SVIs. Expected liquidity and volatility are defined in a similar manner to expected SVIs calculating the mean of the prior eight weeks. I start by presenting the analytical framework and regression models used in testing each of the hypotheses, after which I discuss in more depth how the variables used, are defined.

5.1 The Analytical Framework and Regression Models

5.1.1 H1 – Top Industries Generate Most Significant Results

As mentioned in Section 3, the popularity of an industry among investors is based on the average SVIs the industry generates. More precisely, I define the top industries in the eyes of investors to be the industries in the 75th percentile (top 25%) of the SVI averages over the five-year period. These percentiles are calculated for both the US and UK searches for S&P500 and FTSE100 industries, as well as, for global searches for the industries within the indices. I test six different sub-hypotheses of whether the majority of the industries within the 75th percentile is, de facto, the industries generating highest results in third and fourth hypotheses. Thus, I see enough evidence for H1 to hold, if, the majority of the aforementioned sub-hypotheses seem to hold.

5.1.2 H2 – US Investors Are More Home Biased than UK Investors

For the second hypothesis I perform differences in means tests on the five-year average SVIs generated by local and non-local investors for each of the industries. This means that I will be testing whether the differences in average SVIs generated by local relative to non-local investors are statistically significant:

$$H_o: \overline{SVI}_{US} = \overline{SVI}_{UK} \qquad H_a: \overline{SVI}_{US} \neq \overline{SVI}_{UK} \tag{1}$$

If there is evidence for H2, I would expect the null hypothesis of no differences in means to be rejected when comparing average SVIs of UK and US investors on FTSE100. That is, the average SVIs *are* different when comparing US and UK investors on FTSE100. Further, I expect the average SVIs of US investors to be lower than that of UK investors. On the other hand, I do not expect the null hypothesis of, no differences in means, to be rejected when studying the SVIs for US and UK investors on S&P500. That is, the average SVIs *are not* different when comparing US and UK investors on S&P500.

5.1.3 H3 – SVIs of S&P500 Predict FTSE100 and Vice Versa

To test for the third hypothesis, I define three regression models that aim to capture effects of global S&P500 and FTSE100 abnormal SVIs (ASVIs) on abnormal returns, liquidity and volatility for both indices. These models are built on the abnormal values of each of the measures defined later in this section.

When studying the effects on abnormal returns for industry *i* at time *t* the regression model is defined as

$$AR_{i,r,t} = \beta_0 + \beta_1 * ASVI_{i,1,t-1} + \beta_2 * ASVI_{i,2,t-1} + \cdots + \beta_5 * ASVI_{i,1,t-3} + \beta_6 * ASVI_{i,2,t-3} + \varepsilon_{i,r,t} \tag{2}$$

where $AR_{i,r,t}$ is the abnormal return for industry *i* within index *r* at time *t* and *r* = [1,2] denoting abnormal SVI for S&P500 industry with 1 and abnormal SVI for FTSE100 with 2. Thus, the model studies the effects of abnormal SVIs for industries within both indices, with lags up to three weeks.

The regression models for abnormal liquidity and volatility follow similar fashion:

$$ALTV_{i,r,t} = \beta_0 + \beta_1 * ASVI_{i,1,t-1} + \beta_2 * ASVI_{i,2,t-1} + \cdots + \beta_5 * ASVI_{i,1,t-3} + \beta_6 * ASVI_{i,2,t-3} + \varepsilon_{i,r,t} \tag{3}$$

$$AV_{i,r,t} = \beta_0 + \beta_1 * ASVI_{i,1,t-1} + \beta_2 * ASVI_{i,2,t-1} + \cdots + \beta_5 * ASVI_{i,1,t-3} + \beta_6 * ASVI_{i,2,t-3} + \varepsilon_{i,r,t} \tag{4}$$

where $ALTV_{i,r,t}$ is the abnormal liquidity for industry *i* within index *r* at time *t* and $AV_{i,r,t}$ is the abnormal volatility for industry *i* within index *r* at time *t*. The results of these regressions on both the median and mean approach are presented in Appendix A and main findings discussed in Section 6.2.

5.1.4 H4 – Local Investors Process Information before Non-locals

To test for the fourth hypothesis, I define two regression models that aim to capture effects of UK (non-local) and US (local) searches on S&P500, and, effects of US (non-local) and UK (local) searches on FTSE100. These models are built on the abnormal SVI variables defined later in this section. The abnormal non-local SVI for industry *i* at time *t* is defined as

$$ASVI_{non-local,i,t} = \beta_0 + \beta_1 * ASVI_{non-local,i,t-1} + \beta_2 * ASVI_{local,i,t} + \beta_3 * ASVI_{local,i,t-1} \qquad (5)$$

where $ASVI_{non-local,i,t}$ is the abnormal SVI for non-local searches for industry *i* at time *t* and $ASVI_{local,i,t}$ is the abnormal SVI for local searches for industry *i* at time *t*. For the hypothesis to be rejected only past non-local SVIs should have an effect on this week's non-local SVIs.

The second part of the hypothesis, regarding whether the gradual diffusion of local information is more pronounced for smaller companies measured in market capitalisation, is performed using same regression model where *i* denotes the market capitalisation of the company. The results of all of these regressions on both the median and mean approach are presented in Appendix A and main findings discussed in Section 6.3. To understand the presented regression models further, the following section delves into a more in-depth discussion concerning the definitions of each of the variables.

5.2 Abnormal Search Volumes

Abnormal logarithmic value of a variable at time *t* is defined as the difference between the actual and expected value at time *t*. Using abnormal values gives an opportunity to take a closer look at the effect of SVIs on the stock market in those moments where the search volumes deviate from what is expected as well as adjusting for time trends and other seasonalities. Therefore, the abnormal SVI for an industry *i* at time *t*, $ASVI_{i,t}$, is defined as:

$$ASVI_{i,t} = log(SVI_{i,t}) - log[E(SVI_{i,t})] \qquad (6)$$

In words, the abnormal SVI for an industry *i* at time *t* is defined as the difference between the actual and expected logarithm of SVI for industry *i* at time *t*. To allow for robustness checks the expected SVIs are defined using two different ways – the median and mean approaches.

5.1.1 Median Approach

The first method used, the median approach, defines the expected SVI as the median value of observed SVIs for the prior eight weeks:

$$E(SVI_{i,t}) = median(SVI_{i,t-1}, SVI_{i,t-2}, \dots, SVI_{i,t-8}) \tag{7}$$

The advantage of this approach lies in its ability to smooth down the effects of a single observation.

5.1.2 Mean Approach

The second method, mean approach, defines the expected SVI as the mean value of observed SVIs for the prior eight weeks:

$$E(SVI_{i,t}) = mean(SVI_{i,t-1}, SVI_{i,t-2}, \dots, SVI_{i,t-8}) \tag{8}$$

The mean value is more sensitive to single extreme values, therefore, one-week SVI peaks (drops) have the potential to skew the expected value upwards (downwards). Thus, the focus of my analysis will be interpreting results generated using the median approach. Nevertheless, the results of the mean approach serve a useful purpose in acting as a suitable robustness check.

Following Da et. Al's (2011) example, I see the eight-week period being an optimal choice between the trade-off of accounting for seasonality and smoothing out any effects. Since SVI is defined as an index returning a value of 0-100 cross-sectional analysis of abnormal SVI can be performed regardless of the industry size. The final definition of abnormal SVI is thus a combination of (6) and (7) as well as (6) and (8):

$$ASVI_{i,t} = log(SVI_{i,t}) - log[median(SVI_{i,t-1}, SVI_{i,t-2}, \dots, SVI_{i,t-8})] \tag{9}$$

$$ASVI_{i,t} = log(SVI_{i,t}) - log[mean(SVI_{i,t-1}, SVI_{i,t-2}, \dots, SVI_{i,t-8})] \tag{10}$$

5.3 Abnormal Returns

The aim of this paper is to study the effects of abnormal search volumes on abnormal weekly industry returns. The returns for the S&P500 industry ETFs are the actual returns captured by the market, while, the returns for the FTSE100 industries have been estimated as the returns of each of the constituents multiplied with its weight on the FTSE100 index. The abnormal weekly industry returns are defined as:

$$AR_{i,t} = r_{i,t} - E(r_{i,t}) \quad (11)$$

In words, the abnormal weekly return for industry *i* at time *t* is defined as the difference between the observed and expected return. The expected return is defined using the CAPM expression and thus is interpreted as a sum of risk-free rate and the market premium multiplied with each industry's sensitivity to systematic risk denoted as β:

$$E(r_{i,t}) = r_{f_{i,t}} + \beta_{i,t} * (r_{m_{i,t}} + r_{f_{i,t}}) \quad (12)$$

The market return $r_{m_{i,t}}$ is defined as the return for the S&P500 index for U.S. industries and as the return for the FTSE100 index for the UK industries. The risk-free rate is usually defined using either LIBOR or government bonds of short duration. I refrain from using LIBOR due to its notorious nature[15] and define the risk-free rate as returns on 3-month T-bills for the US industries and as returns on 3-month gilts for UK industries. The industry betas, $\beta_{i,t}$, are estimated by using a rolling regression on industry and market excess returns with a one-year window.

5.3 Abnormal Liquidity

I aim to study the effects of abnormal search volumes on abnormal weekly industry liquidity. I have decided to use trading volume as a proxy for liquidity, which is seen as a suitable proxy, as results of Datar et. Al (1998) suggest. The trading volumes for the S&P500 industry ETFs are the actual volumes captured by the market, while, the trading volumes for the FTSE100 industries have been estimated as the trading volumes of each of the constituents multiplied with its weight on the FTSE100 index. As referred to previously, the abnormal value is defined as the difference between the logarithmic actual and expected values, where expected trading volume is defined as the mean value of the prior eight weeks:

$$ALTV_{i,t} = log(tv_{i,t}) - log[mean(tv_{i,t-1}, tv_{i,t-2}, \dots, tv_{i,t-8})] \quad (13)$$

In words, the abnormal trading volume for industry *i* at time *t* is the difference between the logarithm of industry *i*'s trading volume at time *t* and the logarithmic mean of trading volumes of the past eight weeks.

