Can Google Search Volume Data Help Predict Future Stock Measures?

Returns, Liquidity, Volatility

Caroline Lundström*
Stockholm School of Economics

Hampus Nestius**
Stockholm School of Economics

Bachelor Thesis
Stockholm School of Economics
Department of Finance
Tutor: Laurent Bach
May 22, 2012

Abstract

We hypothesise that search volume data from large search engines, such as Google, can be used to predict higher stock returns, higher liquidity and higher volatility. To test the validity of our hypothesis, we empirically test the relation between the public Search Volume Indices provided by Google Inc. and weekly stock returns, stock liquidity and stock volatility in a sample of OMX Nordic stocks from 2004 to 2011. We do this by performing a Fama-MacBeth regression analysis and conducting two event studies. We find that abnormal search volumes can help predict abnormal returns and abnormal liquidity in the two subsequent weeks. Moreover, we find that these relations are stronger for smaller firms than larger firms. Furthermore, we find that the predicted abnormal return is highest when high growth in search volume is observed. The results are robust to the inclusion of control variables and to the methodology. Our findings suggest that it is possible to construct a market neutral trading strategy and gain up to 11.5 percent return per year, before transaction costs and taxes, by investing each week in firms with highest search volume growth each.

We would like to thank our tutor Laurent Bach, Assistant Professor at the Department of Finance at the Stockholm School of Economics, for his comments and support throughout the writing of this thesis. He has provided great insight to the methodology of empirical finance research. Finally, we thank Johan Danielsson for helpful remarks at the final stage. Naturally, all mistakes are entirely our own.

* 21966@student.hhs.se
** 21533@student.hhs.se

Table of Contents

1 Introduction

Since 2004 Google Inc.[1] has provided a service[2] making it possible for anyone to explore online search volume data and study what people pay attention to online. By looking at historical Google search trends, Choi and Varian (2009) showed that the service could help predict home sales, tourism and automotive sales on the American market. Ginsberg et al. (2009) found that, by using the service, flu outbreaks and influenzas could be predicted one to two weeks before the outbreaks were predicted by the Centre for Disease Control and Prevention (CDC). People's search habits and online behaviour result in a wealth of information and by studying this, interesting findings about the future can be made. This paper seeks to empirically test if there exists a predictable power over the stock market as well, and more specifically if stock measure movements can be predicted by studying the online search volume data provided by Google.

Commonly used asset pricing models assume that asset prices immediately incorporate new information. For new information to instantaneously be reflected in an asset price, investors need to allocate sufficient amount of attention to the asset. However, in reality investors have limited attention and the likelihood that it is perfectly allocated among all available assets is low (Kahneman, 1973). A large number of studies have tried to understand how investor attention affects stock dynamics and have used indirect and direct measures of investor attention to test it. Examples of indirect measures of investor attention are trading volume (Barber & Odean, 2008; Gervais, Kaniel, & Mingelgrin, 2001), extreme returns (Li, Mahani, & Sandhya, 2011) and news appearance (Barber & Odean, 2008). All these papers make the critical assumption that investors immediately pay attention to newly released data on these measures. However, this assumption is criticised by Da et al. (2011) who argue that investors have the possibility to choose what to get exposed to. Additionally, simply because information becomes public it must not necessarily mean that investors pay attention to it.

Da et al. (2011) were among the first to use the search volume data presented though Google Insights as a measure of investor attention. Their arguments for using this data as a direct proxy for investor attention are that (1) the search for information reveals what people pay attention to and (2) Google is representative of all search engines due to its large market share. Da et al. (2011) studied the relation between stock returns and the quantity of Google searches made on stock ticker symbols of Russell 3000 firms. They found evidence that the Google data based on stock ticker symbols can be used to predict future stock returns and that it can be used as a direct proxy for investor attention.

The drawback of Da et al.'s (2011) study is that it is less applicable on smaller stock markets where the stocks attract relatively low attention. Google can only generate search volume data if a critical number of searches are made for a specific search query. When it comes to stocks on smaller markets, too few searches for the ticker symbols are made for Google to generate a valid data set. This thesis seeks to investigate if the findings by Da et al. (2011) hold for smaller markets where company name instead of ticker symbol has to be used as search query. We believe

[1] From here on referred to as Google.