---

[15]Article on the LIBOR scandal, https://www.cfr.org/backgrounder/understanding-libor-scandal

5.4 Abnormal Volatility

In addition, I study the effects of abnormal search volumes on abnormal weekly industry volatility. As a proxy for volatility I use the high-minus-low measure, which was popularised by Parkinson (1980) and is still a widely accepted proxy for volatility, as Goyenko et. Al (2009) find. The difficulty in defining volatility lies in choosing an interval volatility is estimated upon, since basing the estimations on too long of a time period will smooth out any distinctive effects. As defined previously, I choose to use the same eight-week time period, thus, the proxy for weekly volatility for an industry *i* at time *t* is defined as the mean of intra-day high-minus-low values of the prior eight weeks

$$\hat{\sigma} = mean(ph_{i,t-1} - pl_{i,t-1}, ph_{i,t-2} - pl_{i,t-2}, \dots, ph_{i,t-8} - pl_{i,t-8}) \tag{14}$$

where $ph_{i,t-n}$ is the highest price for industry *i* at time *t-n* and $pl_{i,t-n}$ is the lowest price for industry *i* at time *t-n*. As previously, the highest and lowest prices for the S&P500 industry ETFs are the actual prices captured by the market while highest, and, lowest prices for the FTSE100 industries have been estimated as the prices for each of the constituents multiplied with its weight on the FTSE100 index. The abnormal volatility is thus defined as:

$$AV_{\boldsymbol{i,t}} = log(ph_{i,t} - pl_{i,t}) - log[mean\big(ph_{i,t-1} - pl_{i,t-1}, ph_{i,t-2} - pl_{i,t-2}, \dots, ph_{i,t-8} - pl_{i,t-8}\big)] \tag{15}$$

6 Empirical Results

6.1 US Investors Pay More Attention to the Domestic Equity Market

[Table 2]

To start off, I study the differences between US and UK searches within S&P500. Looking at averages presented in Table 2 for each of the industries, the averages for S&P500 do seem to be quite similar while there are differences in averages within FTSE100.

[Table A1]

The results from the differences in the means test, presented in Table A1, show no statistical significant differences between the averages of UK and US investors on S&P500. Therefore, the null hypothesis of no differences in means cannot be rejected. Further, the results of the differences in means test within FTSE100 return a p-value of 0.000, which means that the null hypothesis of no differences in means can be rejected at any confidence level. Furthermore, the results show the US investors' average SVIs to be lower than that of UK investors. Since both of the hypothesis return the expected results mentioned in Section 3, I find evidence of US investors to be more home biased compared to UK investors when measured in SVIs.

Interestingly, for Technology and Utilities the SVIs of US searches are over double the SVIs of UK searches, which would suggest that home bias of UK investors is especially strong within these industries. The high SVIs of UK searches on FTSE100 for the same industries further strengthen this assumption. Additionally, the average SVI for UK searches on S&P500 industrials is over double the SVI of US searches on the industry. This seems to be explained by the increased interest of UK investors in the industry, since the average SVI of UK searches for FTSE100 Industrials and Technology are of the same level, whereas, the SVIs for US searches are comparatively low on both indices. A similar effect can also be found for US investors regarding Consumer Staples and Technology. The average SVIs of US investors for said industries are significantly higher on S&P500 than they are on FTSE100.

6.2 FTSE100 SVIs Are Able to Predict Movements of S&P500

The aim of the third hypothesis is to study whether SVIs for industries within S&P500 have the ability to predict FTSE100 industry abnormal returns, liquidity as well as volatility and vice versa. The regression results seem to suggest that while SVIs for industries within FTSE100 do predict movements of S&P500 industries, there isn't much of statistically significant evidence of S&P500 industries predicting FTSE100 industry movements. Next, I discuss further the results for each of the abnormal values. As previously discussed, the analysis focuses on median values while mean values are presented as a robustness check in Appendix A.

6.2.1 Abnormal Returns

[Table A2]

As can be seen in Table A2, none of the S&P500 industry SVIs generate statistically significant results when studying the predictive power over abnormal returns for FTSE100 industries. Further, even SVIs for FTSE industries do not seem to generate any statistically significant results.

[Table A3]

This isn't the case when looking at the results generated for the S&P500 industries (Table A3). For Financials, last week's S&P500 ASVIs predict a 3.0 basis point (bps) increase in abnormal returns at a 1% significance level. Further, FTSE100 SVIs from two weeks ago predict a 2.2bps increase in abnormal returns at a 10% significance level. For Industrials, S&P500 ASVIs from three weeks ago decrease the abnormal returns by 1.1bps (10% significance level), whilst, the S&P500 ASVI for Materials with a three-week lag reduces abnormal returns by 0.7bps; at a 5% significance level. Lastly, S&P500 ASVI for Utilities with a two-week lag increases the abnormal returns by 0.7bps at a 10% significance level. Thus, I can conclude that abnormal returns are mostly affected by ASVIs on the industry within the studied index, as opposed to, ASVIs of the same industries within some other index. Interestingly, Financials are affected by both ASVIs for S&P500 and FTSE100. A possible explanation for this could be the heightened impact of the financial industry on FTSE100 and the UK economy as a whole.

6.2.2 Abnormal Liquidity

[Table A4]

Table A4 presents the results for predictive power of ASVIs on abnormal liquidity. Last week's FTSE ASVIs on Consumer Staples and Financials, increase the abnormal liquidity for the industries by 91.5bps and 97.3bps respectively, both at a 1% significance level. Further, FTSE ASVIs with a two-week lag decrease abnormal liquidity by 38.2bps at a 10% significance level. Additionally, last week's FTSE ASVIs have a positive effect on abnormal liquidity for Industrials, Materials, Technology and Utilities at a 1% significance level. The abnormal liquidity is increased by 70.3bps, 155.4bps, 64.6bps and 20.9bps respectively. FTSE ASVIs with a two-week lag reduces the abnormal liquidity by 81.8bps for Materials at a 5% significance level. Finally, last week's FTSE ASVIs increase abnormal liquidity for Health Care by 37.8bps (at a 10% significance level). S&P500 industry ASVIs have an effect on the FTSE100 Materials. Last week's ASVIs reduce the abnormal liquidity by 21.7bps, while, ASVI with a two-week lag has a negative effect of 15.6bps, at a 5% and 10% significance level respectively.

[Table A5]

When looking at the results for S&P500 industries (Table A5), S&P500 ASVIs with a two-week lag seem to increase abnormal liquidity for Health Care by 30.4bps and last week's ASVIs for Materials reduced the measure by 22.2bps. Interestingly, all other statistically significant effects on S&P500 industry abnormal liquidity are generated by FTSE100 ASVIs. Last week's ASVIs increase abnormal liquidity for Energy by 29.0bps (10% significance), by 63.3bps for Financials (5% significance), by 54.5bps for Health Care (10% significance), by 106.0bps for Materials (1% significance) and by 16.9bps for Utilities (10% significance). The ASVIs with a two-week lag seem to have a negative effect on abnormal liquidity. For Financials this effect is at 41.8bps (10% significance), for Health Care at 71.8bps (5% significance) and for Materials at 67.8bps (10% significance). The negative effect seems to continue when looking at ASVIs with a three-week lag with ASVIs having a -28.3bps effect on Technology at 10% significance level.

To conclude, FTSE100 ASVIs seem to generate more of statistically significant effects on S&P500 industries than S&P500 industries have on FTSE100 ones, even when abnormal liquidity is studied. Interestingly, last week's ASVIs tend to increase abnormal liquidity while two- and-three-week lags seemingly have a negative effect thus, partly reversing the previous liquidity increase.

6.2.3 Abnormal Volatility

[Table A6]

Finally, I take a deeper look at Table A6 that presents the results for abnormal volatility. As with abnormal liquidity, ASVIs on S&P500 industries generate statistically significant results for one industry only when looking at abnormal volatility within

FTSE100. S&P500 ASVIs with two-week lags have a positive effect of 20.0bps on Financials (10% significance level). Last week's FTSE ASVIs increase the abnormal volatility for Consumer Discretionary by 50.1bps (1% significance), for Financials by 49.7bps (10% significance), for Health Care by 99.7bps (1% significance), for Materials by 92.5bps (1% significance) and for Technology by 24.4bps (10% significance). When looking at FTSE ASVIs with two-or-three-week lags the negative effect aforementioned in the previous section, seems to prevail even within abnormal volatility. ASVIs with two-week lags reduce abnormal volatility with 56.4bps for Health Care (10% significance) and 48.0bps for Materials (10% significance). Similarly, ASVIs with three-week lags reduce the volatility by 62.6bps for Consumer Staples (5% significance) and with 28.3bps for Technology (10%).

[Table A7]

Looking at the results for S&P500 industries (Table A7), FTSE100 ASVIs have a profound effect. Last week's FTSE ASVIs increase abnormal volatility by 59.9bps for Materials (10% significance). Adding more lags seems to, yet again, have a negative effect. ASVIs with two-week lags reduce abnormal volatility with 66.5bps for Health Care and (5% significance) with 27.1bps for Utilities (5% significance). ASVIs with two-week lags reduce abnormal volatility with 53.2bps for Technology (5% significance).

In conclusion, abnormal search volumes for FTSE100 industries have an impact on abnormal returns, volatility, and liquidity for S&P500 industries. On the other hand, abnormal search volumes for S&P500 industries don't have an impact on FTSE100, apart from the effect found in liquidity within Materials and in volatility within Financials.

6.3 Information Diffuses within a Week

The fourth hypothesis states that I expect the ASVIs of non-local investors to follow those that of local investors, with a lag. As previously mentioned, the analysis of results in this section is based on the median values while results from mean values are presented in Appendix A.

6.3.1 Industries

[Table A8]

I start the analysis by looking at the results on an industry level, as presented in Table A8. For S&P500, where US investors are defined as local and UK investors as non-local, the non-local searches are affected by both this and last week's local searches. In Consumer Staples same week's US searches have a positive effect of 25.5bps while in Energy last week's US searches increase the UK searches by 10.6bps, both at 10% significance level. In Financials and Materials, the UK searches are affected by previous UK searches with 23.4bps increase in Financials and 21.3bps increase in Materials, both at 1% significance level.

[Table A9]

Next, I look at results for FTSE100 where UK searches are defined as local, and, US searches defined as non-local. These results are presented in Table A9. The local searches do seem to have an effect on non-local searches. For example, UK searches in the same week and last week on Consumer Staples seemingly have a positive effect in respect to US searches. Same week local searches increase non-local searches with 17.4bps while last week's local searches increase the non-local ones with 15.4bps, both at a 10% significance level. Further, same week local searches have a positive effect on non-local searches within Financials, Health Care, Industrials and Materials. For Financials, local searches increase non-local ones with 14.9bps (5% significance), for Health Care the increase is at 52.4bps (1% significance), for Industrials at 13.8bps (5% significance) and for Materials at 32.2bps (1% significance). Also, last week's non-local searches increase this week's non-local ones by 22.0bps in Financials (1% significance), by 14.1bps in Industrials (10% significance) and by 12.5bps in Materials (10% significance).

In conclusion, the results do show some evidence of gradual diffusion of local information since the non-local searches are not only affected by last week's non-local searches but also by this and last week's local ones. Further, the delay in diffusion of information seems to be less than a week since mostly same week's local searches affect non-local ones. Also, the effect of local searches on non-local ones is stronger within FTSE100, which is in line with the other part of the hypothesis that assumes the information asymmetry to be stronger in smaller and more 'remote' areas.

6.3.2 Market Capitalisation

[Table A10]

The second part of this hypothesis is to study whether the market capitalisation of a constituent has an effect on how statistically significant the results are. These results are presented in Table A10. I start with results for the S&P500, where US investors are deemed as local and UK investors as non-local. For Small Cap same week US searches have a positive effect of 42.9bps on UK searches with 5% significance. For Mid Cap same week's US searches increase UK searches by 40.9bps with 1% significance, and, for Large Cap local searches do not seem to affect non-local searches. Therefore, I find evidence for the hypothesis of smaller market capitalisation increasing information asymmetry. Whilst non-local searches are affected by local searches in both Small and Mid Caps, this effect is not present for Large Caps.

[Table A11]

Next, I analyse the results generated for FTSE100 where UK investors are deemed as local and US investors as non-local, presented in Table A11. Same week UK searches have a positive effect of 25.4bps on US searches for small caps at a 5% significance level. Interestingly, the effect on Large Caps is also statistically significant. Same week's local searches have a positive effect of 23.7bps at 1% significance level. This result goes against the hypothesis that Large Caps would be less prone to information asymmetry between locals and non-locals.

In conclusion, although there is some evidence of smaller market capitalisation increasing information asymmetry, there is also evidence for the opposite to hold. Furthermore, no statistically significant results are generated for last week's local searches and therefore it is impossible to say whether a delay of information actually exists or whether the delay is less than a week and therefore not captured by my data. Nevertheless, these regressions do generate some interesting results to be studied further, possibly with a more frequent dataset.

6.4 Familiarity Correlates Positively with Significance of Results

As mentioned in Section 5.1.1, H1 is tested by testing six sub-hypotheses. The results are derived from the output of H3 and H4 presented in previous sections, as well as, Appendix A.

I start with results for the US searches for S&P500 industries. The most familiar industries are Consumer Discretionary and Consumer Staples. Therefore, these industries should generate the strongest results for regressions discussed in Section 6.3. Although, results for Consumer Discretionary aren’t statistically significant, Consumer Staples generate some results at a 10% significance level. Therefore, I find some evidence for the first sub-hypothesis to hold. The second sub-hypothesis studies the UK searches within S&P500. For these, the most familiar industries are Energy and Industrials: if these generate statistically significant results the hypothesis holds. Looking at the results neither of these industries return any results of interest and therefore there isn’t enough evidence for this sub-hypothesis to hold.

Thirdly, I analyse the results for US searches within FTSE100. For these, the most familiar industries are Energy, Financials and Materials and therefore I expect statistically significant results for at least two of the industries for the sub-hypothesis to hold. As discussed in Section 6.3, both Financials and Materials generate results at 1% and 10% significance levels respectively. Thus, the third sub-hypothesis seems to hold. Similarly, I study the results for UK searches within FTSE100. Here, the most familiar industries are Financials, Health Care and Technology, for which, statistically significant results are expected. As Section 6.3 shows, Financials and Health Care generate results of interest at 5% and 1% significance level respectively.

Lastly, I move to Section 6.2 to test the last two sub-hypotheses of whether the most familiar industries for global S&P500 and FTSE100 searches generate the strongest results. I start this analysis with the FTSE100 data, for which, Financials is defined as the most familiar industry. As discussed in Section 6.2 Financials are able to predict S&P500 abnormal returns, liquidity as well as FTSE100 abnormal liquidity and volatility at a statistically significant level. Thus, I see enough evidence for the fifth sub-hypothesis to hold. Next, I look at results generated by S&P500 global searches. For these the most familiar industries are Consumer Staples, Energy, Health Care, Industrials and Materials. Out of these, only Materials and Health Care are able to predict FTSE100 and S&P500 abnormal liquidity respectively at a statistically significant level. Therefore, there isn’t enough evidence for the sixth sub-hypothesis to hold.

In conclusion, I see enough evidence in support of four out of the six sub-hypotheses. Therefore, the hypothesis of industries most familiar to retail investors generating the most significance in results seems to hold. The popularity of an industry measured by, average SVIs it generates, does seem to have a positive correlation to the significance of results.

7 Conclusion

The results of this paper suggest that US investors are more home biased than UK investors when measured in respect to SVIs. Further, I find evidence that global searches for FTSE100 industries are able to predict abnormal returns, liquidity and volatility of S&P500; but none for vice versa to hold. Interestingly, abnormal search volumes tend to have a positive effect on these measures when looking at data with one-week lags, but the effect turns negative with two- and-three-week lags. This supports the previous findings of temporary price pressure generated by abnormal SVIs discussed by Da et. Al (2011). Testing for the gradual diffusion of local information hypothesis shows that there is some evidence of information asymmetries between local and non-local investors, although the results suggest that the information is mostly diffused within a week. As the hypothesis suggests, the asymmetry seems to be stronger for smaller, more remote areas, with analysis in regards to FTSE100 generating more significant results than that of S&P500. Finally, I find evidence that supports positive correlation between search volumes and predictive ability on stock market movements. The industries that are deemed as most familiar in the minds of retail investors do generate more statistically significant results.

I have aimed to add to previous investor attention research by studying investment behavioural differences within regions, as well as, taking advantage of the newest features added to Google Trends throughout the years. I have also analysed whether investment attention within a region has any predictive ability on both investor attention and stock markets within other regions.

This study does come with its limitations. Firstly, the data for S&P500 and FTSE100 isn’t fully comparable since one consists of market-traded industry ETFs and other of self-built industry divisions. A solution to this limitation would be to replicate this study with either ETFs or self-built industries only. Further, I have chosen the search terms used based on the related topics and queries generated by Google Trends, rather than systematically sticking to ticker symbols or company names only. Although I believe taking this route will have reduced the noisiness of the data, this study does lack in robustness checks in this area. Finally, I have chosen to run my regressions using simple OLS models instead of following examples of previous research built on both the Fama-MacBeth and OLS regression models. Although, these models tend to return similar results, there is a possibility of said situation not holding for my data. The test of whether the results presented in this paper would change with the usage of the Fama-MacBeth regression model is thus left for future research. Another possible topic for future research would be to build on findings regarding the information asymmetry between locals and non-locals and study the differences in market capitalisation further, possibly with more frequent data.

This paper asked the question of whether Big Data and the strategies surrounding it are worth the hype and attention generated in the financial world in recent years. The results presented here do indicate Big Data strategies' ability to predict stock market movements, at least in regards to investor attention measured in Google SVIs. However, one must ask the question of whether this benefit offsets the costs involved or whether the costs evaporate the alpha into thin air? For an individual investor, probably. For institutional investors, not necessarily.

8 References

Barber, M. B., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behaviour of Individual and Institutional Investors. *Review of Financial Studies*, *21* (2), pp. 785-818.

Chemmanur, T., & Yan, A. (2009). *Advertising, Attention, and Stock Returns.* Working Paper. Boston College and Fordham University.

Chen, T. (2017). Investor Attention and Global Stock Returns. *Journal of Behavioral Finance, 18 (3),* pp. 358-372

Choi, H., & Varian, H. (2011). *Predicting the Present with Google Trends.* Working Paper. University of California at Berkeley.

CNN Money, (2000). *The $1.7 trillion dot.com lesson*. [viewed 30 April 2018]. Available at: http://money.cnn.com/2000/11/09/technology/overview/

Council on Foreign Relations, (2016). *Understanding the Libor Scandal*. [viewed 30 April 2018]. Available at: https://www.cfr.org/backgrounder/understanding-libor-scandal

Cziraki, P., and Mondria, J., and Wu, T. (2017) *Asymmetric Attention and Stock Returns.* Working paper. University of Toronto and University of California at Santa Cruz.

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

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and Stock Returns: An Alternative Test. *Journal of Financial Markets, 1* (2), pp. 203-219.

Demsetz, H. (1968). The Cost of Transacting. *The Quarterly Journal of Economics, 82* (1), pp. 33-53.

Ding, R., & Hou, W. (2011). *Retail Investor's Active Attention and Stock Liquidity.* Working Paper. Durham University and Middlesex University.

Fama, E. F. (1998). Market Efficiency, Long-term Returns, and Behavioral Finance. *Journal of Financial Economics, 49* (3), pp. 283-306.

Fama, E. F., & MacBeth, J. D. (1973). Risk Return and Equilibrium: Empirical Tests. *Journal of Political Economy, 81* (3), pp. 607-636.

Federal Trade Commission, (2017). *The Equifax Data Breach*. [viewed 30 April 2018]. Available at: https://www.ftc.gov/equifax-data-breach

French K., and Poterba J. (1991) Investor Diversification and International Equity Markets. *American Economics Review, 81*, pp. 222–226.

FTSE Russell, (2018). *Industry Classification Benchmark (ICB).* [viewed 20 March 2018]. Available at: http://www.ftserussell.com/financial-data/industry-classification-benchmark-icb

FTSE Russell, (2018). UKX Quarterly Data Publication, Stock Weights. [viewed 20 March 2018]. Available at: https://www.ftse.com/analytics/factsheets/Home/DownloadConstituentsWeights/?indexdetails=UKX

Gervais, S., Kaniel, R., & Mingelgrin, D. H. (2001). The High-Volume Return Premium. *Journal of Finance*, *56* (3), pp. 877-919.

Global Stats Statcounter, (2018) *Search Engine Market Share Worldwide Apr 2017 – Apr 2018*. [viewed 30 April 2018]. Available at: http://gs.statcounter.com/search-engine-market-share

Google, (2006). *Google Trends*. Available at: https://trends.google.com/trends

Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do Liquidity Measures Measure Liquidity? *Journal of Financial Economics, 92* (2), pp. 153-181.

Hong, H., Lim, T., & Stein, J.C. (1999). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance*, *55* (1), pp. 265-295.

Jobman, D. (2002). *The Handbook of Alternative Investments.* New York, NY: John Wiley & Sons, Inc.

Kahneman, D. (1973). *Attention and Effort.* Englewood Cliffs, NJ: Prentice-Hall.

Li, X., Mahani, R. S., & Sandhya, V. (2011). *Does Investor Attention Affect Stock Prices.* Working Paper. Georgia State University.

Lundström, C. & Nestius, H., (2012). *Can Google Search Volume Data Help Predict Future Stock Measures*?, Stockholm School of Economics

Lou, D. (2008). *Attracting Investor Attention Through Advertising.* Working Paper. London School of Economics and Political Science.

Mishra, Anil V., (2015) Measures of equity home bias puzzle. *Journal of Empirical Finance, (34) pp. 293-312*

Newey, W., & West, K. (1987). A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Matrix. *Econometrica , 55* (3), pp. 703-708.

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business, 53* (1), pp. 61-65.

Peng, L., & Xiong, W. (2006). Investor Attention, Overconfidence and Category Learning. *Journal of Financial Economics, 80* (3), pp. 563-602.

Seasholes, M. S., & Wu, G. (2007). Predictable Behavior, Profits, and Attention. *Journal of Empirical Finance, 14* (5), pp. 590-610.

Skoulakis, G. (2006). *Panel Data Inference in Finance: Least-Squares vs Fama-MacBeth.* Working Paper. University of Maryland.