[2] Google Insights, http://www.google.com/insights/search/

that this will contribute to the on-going discussion about investor attention as well as to the discussion about Google's predictable power. We also believe that the thesis can be seen as a supplementing study to Da et al.'s (2011) paper.

We identify three plausible reasons as to why Google search volume data should help predict positive abnormal values of stock measures. Firstly, online search query data should be a good direct proxy for attention since it measures what people are searching for and thus actively paying attention to (Da, Engelberg, & Gao, 2011; Mondria, Wu, & Zhang, 2010). Search volume data generated from using firm name as search queries can be seen as a proxy for how much attention is paid to specific firms. Although this data includes searches made by non-investors, we expect it to be a feasible proxy for investor attention. Secondly, Google is by far the most popular online search engine. As of December 2011, Google accounted for approximately 80 percent of all searches made online (Net Applications, 2012). Thus, Google is likely to be representative of the search behaviour of all Internet users. Thirdly, attention shocks are more likely to lead to net buying than net selling (Barber & Odean, 2008). The logic behind Barber and Odean's (2008) argument is that investors often are short-selling constrained and that this results in more potential buyers than sellers. Moreover, Da et al. (2011) find evidence that SVI mainly captures the attention of retail investors, who are even more short-selling constrained than institutional investors. Thus, if an attention shock is observed this extra attention is more likely to come from potential buyers than potential sellers. This indicates that an increase in search queries entered for a specific firm name would lead to a positive price pressure as well as a positive trading pressure, suggesting that stock price, liquidity and volatility are expected to increase shortly after an attention shock.

We empirically test our arguments by studying the relation between the stock measure movements of all firms listed on the NASDAQ OMX Nordic exchanges[3] and the weekly search volume data for these firms provided by Google. The sample period is from January 2004 to December 2011. The Nordic setting is interesting since the stock ticker symbols of the NASDAQ OMX Nordic listed firms do not attract enough relative interest throughout the world for Google to generate SVIs. This means that we can only perform the analysis if we use firm names as search queries. If a predictable power of Google data based on firm name exists, we should find that attention shocks identified in the data are followed by unexpected stock returns (abnormal returns), unexpected stock liquidity (abnormal liquidity) and unexpected stock volatility (abnormal volatility).

We apply two complementing methodologies to examine the validity of this reasoning. Firstly, we apply a Fama-MacBeth regression model method (Fama & MacBeth, Risk Return and Equilibrium: Empirical Tests, 1973) to study the relation between future abnormal values of stock measures and abnormal search volume. Secondly, we apply two event studies. The first one studies the relation between future abnormal values of stock returns and abnormal search volume for observations with search volume growth above a predefined threshold. The second one tests if future abnormal returns exist for observations with search volume growth over thresholds that are continuously redefined. Based on the latter we will simulate a trading strategy to see if there

[3] The list of firms listed on the NASDAQ Nordic exchanges was obtained from NASDAQ Nordic's website in March of 2012.

are any gains from studying Google search volume data. The two methodologies allow for complementary interpretations. Both are included to insure the robustness of our results.

We present three main results in this thesis. In accordance with our hypothesis, we find a strongly significant positive relation between abnormal returns and abnormal search volume, especially for smaller firms, indicating that search engine data can be used to predict future stock price movements and that it is a feasible measure for investor attention. Furthermore, we find evidence of a positive relation between abnormal liquidity and abnormal search volume, implying that stocks become more frequently traded after an attention shock. Additionally, we find an indication of a positive, yet insignificant, relation between abnormal volatility and abnormal search volume. Finally, our findings suggest that a simplified trading strategy based on search volume data can earn up to 11.5 percent in yearly return, before transaction costs and taxes.

The rest of this thesis is organised as follows. Section two discusses prior research and section three presents theory and the formulation of our hypothesis. Section four contains a description of our data and section five presents our analytical framework. Section six presents findings of the empirical study and section seven provides conclusions and suggestions for future research.