State Street Global Advisors, (1993). *Select Sector SPDRs*. Available at: http://www.sectorspdr.com/sectorspdr/tools/sector-tracker

Statista, (2017). *Share of desktop search traffic originating from Google in selected countries as of October 2017.* [viewed 30 April 2018]. Available at: https://www.statista.com/statistics/220534/googles-share-of-search-market-in-selected-countries/

Tesar, L., and Werner I. (1995) Home Bias and High Turnover. *Journal of International Money and Finance, 14*, pp. 467–492.

The Guardian, (2008). *Three weeks that changed the world*. [viewed 30 April 2018]. Available at: https://www.theguardian.com/business/2008/dec/28/markets-credit-crunch-banking-2008

Yuan, Y. (2008). *Attention and Trading*. Working Paper. University of Iowa.

Appendix A

Table A1: Results for Differences in Means Tests for US and UK Searches

(a) US and UK Searches within S&P500

Paired t test						
Variable	Obs.	Mean	Std.Err.	Std.Dev.	[95% Conf. Interval]	
S&P500, US Searches	9	27.699	3.200	9.599	20.321	35.077
S&P500, UK Searches	9	20.948	2.451	7.354	15.296	26.601
diff.	9	6.750	3.765	11.294	-1.931	15.432

t = 1.793

mean(diff) = mean(SP_US - SP_UK)
Ho: mean(diff) = 0 degrees of freedom = 8

Ha: mean(diff) < 0	Ha: mean(diff) != 0	Ha: mean(diff) > 0
Pr(T < t) = 0.945	Pr(|T| > |t|) = 0.111	Pr(T > t) = 0.055

(b) US and UK Searches within FTSE100

Paired t test						
Variable	Obs.	Mean	Std.Err.	Std.Dev.	[95% Conf. Interval]	
S&P500, US Searches	9	8.660	1.648	4.944	4.859	12.460
S&P500, UK Searches	9	31.604	2.327	6.982	26.237	36.971
diff.	9	-22.945	2.736	8.209	-29.254	-16.635

t = -8.385

mean(diff) = mean(FTSE_US - FTSE_UK)
Ho: mean(diff) = 0 degrees of freedom = 8

Ha: mean(diff) < 0	Ha: mean(diff) != 0	Ha: mean(diff) > 0
Pr(T < t) = 0.000	Pr(|T| > |t|) = 0.000	Pr(T > t) = 1.000

Table A1 presents the results for differences in means tests. Table *(a)* presents the results for the differences in means test for US and UK searches within S&P500. The results show that the alternative hypothesis of the means not being different gives a p-value of 0.111 and therefore the results are not statistically significant. Thus, I cannot reject the null hypothesis of averages for US and UK searches being similar. Further, the p-value for the alternative hypothesis of US searches being higher than UK searches on average is the lowest generated, which gives an indication of this being mostly the case within S&P500. Table *(b)* presents the results for the differences in means test for US and UK searches within FTSE100. The results show that the alternative hypothesis of the means not being different gives a p-value of 0.000 and therefore the results are statistically significant. Thus, I reject the null hypothesis of averages for US and UK searches being similar. Further, the p-value for the alternative hypothesis of US searches being lower than UK searches is at 0.000 and, thus, also statistically significant. Therefore, I can conclude that US searches on FTSE100 differ from UK searches and are lower than the UK searches. In conclusion, both of the differences in means tests show evidence of US investors being more home biased than UK investors.

Table A2: Effect of ASVIs on FTSE100 Returns

(a) ASVIs defined using the Median Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	-0.003	0.008	-0.013	-0.002	0.009	0.010	-0.006	-0.006	-0.003
	(-0.58)	(1.34)	(-0.81)	(-0.81)	(1.34)	(1.83)	(-0.81)	(-1.57)	(-1.25)
FTSE100 ASVI, one-week lag	-0.005	0.001	-0.010	0.000	-0.006	-0.006	0.031	0.009	-0.002
	(-0.61)	(0.09)	(-0.49)	(0.06)	(-0.32)	(-1.02	(1.24)	(1.84)	(-0.57)
S&P500 ASVI, two-week lag	0.002	-0.000	0.001	-0.003	0.0139	-0.000	-0.010	-0.002	0.004
	(0.55)	(-0.02)	(0.07)	(-0.81)	(1.32)	(-0.02)	(-1.88)	(-0.42)	(1.32)
FTSE100 ASVI, two-week lag	0.012	-0.010	-0.015	-0.005	-0.011	0.009	-0.040	0.006	-0.003
	(1.27)	(-1.38)	(-1.27)	(-0.79)	(-0.52)	(1.63)	(-1.62)	(0.76)	(-0.79)
S&P500 ASVI, three-week lag	-0.002	-0.004	-0.038	-0.004	-0.000	0.002	-0.000	0.002	-0.004
	(-0.24)	(-1.12)	(-1.96)	(-1.24)	(-0.04)	(0.27)	(-0.06)	(0.67)	(-1.68)
FTSE100 ASVI, three-week lag	0.003	-0.005	-0.005	-0.002	0.004	-0.001	0.031	0.000	0.001
	(0.31)	(-0.86)	(-0.63)	(-0.27)	(0.36)	(-0.23)	(1.15)	(0.00)	(0.29)
Constant	0.004^{**}	0.001	-0.000	0.002^{*}	0.000	0.002	0.002	0.001	-0.001
	(2.82)	(0.71)	(-0.03)	(2.30)	(0.03)	(2.14)	(1.01)	(1.06)	(-1.12)
Observations	258	258	258	258	258	258	258	258	258

(b) ASVIs defined using the Mean Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	-0.002	0.008	-0.020	-0.003	0.006	0.009	-0.003	-0.005	-0.003
	(-0.39)	(1.58)	(-1.05)	(-0.83)	(0.81)	(1.75)	(-0.49)	(-1.48)	(-1.32)
FTSE100 ASVI, one-week lag	-0.000	0.000	-0.010	-0.001	-0.006	-0.005	0.036	0.006	-0.002
	(-0.06)	(0.08)	(-0.57)	(-0.16)	(-0.32)	(-0.97)	(1.59)	(1.09)	(-0.49)
S&P500 ASVI, two-week lag	0.003	0.001	0.006	0.004	0.011	-0.002	-0.008	-0.000	0.002
	(0.73)	(0.25)	(0.45)	(0.45)	(1.24)	(-0.30)	(-1.62)	(-0.14)	(0.99)

FTSE100 ASVI, two-week lag	0.012	-0.010	-0.011	-0.003	-0.008	0.011*	-0.039	0.003	-0.001
	(1.45)	(-1.49)	(-1.20)	(-0.55)	(-0.46)	(2.15)	(-1.73)	(0.43)	(-0.40)
S&P500 ASVI, three-week lag	-0.002	-0.004	-0.031	-0.005	-0.003	0.000	0.003	0.002	-0.004
	(-0.24)	(-1.20)	(-1.67)	(-1.72)	(-0.39)	(0.04)	(0.59)	(0.57)	(-1.72)
FTSE100 ASVI, three-week lag	0.003	-0.005	-0.002	-0.000	0.008	0.000	0.028	-0.001	0.000
	(0.44)	(-1.09)	(-0.83)	(-0.004)	(0.76)	(0.18)	(1.13)	(-0.11)	(0.14)
Constant	0.004**	0.001	-0.000	0.001*	0.000	0.002**	0.001	0.001	-0.001
	(2.88)	(0.74)	(-0.01)	(2.16)	(0.24)	(2.60)	(0.87)	(0.85)	(-1.27)
Observations	258	258	258	258	258	258	258	258	258

Table A2 reports the results of OLS regressions on FTSE100 abnormal returns. The sample consists of 90 firms included in the FTSE100 index, divided into nine industries. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal return for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The benchmark used as a market return is the FTSE100 index returns. The independent variables are S&P500 ASVIs with one-, two- and three-week lags as well as FTSE100 ASVIs with one-, two- and three-week lags. Standard errors are obtained by using the Newey-West (1987) formula with eight lags. These are reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A3: Effect of ASVIs on S&P500 Returns

(a) ASVIs defined using the Median Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	0.001	-0.001	-0.018	0.030***	0.000	0.003	-0.004	-0.000	0.001
	(0.49)	(-0.26)	(-1.70)	(1.34)	(0.11)	(0.56)	(-1.01)	(-0.15)	(0.25)
FTSE100 ASVI, one-week lag	-0.001	-0.002	-0.001	-0.011	-0.008	0.008*	-0.001	-0.000	0.000
	(-0.52)	(-0.54)	(-0.15)	(-1.24)	(-1.48)	(2.44)	(-0.12)	(-0.15)	(0.09)
S&P500 ASVI, two-week lag	0.001	0.000	0.007	-0.007	0.004	0.004	-0.004	0.003	0.007*
	(0.55)	(0.10)	(0.61)	(-1.77)	(0.78)	(-1.04)	(-1.29)	(1.96)	(2.08)
FTSE100 ASVI, two-week lag	0.000	-0.002	-0.008	-0.022*	-0.002	-0.002	0.001	-0.001	-0.005
	(0.06)	(-0.33)	(-1.02)	(-1.06)	(-0.24)	(-0.45)	(0.11)	(-0.31)	(-1.59)
S&P500 ASVI, three-week lag	-0.002	-0.001	0.008	0.007	0.002	-0.011*	-0.007**	-0.002	-0.003
	(-1.35)	(-0.20)	(0.88)	(1.93)	(0.84)	(-2.40)	(-2.65)	(-1.05)	(-0.95)
FTSE100 ASVI, three-week lag	0.000	-0.004	-0.006	0.019	-0.001	0.003	0.006	0.003	-0.005
	(0.18)	(-0.74)	(-0.74)	(1.19)	(-0.25)	(0.59)	(0.80)	(0.76)	(-1.43)
Constant	0.001*	-0.001	-0.003**	0.002	0.000	0.001	-0.000	0.002**	-0.002
	(2.20)	(-1.60)	(-2.39)	(1.99)	(0.73)	(1.98)	(-0.07)	(3.79)	(-1.34)
Observations	258	258	258	258	258	258	258	258	258

(b) ASVIs defined using the Mean Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	-0.000	-0.001	-0.018	0.025	-0.000	0.003	-0.004	-0.000	0.000
	(-0.03)	(-0.14)	(-1.81)	(1.28)	(-0.14)	(0.54)	(-1.06)	(-0.15)	(0.16)
FTSE100 ASVI, one-week lag	-0.001	-0.002	-0.002	-0.015	-0.007	0.007*	-0.003	0.000	0.000
	(-0.20)	(-0.58)	(-0.30)	(-1.62)	(-1.52)	(2.34)	(-0.32)	(0.01)	(0.03)
S&P500 ASVI, two-week lag	0.001	0.001	0.004	-0.006	0.004	-0.003	-0.003	0.002	0.006
	(0.34)	(0.37)	(0.42)	(-1.96)	(0.87)	(-0.78)	(-1.21)	(1.68)	(1.90)
FTSE100 ASVI, two-week lag	0.001	-0.001	-0.009	-0.019	-0.002	-0.000	0.003	-0.001	-0.004
	(0.19)	(-0.26)	(-1.23)	(-0.99)	(-0.38)	(-0.16)	(0.33)	(-0.23)	(-1.46)
S&P500 ASVI, three-week lag	-0.002	-0.001	0.009	0.005	0.002	-0.011*	-0.005*	-0.001	-0.002
	(-1.32)	(-0.24)	(1.07)	(1.56)	(0.75)	(-2.49)	(-2.21)	(-0.96)	(-0.85)
FTSE100 ASVI, three-week lag	0.001	-0.005	-0.005	0.0128	-0.001	0.001	0.006	0.003	-0.004
	(0.48)	(-0.97)	(-0.68)	(1.14)	(-0.12)	(0.17)	(0.76)	(0.85)	(-1.46)
Constant	0.001*	-0.001	-0.003*	0.003	0.001	0.001	-0.000	0.002***	-0.002
	(2.10)	(-1.65)	(-2.42)	(1.75)	(0.86)	(1.71)	(-0.57)	(3.81)	(-1.30)
Observations	258	258	258	258	258	258	258	258	258

Table A3 reports the results of OLS regressions on S&P500 abnormal returns. The sample consists of 473 firms included in the S&P500 index, divided into nine industries. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal return for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The benchmark used as a market return is the S&P500 index returns. The independent variables are S&P500 ASVIs with one-, two- and three-week lags as well as FTSE100 ASVIs with one-, two- and three-week lags. Standard errors are obtained by using the Newey-West (1987) formula with eight lags. These are reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A4: Effect of ASVIs on FTSE100 Liquidity

(a) ASVIs defined using the Median Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	-0.124	-0.070	-0.197	-0.005	0.100	0.031	-0.217**	0.131	0.056
	(-1.03)	(-0.72)	(-0.78)	(-0.06)	(1.42)	(0.21)	(-3.10)	(1.47)	(0.99)
FTSE100 ASVI, one-week lag	1.050***	0.915***	0.248*	0.973***	0.378**	0.703***	1.554***	0.646***	0.209***
	(3.95)	(5.52)	(2.37)	(5.09)	(2.79)	(4.50)	(6.19)	(4.97)	(3.41)
S&P500 ASVI, two-week lag	-0.072	-0.044	-0.087	0.059	0.116	0.194	-0.156*	-0.034	-0.056
	(-0.59)	(-0.55)	(-0.41)	(0.81)	(0.81)	(1.10)	(-2.42)	(-0.49)	(-0.84)
FTSE100 ASVI, two-week lag	0.020	0.012	-0.016	-0.382*	-0.188	-0.079	-0.818**	0.178	0.073
	(0.09)	(0.08)	(-0.17)	(-2.36)	(-1.42)	(-0.65)	(-2.98)	(1.48)	(1.07)
S&P500 ASVI, three-week lag	-0.185	0.014	0.292	-0.161	0.036	0.269*	0.027	-0.034	0.062
	(-1.10)	(0.15)	(1.06)	(-1.70)	(0.41)	(2.03)	(0.43)	(-0.49)	(1.26)
S&P500 ASVI, three-week lag	-0.146	-0.262	0.126	0.037	-0.026	-0.095	0.076	-0.196	0.085
	(0.48)	(-1.56)	(0.93)	(0.23)	(-0.21)	(-0.98)	(0.36)	(-1.63)	(0.98)
Constant	-0.117**	-0.022	-0.036	-0.031	-0.025	-0.029	-0.011	-0.041	-0.011
	(-2.82)	(-1.14)	(-1.70)	(-1.36)	(-1.37)	(-1.43)	(-0.48)	(-1.75)	(-0.56)
Observations	258	258	258	258	258	258	258	258	258

(b) ASVIs defined using the Mean Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	-0.103	-0.099	-0.028	0.002	0.076	0.003	-0.184**	0.101	0.026
	(-0.90)	(-1.14)	(-0.12)	(0.02)	(1.20)	(0.03)	(-3.09)	(1.18)	(0.50)
FTSE100 ASVI, one-week lag	0.904***	0.816***	0.249*	0.888***	0.463***	0.669***	1.482***	0.609***	0.206***
	(3.68)	(5.08)	(2.50)	(4.76)	(3.64)	(4.42)	(6.32)	(4.95)	(3.52)
S&P500 ASVI, two-week lag	-0.106	-0.086	0.010	0.035	0.074	0.116	-0.119*	-0.063	-0.064
	(-0.94)	(-1.10)	(0.05)	(0.47)	(0.58)	(0.68)	(-2.15)	(-0.95)	(-1.07)
FTSE100 ASVI, two-week lag	0.050	0.048	-0.025	-0.291*	-0.214	-0.049	-0.751**	0.160	0.072
	(0.24)	(0.34)	(-0.31)	(-2.13)	(-1.72)	(-0.46)	(-3.09)	(1.57)	(1.28)
S&P500 ASVI, three-week lag	-0.178	-0.016	0.252	-0.173	0.003	0.182	0.063	0.077	0.051
	(-1.21)	(-0.17)	(1.09)	(-1.86)	(0.04)	(1.28)	(1.21)	(1.34)	(1.05)
S&P500 ASVI, three-week lag	-0.148	-0.270	0.098	0.073	-0.069	-0.044	0.101	-0.157	0.066
	(-0.83)	(-1.72)	(0.84)	(0.50)	(-0.55)	(-0.51)	(0.57)	(-1.43)	(0.83)
Constant	-0.120**	-0.023	-0.037	-0.032	-0.023	-0.020	-0.018	-0.038	-0.007
	(-2.91)	(-1.19)	(-1.77)	(-1.39)	(-1.22)	(-1.10)	(-0.84)	(-1.58)	(-0.35)
Observations	258	258	258	258	258	258	258	258	258

Table A4 reports the results of OLS regressions on FTSE100 abnormal liquidity. The sample consists of 90 firms included in the FTSE100 index, divided into nine industries. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal return for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The benchmark used as a market return is the FTSE100 index returns. The independent variables are S&P500 ASVIs with one-, two- and three-week lags as well as FTSE100 ASVIs with one-, two- and three-week lags. Standard errors are obtained by using the Newey-West (1987) formula with eight lags. These are reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A5: Effect of ASVIs on S&P500 Liquidity

(a) ASVIs defined using the Median Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	0.001	-0.176	0.171	0.191	0.064	0.020	-0.222*	0.110	0.059
	(0.01)	(-1.27)	(0.86)	(2.29)	(0.54)	(0.09)	(-2.45)	(1.61)	(1.40)
FTSE100 ASVI, one-week lag	0.141	0.209	0.290*	0.633**	0.545*	0.353	1.060***	0.180	0.169*
	(0.89)	(1.15)	(2.56)	(3.29)	(2.52)	(1.91)	(4.83)	(1.68)	(2.86)
S&P500 ASVI, two-week lag	-0.054	-0.138	0.076	0.078	0.304*	0.097	-0.126	-0.033	-0.058
	(-0.63)	(-1.11)	(0.40)	(0.96)	(1.77)	(0.41)	(-1.51)	(-0.53)	(-1.06)
FTSE100 ASVI, two-week lag	0.002	0.012	-0.114	-0.418*	-0.718**	-0.233	-0.678*	-0.061	0.111
	(0.01)	(0.07)	(-1.02)	(-2.72)	(2.52)	(-1.25)	(-2.63)	(-0.51)	(1.59)
S&P500 ASVI, three-week lag	-0.086	-0.111	0.034	-0.067	0.030	0.037	-0.019	0.028	0.017
	(-0.73)	(-0.79)	(0.13)	(-0.56)	(0.32)	(0.18)	(-0.22)	(0.42)	(0.42)

FTSE100 ASVI, three-week lag	-0.013	-0.306	0.043	-0.199	-0.075	0.123	0.244	-0.283*	0.104
	(-0.08)	(-1.33)	(0.39)	(-1.15)	(-0.28)	(0.80)	(1.04)	(-2.30)	(1.72)
Constant	-0.041	-0.043	-0.030	-0.042	-0.043	-0.041	-0.040	-0.030	-0.011
	(-1.72)	(-1.52)	(-1.06)	(-1.54)	(-1.28)	(-1.75)	(-1.38)	(-1.21)	(-0.49)
Observations	258	258	258	258	258	258	258	258	258

(b) ASVIs defined using the Mean Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	-0.008	-0.211	0.314	0.167	0.052	0.057	-0.193*	0.131*	0.032
	(-0.09)	(-1.65)	(1.69)	(1.88)	(0.49)	(0.32)	(-2.26)	(2.13)	(0.73)
FTSE100 ASVI, one-week lag	0.093	0.218	0.239*	0.531**	0.527*	0.294	0.986***	0.114	0.182**
	(0.73)	(1.32)	(2.42)	(2.87)	(2.49)	(1.74)	(4.62)	(1.11)	(3.25)
S&P500 ASVI, two-week lag	-0.054	-0.199	0.107	0.048	0.253	0.075	-0.089	-0.021	-0.050
	(-0.74)	(-1.61)	(0.65)	(0.55)	(1.72)	(0.36)	(-1.10)	(-0.38)	(-1.02)
FTSE100 ASVI, two-week lag	-0.017	0.013	-0.139	-0.383**	-0.671***	-0.231	-0.712**	-0.069	0.109
	(-0.13)	(0.07)	(-1.39)	(-2.96)	(-3.57)	(-1.47)	(-3.12)	(-0.65)	(1.74)
S&P500 ASVI, three-week lag	-0.077	-0.125	0.034	-0.095	0.053	0.047	0.003	0.013	0.005
	(-0.77)	(-0.97)	(0.15)	(-0.88)	(0.61)	(0.26)	(0.04)	(0.25)	(0.15)
FTSE100 ASVI, three-week lag	-0.013	-0.279	-0.026	-0.196	-0.106	0.077	0.162	-0.257*	0.105*
	(-0.09)	(-1.32)	(-0.27)	(-1.21)	(-0.43)	(0.57)	(0.80)	(-2.32)	(2.04)
Constant	-0.043	-0.053	-0.032	-0.040	-0.039	-0.038	-0.049	-0.029	-0.006
	(-1.79)	(-1.79)	(-1.12)	(-1.48)	(-1.22)	(-1.66)	(-1.72)	(-1.16)	(-0.27)
Observations	258	258	258	258	258	258	258	258	258

Table A5 reports the results of OLS regressions on S&P500 abnormal liquidity. The sample consists of 473 firms included in the S&P500 index, divided into nine industries. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal return for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The benchmark used as a market return is the S&P500 index returns. The independent variables are S&P500 ASVIs with one-, two- and three-week lags as well as FTSE100 ASVIs with one-, two- and three-week lags. Standard errors are obtained by using the Newey-West (1987) formula with eight lags. These are reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A6: Effect of ASVIs on FTSE100 Volatility

(a) ASVIs defined using the Median Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	-0.008	-0.064	-0.079	-0.088	0.145	-0.133	-0.053	-0.028	0.045
	(-0.10)	(-0.50)	(-0.29)	(-0.80)	(1.09)	(-0.92)	(-0.88)	(-0.40)	(1.07)
FTSE100 ASVI, one-week lag	0.501***	0.391	0.184	0.497*	0.997***	0.206	0.925***	0.244*	0.058
	(3.62)	(1.71)	(1.46)	(2.56)	(3.64)	(1.49)	(6.08)	(2.02)	(0.84)
S&P500 ASVI, two-week lag	-0.109	0.063	0.088	0.200*	-0.001	0.034	-0.010	-0.084	-0.003
	(-1.79)	(0.53)	(0.32)	(2.23)	(-0.01)	(0.25)	(-1.81)	(-1.16)	(-0.05)
FTSE100 ASVI, two-week lag	0.039	0.223	-0.153	-0.212	-0.564*	0.078	-0.480*	0.106	0.054
	(0.40)	(1.17)	(-1.03)	(-1.24)	(-2.12)	(0.57)	(-1.92)	(0.78)	(1.05)
S&P500 ASVI, three-week lag	-0.112	0.250	-0.247	-0.089	-0.082	0.091	0.086	0.122	-0.073
	(-1.37)	(1.44)	(-0.82)	(-1.09)	(-0.68)	(0.58)	(1.80)	(2.01)	(-1.64)
FTSE100 ASVI, three-week lag	-0.216	-0.626**	-0.122	-0.363	0.065	-0.117	0.091	-0.283*	-0.040
	(-1.72)	(-3.17)	(-0.71)	(-1.87)	(0.33)	(-0.88)	(0.42)	(-2.49)	(-0.49)
Constant	-0.030	-0.089*	-0.063*	-0.043*	-0.033	-0.020	-0.016	-0.033	-0.035
	(-1.51)	(-1.86)	(-2.56)	(-1.85)	(-1.43)	(-1.01)	(-0.92)	(-1.54)	(-1.57)
Observations	258	258	258	258	258	258	258	258	258

(b) ASVIs defined using the Mean Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	0.001	-0.017	0.002	-0.042	0.091	-0.145	-0.030	-0.005	0.012
	(0.01)	(-0.14)	(0.01)	(-0.38)	(0.78)	(-1.05)	(-0.56)	(-0.08)	(0.48)
FTSE100 ASVI, one-week lag	0.458***	0.387	0.168	0.429*	0.949***	0.199	0.915***	0.233*	0.063
	(3.81)	(1.85)	(1.48)	(2.41)	(3.86)	(1.59)	(6.52)	(2.04)	(1.00)
S&P500 ASVI, two-week lag	-0.126*	0.102	0.116	0.194*	-0.005	-0.044	-0.068	-0.077	-0.007
	(-2.11)	(0.77)	(0.47)	(2.26)	(-0.04)	(-0.29)	(-1.35)	(-1.13)	(-0.15)
FTSE100 ASVI, two-week lag	0.089	0.179	-0.145	-0.166	-0.535*	0.102	-0.419	0.081	0.059

	(1.10)	(1.00)	(-1.08)	(-1.14)	(-2.20)	(0.79)	(-1.86)	(0.64)	(1.31)
S&P500 ASVI, three-week lag	-0.101	0.315	-0.240	-0.050	-0.068	0.010	0.109**	0.113*	-0.083
	(-1.45)	(1.55)	(-0.87)	(-0.64)	(-0.57)	(0.06)	(2.61)	(2.02)	(-1.88)
FTSE100 ASVI, three-week lag	-0.208	-0.580**	-0.154	-0.318	-0.001	-0.106	0.080	-0.270*	-0.001
	(-1.78)	(-3.30)	(-0.95)	(-1.81)	(-0.01)	(-0.89)	(0.41)	(-2.56)	(-0.01)
Constant	-0.033	-0.082	-0.063*	-0.042	-0.033	-0.019	-0.016	-0.033	-0.035
	(-1.66)	(-1.89)	(-2.56)	(-1.84)	(-1.46)	(-1.08)	(-1.01)	(-1.52)	(-1.59)
Observations	258	258	258	258	258	258	258	258	258

Table A6 reports the results of OLS regressions on FTSE100 abnormal volatility. The sample consists of 90 firms included in the FTSE100 index, divided into nine industries. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal return for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The benchmark used as a market return is the FTSE100 index returns. The independent variables are S&P500 ASVIs with one-, two- and three-week lags as well as FTSE100 ASVIs with one-, two- and three-week lags. Standard errors are obtained by using the Newey-West (1987) formula with eight lags. These are reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A7: Effect of ASVIs on S&P500 Volatility

(a) ASVIs defined using the Median Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	0.082	-0.024	-0.336	-0.057	0.005	-0.033	-0.127	0.127	0.077
	(0.66)	(-0.15)	(-1.39)	(-0.40)	(0.04)	(-0.13)	(-1.36)	(1.69)	(1.49)
FTSE100 ASVI, one-week lag	0.127	0.068	0.102	0.446*	0.443	0.279	0.599*	0.121	0.028
	(0.79)	(0.23)	(0.80)	(2.26)	(1.68)	(1.58)	(2.38)	(0.75)	(0.37)
S&P500 ASVI, two-week lag	-0.040	-0.190	-0.041	0.140	0.101	-0.115	-0.156*	-0.049	-0.098
	(-0.37)	(-1.78)	(-0.16)	(1.28)	(0.71)	(-0.47)	(-2.03)	(-0.56)	(-1.76)
FTSE100 ASVI, two-week lag	0.170	0.267	-0.166	-0.189	-0.665**	-0.056	-0.499	-0.073	0.271**
	(0.99)	(1.26)	(-1.34)	(-0.78)	(-2.89)	(-0.25)	(-1.35)	(-0.58)	(3.12)
S&P500 ASVI, three-week lag	0.029	-0.065	0.371	0.094	-0.255	-0.174	0.012	0.046	-0.002
	(0.29)	(-0.47)	(1.20)	(0.70)	(-1.65)	(-0.61)	(0.15)	(0.43)	(-0.04)
FTSE100 ASVI, three-week lag	-0.191	-0.267	-0.040	-0.461	0.158	-0.286	0.057	-0.532**	0.021
	(-0.95)	(-1.47)	(-0.31)	(-2.06)	(0.55)	(-1.63)	(0.17)	(-3.81)	(0.24)
Constant	-0.040	-0.051	-0.047	-0.052	-0.041	-0.034	-0.034	-0.050	-0.034
	(-1.27)	(-1.83)	(-1.78)	(-1.62)	(-1.25)	(-1.16)	(-1.21)	(-1.69)	(-1.61)
Observations	258	258	258	258	258	258	258	258	258

(b) ASVIs defined using the Mean Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
S&P500 ASVI, one-week lag	0.0721	-0.061	-0.107	0.009	-0.007	-0.014	-0.102	0.158*	0.062
	(0.67)	(-0.42)	(-0.47)	(0.06)	(-0.06)	(-0.07)	(-1.19)	(2.24)	(1.27)
FTSE100 ASVI, one-week lag	0.010	0.064	0.094	0.330	0.415	0.226	0.558*	0.054	0.059
	(0.75)	(0.25)	(0.88)	(1.84)	(1.68)	(1.53)	(2.19)	(0.35)	(0.78)
S&P500 ASVI, two-week lag	-0.059	-0.235*	-0.007	0.099	0.114	-0.093	-0.111	-0.028	-0.088
	(-0.64)	(-2.33)	(-0.03)	(0.92)	(0.95)	(-0.43)	(-1.58)	(-0.34)	(-1.69)
FTSE100 ASVI, two-week lag	0.173	0.233	-0.143	-0.160	-0.697***	-0.064	-0.528	-0.091	0.248**
	(1.24)	(1.25)	(-1.21)	(-0.74)	(-3.37)	(-0.34)	(-1.55)	(-0.78)	(3.30)
S&P500 ASVI, three-week lag	0.023	-0.101	0.360	0.106	-0.205	-0.108	0.046	0.034	-0.001
	(0.26)	(-0.85)	(1.34)	(0.90)	(-1.56)	(-0.40)	(0.62)	(0.37)	(-0.03)
FTSE100 ASVI, three-week lag	-0.182	-0.240	-0.080	-0.400	0.117	-0.286	0.007	-0.473***	0.038
	(-1.07)	(-1.46)	(-0.73)	(-1.91)	(0.42)	(-1.83)	(0.02)	(-3.76)	(0.50)
Constant	-0.040	-0.058*	-0.050	-0.049	-0.043	-0.038	-0.040	-0.049	-0.030
	(-1.22)	(-2.05)	(-1.90)	(-1.59)	(-1.35)	(-1.31)	(-1.43)	(-1.63)	(-1.44)
Observations	258	258	258	258	258	258	258	258	258

Table A7 reports the results of OLS regressions on S&P500 abnormal volatility. The sample consists of 473 firms included in the S&P500 index, divided into nine industries. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal return for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The benchmark used as a market return is the S&P500 index returns. The independent variables are S&P500 ASVIs with one-, two- and three-week lags as well as FTSE100 ASVIs with one-, two- and three-week lags. Standard errors are obtained by using the Newey-West (1987) formula with eight lags. These are reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A8: S&P500 SVI, UK Non-Local

(a) ASVIs defined using the Median Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
Non-Local ASVI, last week	-0.075	0.118	0.061	0.234***	0.057	-0.056	0.213***	0.100	0.059
	(-1.20)	(1.91)	(0.99)	(3.85)	(0.91)	(-0.86)	(3.45)	(1.61)	(0.96)
Local ASVI, same week	0.074	0.255*	0.102	-0.367	0.330	0.173	0.016	-0.131	-0.140
	(1.05)	(2.38)	(1.87)	(-1.17)	(1.31)	(1.26)	(0.12)	(-0.55)	(-0.80)
Local ASVI, last week	0.016	0.060	0.106*	-0.082	-0.145	-0.023	-0.075	0.338	0.266
	(0.22)	(0.55)	(1.98)	(-0.26)	(-0.57)	(-0.17)	(-0.55)	(1.42)	(1.52)
Constant	-0.161**	-0.088*	0.030	-0.184	-0.310***	-0.211***	-0.246**	0.430***	0.526***

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	(-3.33)	(-1.98)	(0.75)	(-1.63)	(-3.61)	(-3.66)	(-2.90)	(3.92)	(5.62)
Observations	260	260	260	260	260	260	260	260	260

(b) ASVIs defined using the Mean Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
Non-Local ASVI, last week	-0.077	0.097	-0.025	0.023	-0.004	-0.010	0.157*	0.157*	0.049
	(-1.24)	(1.56)	(-0.40)	(0.36)	(-0.06)	(-0.15)	(2.52)	(2.52)	(0.79)
Local ASVI, same week	0.020	0.228*	0.104	-0.120	0.352	0.138	0.028	0.028	-0.188
	(0.28)	(2.12)	(1.92)	(-0.41)	(1.47)	(0.99)	(0.21)	(0.21)	(-1.09)
Local ASVI, last week	-0.031	0.081	0.098	-0.078	-0.040	0.068	-0.050	-0.050	0.167
	(-0.45)	(0.74)	(1.84)	(-0.27)	(-0.17)	(0.49)	(-0.38)	(-0.38)	(0.96)
Constant	0.017	-0.006	0.041	0.014	0.016	-0.003	-0.003	0.021	0.007
	(0.34)	(-0.13)	(0.95)	(0.12)	(0.17)	(-0.05)	(-0.03)	(0.20)	(0.07)
Observations	260	260	260	260	260	260	260	260	260

Table A8 reports the results of OLS regressions on the effects of past and present local and past non-local abnormal SVIs on present non-local abnormal SVIs within S&P500, divided into nine industries. On the S&P500 index US searches are denoted as local and UK searches as non-local. The sample consists of 473 firms included in the S&P500 index. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal SVI for non-local searches for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The independent variables are the local abnormal SVIs for week t and t-1 as well as non-local abnormal SVIs for week t-1. Standard errors are also obtained and reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A9: FTSE100 SVI, US Non-Local

(a) ASVIs defined using the Median Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
Non- Local ASVI, last week	0.0619	0.115	0.037	0.220***	0.057	0.141*	0.125*	0.093	0.026
	(0.99)	(1.88)	(0.59)	(3.58)	(0.89)	(2.24)	(2.02)	(1.49)	(0.41)
Local ASVI, same week	0.104	0.174*	0.120	0.149**	0.524***	0.138**	0.322***	0.200	0.027
	(1.59)	(2.53)	(1.76)	(3.23)	(9.43)	(2.61)	(6.29)	(1.29)	(0.43)
Local ASVI, last week	-0.024	0.154*	0.105	-0.021	-0.079	-0.013	-0.035	-0.286	-0.069
	(-0.36)	(2.21)	(1.53)	(-0.45)	(-1.28)	(-0.25)	(-0.63)	(-1.84)	(-1.10)
Constant	0.005	0.003	-0.006	-0.002	-0.017	-0.006	-0.006	-0.067	-0.027
	(0.35)	(0.19)	(-0.30)	(-0.23)	(-1.81)	(-0.66)	(-0.70)	(-1.42)	(-1.31)
Observations	260	260	260	260	260	260	260	260	260

(b) ASVIs defined using the Mean Approach

	(1) Consumer Discretionary	(2) Consumer Staples	(3) Energy	(4) Financials	(5) Health Care	(6) Industrials	(7) Materials	(8) Technology	(9) Utilities
Non- Local ASVI, last week	0.004	0.062	0.035	0.216***	0.014	0.133*	0.144*	0.025	0.041
	(0.07)	(1.01)	(0.57)	(3.49)	(0.22)	(2.11)	(2.32)	(0.40)	(0.66)
Local ASVI, same week	0.114	0.194**	0.154*	0.148**	0.513***	0.139*	0.331***	0.160	0.007
	(1.79)	(2.81)	(2.22)	(3.26)	(9.42)	(2.58)	(6.45)	(1.05)	(0.11)
Local ASVI, last week	0.045	0.171*	0.135	-0.028	0.011	-0.005	-0.013	-0.171	-0.071
	(0.69)	(2.44)	(1.92)	(-0.59)	(0.18)	(-0.10)	(-0.23)	(-1.12)	(-1.15)
Constant	-0.016	-0.024	-0.025	-0.008	-0.009	-0.004	-0.011	-0.191***	-0.058*
	(-0.93)	(-1.22)	(-1.05)	(-1.13)	(-0.87)	(-0.38)	(-1.08)	(-3.68)	(-2.53)
Observations	260	260	260	260	260	260	260	260	260

Table A9 reports the results of OLS regressions on the effects of past and present local and past non-local abnormal SVIs on present non-local abnormal SVIs within FTSE100, divided into nine industries. On the FTSE100 index UK searches are denoted as local and US searches as non-local. The sample consists of 90 firms included in the FTSE100 index. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal SVI for non-local searches for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The independent variables are the local abnormal SVIs for week t and t-1 as well as non-local abnormal SVIs for week t-1. Standard errors are also obtained and reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A10: S&P500 SVI Using Market Capitalisation, UK Non-Local

(a) ASVIs defined using the Median Approach

	(1) Small Cap	(2) Mid Cap	(3) Large Cap
Non-Local ASVI, last week	0.081	0.191**	0.254***
	(1.28)	(3.16)	(4.21)
Local ASVI, same week	0.429**	0.409***	0.261
	(2.90)	(3.68)	(1.76)
Local ASVI, last week	0.222	0.197	0.065
	(1.51)	(1.72)	(0.44)
Constant	0.003	-0.002	-0.010
	(0.28)	(-0.23)	(-0.82)
Observations	260	260	260

(b) ASVIs defined using the Mean Approach

	(1) Small Cap	(2) Mid Cap	(3) Large Cap
Non-Local ASVI, last week	0.018	0.173**	0.226***
	(0.29)	(2.86)	(3.72)

Local ASVI, same week	0.561***	0.447***	0.303*
	(3.89)	(4.10)	(2.05)
Local ASVI, last week	0.266	0.181	0.010
	(1.83)	(1.58)	(0.67)
Constant	-0.003	-0.008	-0.013
	(-0.29)	(-0.82)	(-0.93)
Observations	260	260	260

Table A10 reports the results of OLS regressions on the effects of past and present local and past non-local abnormal SVIs on present non-local abnormal SVIs within FTSE100 when divided according to market capitalisation. On the S&P500 index US searches are denoted as local and UK searches as non-local. The sample consists of 64 firms included in the S&P500 index. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal SVI for non-local searches for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The independent variables are the local abnormal SVIs for week t and t-1 as well as non-local abnormal SVIs for week t-1. Standard errors are also obtained and reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Table A11: FTSE100 SVI Using Market Capitalisation, US Non-Local

(a) ASVIs defined using the Median Approach

	(1) Small Cap	(2) Mid Cap	(3) Large Cap
Non-Local ASVI, last week	0.038	-0.036	0.235***
	(0.62)	(-0.58)	(3.84)
Local ASVI, same week	0.254**	0.084	0.237***
	(2.65)	(1.75)	(5.71)
Local ASVI, last week	0.127	0.015	-0.048
	(1.31)	(0.30)	(-1.11)
Constant	-0.009	-0.049*	-0.002
	(-0.98)	(-2.10)	(-0.43)
Observations	260	260	260

(b) ASVIs defined using the Mean Approach

	(1) Small Cap	(2) Mid Cap	(3) Large Cap
Non-Local ASVI, last week	0.036	-0.042	0.279***
	(0.58)	(-0.67)	(4.60)
Local ASVI, same week	0.261**	0.106*	0.230***
	(2.75)	(2.24)	(5.42)
Local ASVI, last week	0.175	0.031	-0.068
	(1.81)	(0.64)	(-1.54)
Constant	-0.016	-0.044	-0.007
	(-1.48)	(-1.73)	(-1.09)
Observations	260	260	260

Table A11 reports the results of OLS regressions on the effects of past and present local and past non-local abnormal SVIs on present non-local abnormal SVIs within FTSE100 when divided according to market capitalisation. On the FTSE100 index UK searches are denoted as local and US searches as non-local. The sample consists of 57 firms included in the FTSE100 index. The sample period is from 10th of March 2013 to 10th of March 2018. The dependent variable is the abnormal SVI for non-local searches for week t, denoted in percentages. For regressions in Table *(a)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the median of prior eight weeks. For regressions in Table *(b)*, the abnormal SVI (ASVI) is determined using expected SVI value defined as the mean of prior eight weeks. The independent variables are the local abnormal SVIs for week t and t-1 as well as non-local abnormal SVIs for week t-1. Standard errors are also obtained and reported within parentheses below the regression coefficients. * represents significance at the 10 per cent level. ** represents significance at the 5 per cent level. *** represents significance at the 1 per cent level.

Appendix B

Table B1: List of Constituents and Their Weights within Industries

(a) FTSE100

Industry	Company Name	Ticker	Weight, % (index)	Weight, decimals (SVI)	Market Capitalisation
Consumer Discretionary	GKN plc	GKN	0,28	0,02	
	Barratt Developments plc	BDEV	0,33	0,02	
	The Berkeley Group Holdings plc	BKG	0,27	0,02	
	Persimmon plc	PSN	0,42	0,03	
	Reckitt Benckiser Group plc	RB.	2,17	0,16	
	Informa plc	INF	0,30	0,02	
	Itv plc	ITV	0,31	0,02	
	Pearson plc	PSON	0,30	0,02	
	Relx plc	REL	0,93	0,07	
	Sky plc	SKY	0,53	0,04	
	Wpp plc	WPP	0,83	0,06	
	Burberry Group plc	BRBY	0,38	0,03	
	Unilever plc	ULVR	2,40	0,17	
	Carnival plc	CCL	0,44	0,03	
	Compass Group PLC	CPG	1,27	0,09	
	easyJet plc	EZJ	0,18	0,01	
	Intercontinental Hotels Group PLC	IHG	0,45	0,03	
	International Consolidated Airlines Group, S.A.	IAG	0,54	0,04	
	Paddy Power Betfair plc	PPB	0,37	0,03	
	Whitbread PLC	WTB	0.34	0,02	
Consumer Staples	Diageo plc	DGE	3,39	0,24	large
	Wm Morrison Supermarkets PLC	MRW	0,24	0,02	small
	J Sainsbury plc	SBRY	0,20	0,01	small
	Tesco PLC	TSCO	0,86	0,06	mid
	Associated British Foods plc	ABF	0,49	0,04	mid

	Kingfisher plc	KGF	0,37	0,03	mid
	Marks and Spencer Group plc	MKS	0,26	0,02	small
	Next plc	NXT	0,32	0,02	mid
	British American Tobacco p.l.c.	BATS	5,78	0,42	large
	Imperial Brands PLC	IMB	1,52	0,11	large
Energy	BP p.l.c.	BP.	5,07	0,32	
	Royal Dutch Shell plc A	RDSA	5,68	0,36	
	Royal Dutch Shell plc B	RDSB	4,74	0,30	
	DCC plc	DCC	0,33	0,02	
Financials	Barclays plc	BARC	1,74	0,08	large
	3i Ord	III	0,44	0,02	mid
	Barclays plc	BARC	1,74	0,08	large
	London Stock Exchange Group plc	LSE	0,60	0,03	mid
	Schroders plc	SDR	0,20	0,01	small
	Standard Life Aberdeen plc	SLA	0,60	0,03	mid
	Aviva plc	AV.	1,02	0,05	large
	Legal & General Group Plc	LGEN	0,81	0,04	mid
	Old Mutual plc	OML	0,55	0,03	mid
	Prudential plc	PRU	2,48	0,11	large
	St James's Place plc	STJ	0,32	0,01	mid
	Admiral Group plc	ADM	0,20	0,01	small
	Direct Line Insurance Group PLC	DLG	0,26	0,01	small
	RSA Insurance Group plc	RSA	0,32	0,01	mid
Health Care	NMC Health plc	NMC	0,12	0,01	small
	Smith & Nephew plc	SN.	0,57	0,06	mid
	AstraZeneca PLC	AZN	3,27	0,36	large
	GlaxoSmithKline plc	GSK	3,22	0,36	large
	Shire plc	SHP	1,75	0,19	large
Industrials	BAE Systems plc	BA	0,92	0,13	mid
	Rolls-Royce Holdings plc	RR.	0,70	0,10	mid
	CRH plc	CRH	1,12	0,16	large
	Halma plc	HLMA	0,24	0,03	small
	DS Smith Plc	SMDS	0,26	0,04	small
	Smiths Group plc	SMIN	0,30	0,04	mid
	Smurfit Kappa Group plc	SKG	0,30	0,04	mid
	Experian plc	EXPN	0,77	0,11	mid
	G4S plc	GFS	0,21	0,03	small
	Ferguson plc	FERG	0,68	0,10	mid
	Bunzl plc	BNZL	0,35	0,05	mid
	Ashtead Group plc	AHT	0,50	0,07	mid
	Intertek Group plc	ITRK	0,42	0,06	mid
	Rentokil Initial plc	RTO	0,29	0,04	mid
Materials	Croda International plc	CRDA	0,28	0,03	mid
	Johnson Matthey plc	JMAT	0,29	0,03	mid
	Mondi plc	MNDI	0,36	0,04	mid
	EVRAZ plc	EVR	0,40	0,04	mid
	Anglo American plc	AAL	0,77	0,09	mid
	Antofagasta plc	ANTO	0,18	0,02	small
	BHP Billiton plc	BLT	1,59	0,18	large
	Fresnillo PLC	FRES	0,13	0,01	small
	Glencore Plc	GLEN	2,29	0,25	large
	Randgold Resources Limited	RRS	0,35	0,04	mid
	Rio Tinto plc	RIO	2,36	0,26	large
Technology	BT Group plc	BT.A	1,15	0,22	
	Vodafone Group plc	VOD	3,15	0,60	
	Micro Focus International plc	MCRO	0,54	0,10	
	The Sage Group plc	SGE	0,43	0,08	
Utilities	SSE plc	SSE	0,68	0,22	
	Centrica plc	CNA	0,39	0,13	
	National Grid plc	NG.	1,49	0,48	
	Severn Trent Plc	SVT	0,26	0,08	
	United Utilities Group PLC	UU	0,29	0,09	

(a) S&P500

Industry	Company Name	Ticker	Weight, % (index)	Weight, % (index, adj.)	Market Capitalisation
Consumer Staples	Procter & Gamble	PG	0,113	0,82	large
	Coca-Cola Co	KO	0,102	0,74	large
	PepsiCo Inc	PEP	0,088	0,64	large
	Philip Morris International	PM	0,079	0,57	large
	Walmart Inc.	WMT	0,077	0,55	large
	Costco Wholesale Corp	COST	0,050	0,36	large
	CVS Health Corporation	CVS	0,043	0,31	large
	Altria Group Inc	MO	0,042	0,31	large
	Colgate-Palmolive Co	CL	0,039	0,28	large
	Mondelez International Inc	MDLZ	0,038	0,28	large
	Walgreens Boots Alliance	WBA	0,035	0,25	large
	Constellation Brands Inc	STZ	0,025	0,18	mid
	Kimberly-Clark	KMB	0,023	0,16	mid
	The Kraft Heinz Company	KHC	0,022	0,16	mid

	Estee Lauder Cos. A	EL	0,022	0,16	mid
	Sysco Corp	SYY	0,019	0,14	mid
	Archer-Daniels-Midland Co	ADM	0,016	0,12	mid
	General Mills	GIS	0,016	0,11	mid
	Monster Beverage Corp	MNST	0,015	0,11	mid
	Dr Pepper Snapple Group	DPS	0,014	0,10	mid
	Kroger Co	KR	0,014	0,10	mid
	Tyson Foods Inc A	TSN	0,013	0,10	mid
	Clorox Co	CLX	0,010	0,07	mid
	Kellogg Co	K	0,010	0,07	mid
	Conagra Brands, Inc	CAG	0,009	0,07	mid
	Brown-Forman Corp B	BF.b	0,009	0,07	mid
	Molson Coors Brewing Co B	TAP	0,008	0,06	small
	Hershey Foods Corp	HSY	0,008	0,06	small
	Smucker J.M. Co	SJM	0,008	0,06	small
	McCormick & Co	MKC	0,008	0,06	small
	Church & Dwight Co	CHD	0,007	0,05	small
	Hormel Foods Corp	HRL	0,006	0,04	small
	Coty Inc	COTY	0,005	0,04	small
	Campbell Soup Co	CPB	0,005	0,04	small
Energy	Exxon Mobil Corp	XOM	0,226	1,39	large
	Chevron Corp	CVX	0,170	1,04	large
	Schlumberger Ltd	SLB	0,070	0,43	large
	ConocoPhillips	COP	0,050	0,31	large
	EOG Resources	EOG	0,046	0,28	large
	Occidental Petroleum	OXY	0,043	0,27	large
	Valero Energy Group	VLO	0,035	0,21	mid
	Phillips 66	PSX	0,034	0,21	mid
	Halliburton Co	HAL	0,034	0,21	mid
	Marathon Petroleum Corp.	MPC	0,028	0,17	mid
	Anadarko Petroleum Corp	APC	0,025	0,16	mid
	Pioneer Natural Resources	PXD	0,025	0,15	mid
	Kinder Morgan Inc	KMI	0,023	0,14	mid
	ONEOK Inc	OKE	0,018	0,11	mid
	Concho Resources Inc	CXO	0,017	0,11	mid
	The Williams Companies Inc	WMB	0,016	0,10	mid
	Devon Energy Corp	DVN	0,014	0,08	mid
	Noble Energy Inc	NBL	0,012	0,08	mid
	Andeavor	ANDV	0,012	0,08	mid
	Apache Corp	APA	0,012	0,07	mid
	Hess Corp	HES	0,011	0,07	mid
	Marathon Oil Corp	MRO	0,011	0,07	mid
	National Oilwell Varco Inc	NOV	0,011	0,07	mid
	Baker Hughes, a GE company	BHGE	0,011	0,07	mid
	EQT Corporation	EQT	0,009	0,05	small
	Cabot Oil & Gas A	COG	0,008	0,05	small
	Cimarex Energy Co	XEC	0,007	0,04	small
	Helmerich & Payne Inc	HP	0,006	0,04	small
	Newfield Exploration Co	NFX	0,004	0,03	small
	Range Resources Corp	RRC	0,002	0,01	small

Table B1 is a list of all the FTSE100 constituents used in this study divided into nine industries: Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Materials, Technology and Utilities. The ticker column presents the tickers used for each of the constituents on FTSE100 index. The weight, % (index) represents the weight of the company within FTSE100 (gotten from UKX Quarterly Data Report) or S&P500 (gotten the SPDR Sector EFTs website), denoted in percentages. The weight, decimals (SVI) in Table (a) represents the weight of the SVI value of the company within its industry and is also used to calculate the market capitalisation of the company. The weight, % (index adj.) in Table (b) is the weight, % (index) multiplied with the weight of the industry within S&P500 and is used to calculate SVI values and market capitalisation. The largest 25% of constituents within an index are denoted as Large Cap, 50-75% as Mid Cap and bottom 25% as Small Cap.

Table B2: Search Terms Used to Obtain Google Trends Data

(a) FTSE100

GKN	Tesco plc	Direct Line Group	Rentokil initial
BDEV	Associated British Foods	RSA insurance	JMAT
Persimmon plc	Kingfisher plc	NMC health	Evraz
Reckitt Benckiser Group	Marks & Spencer	Smith & Nephew	AAL
Informa plc	Next plc	Astrazeneca	Antofagasta
Pearson plc	British American Tobacco	GSK	BHP billiton
Relx	Imperial Brands	Shire plc	Glencore
Sky plc	BP plc	BAE Systems	Randgold
Wpp plc	RDSA	Rolls-Royce plc	BT group
BRBY	RDSB	CRH plc	Vodafone group
ULVR	Barclays plc	HLMA	Micro Focus International
Carnival plc	HSBA	DS Smith	Sage Group plc
Compass Group	LLOY	SMIN	SSE plc
EZJ	Standard Chartered	Smurfit Kappa Group	Centrica
Intercontinental Hotels Group	3i Group	Experian	National grid plc
IAG	Schroders	Ferguson plc	
Paddy Power Betfair	Aviva plc	AHT	
Diageo plc	LGEN	ITRK	
MRW	Prudential plc		
	Admiral Group		

(b) S&P500

XLY	EOG	Marathon oil	Walgreens
XLP	Occidental Petroleum	National Oilwell Varco	STZ
XLB	VLO	BHGE	KMB

XLK	Phillips 66	XEC	KHC
XLU	Halliburton	NFX	Estee Lauder Companies
XLE stock	Marathon Petroleum	Range resources	SYY
XLF	Anadarko Petroleum	Procter & Gamble	General Mills
XLV	PXD	PEP	MNST
XLI	KMI	Philip Morris International	Dr Pepper Snapple Group
XOM	ONEOK	WMT	Kellogg Company
CVX	CXO	CVS	CAG
SLB	DVN	Altria Group	Brown-Forman
ConocoPhillips	ANDV	Colgate-Palmolive	Church & Dwight
	Apache Corp	MDLZ	HRL
			Costco stock

(c) Constituents Searched ‘as a Company’

The Berkeley Group Holdings plc	Bunzl plc	Tyson Foods Inc A	The Williams Companies Inc
Itv plc	Croda International plc	Clorox Co	Noble Energy Inc
Whitbread PLC	Mondi plc	Molson Coors Brewing Co B	Hess Corp
J Sainsbury plc	Fresnillo PLC	Hershey Foods Corp	EQT Corporation
DCC plc	Rio Tinto plc	Smucker J.M. Co	Cabot Oil & Gas A
The Royal Bank of Scotland Group plc	Severn Trent Plc	McCormick & Co	Helmerich & Payne Inc
London Stock Exchange Group plc	United Utilities Group PLC	Coty Inc	
Standard Life Aberdeen plc	Archer-Daniels-Midland Co	Campbell Soup Co	
Old Mutual plc	Kroger Co		
St James’s Place plc			
G4S plc			

Table B2 presents all the search terms used to collect the Google Trends data for this paper. Each of the search terms have been used for US, UK and global searches. Tables (a) and (b) present the search terms used when scraping the data automatically using the programming language *R*. Table (c) represents the companies, for which data was collected manually using the option of searching ‘as a company’.