2 Previous Research

Finding potential ways of uncovering the explicit power to predict the future movements of stock measures intrigues many researchers. When searching for this power of prediction, many factors need to be considered and one important dimension is the behaviour of investors. More specifically, what is of great interest is where the attention of investors lies. This parameter is of importance, as it will determine whether stock prices will fully incorporate public information or not (Hirshleifer & Teoh, 2003). Investors need to pay attention to and be aware of the information before they can perceive and react to it (Peng & Xiong, 2006). Yet, there is a limitation to this. Since attention is a scarce cognitive ability, one cannot claim that it is constantly and evenly allocated to every stock listed on every exchange (Kahneman, 1973). Thus, what prevails is a potential opportunity to predict future stock measure movements based on information on where the desiderated attention of investors actually lies.

An Array of Passive Proxies for Investor Attention

In order to conceive the relationship between investor attention and movements in stock measures, one needs to determine how to obtain values of investor attention. However, due to the difficulty of isolating any absolute quantity of investor attention, researchers have examined the above-mentioned relationship by using different types of proxies.

Gervais, Kaniel and Mingelgrin (2001) defined a proxy for investor attention based on trading volume. Reasoning that extreme trading activity could be seen as something that investors pay attention to, Gervais et al. (2001) determined this proxy to be able to investigate whether or not it could be used to predict the future development of stock prices. Their findings include that unusually high (low) trading volume of a stock over the course of a day or week signalled an appreciation (depreciation) of that stock's price during the subsequent month. This is consistent with their initial hypothesis posing that trading activity shocks are strongly linked to the visibility of a stock, and in turn, the demand and price for that stock.

In like manner, Barber and Odean (2008) looked at trading volume, as well as two additional proxies for investor attention: extreme one-day returns and advertising. Having the same reasoning as Gervais et al. (2001), i.e. assuming that the investors pay attention to these factors, Barber and Odean (2008) hypothesised that investors only consider buying stocks that first have caught their attention. Their findings include that investors are net buyers of attention-grabbing stocks: the stocks experiencing high abnormal trading volume, the stocks experiencing extreme one-day returns or the stock visible in the news and in advertisements.

There are many other examples of research where different proxies for investor attention are used. For example, we have seen proxies such as: news headlines (Yuan, 2008), extreme levels of absolute returns (Li, Mahani, & Sandhya, 2011), advertising expenses (Chemmanur & Yan, 2009; Grullon, Kanatas, & Weston, 2004; Lou, 2008), and price limits (Seasholes & Wu, 2007). However, what characterises these measures is that they all represent an indirect measure of investor attention. This is a significant challenge for researchers who test different theories of attention. When using indirect proxies in their analyses, empiricists must make the crucial assumption that investors actually have paid attention to the stocks whose name has been visible in media or whose returns or trading activity have been at extreme levels. However, one cannot easily claim that a news article guarantees attention unless one knows for certain that all investors actually read the article.

Discovery of Active Proxies for Investor Attention

Recent studies have given birth to ideas of defining proxies for investor attention in a completely new manner. A study made by Choi and Varian (2009) in corporation with Google showed that data from Google could be used to estimate present economic activity by connecting search behaviour in the United States to the American Housing market, Automotive market, Travel market and Retail market. Similarly, Ginsberg, Mohebbi, Patel, Brammer, Smolinski and Brilliant (2009) showed that by looking at search engine behaviours in different regions in the United States they could detect movements of influenza epidemics. These findings intrigued financial empiricists who wanted to test whether search volume data could be used to predict financial activity as well.

Wu and Zhang (2010) used American Online (AOL) search query data as a direct proxy for investor attention in order to analyse the relationship between attention allocation and international investment decisions by combining US data on foreign stock holdings with the attention allocated by the customers of AOL. Their findings include that agents tend to invest more in countries where they process more information, meaning that their use of Internet search query data was valuable as a measure of investor attention.

Recent research from Da, Engelberg and Gao (2011) also proposes a direct measure of investor attention, which is based on data from a Google tool called Google Insights. Google Insights is a public web facility of Google that displays Search Volume Indices (SVI) that presents people's relative interest in certain search queries that are entered on the Google search engine. Examining the SVI for ticker symbols of Russell 3000 Index stocks, Da et al. (2011) downloaded data sets from Google containing the ticker symbols' respective historical search interest level and connected these to the historical returns of the stocks. Their findings include that this measure of investor attention was correlated with the abnormal returns of the Russell 3000 stocks.

They also concluded that a peak in the search volume for a certain stock's ticker symbol represents a predictor of higher future stock prices.

Da et al. (2011) applied the use of each stock's ticker symbol as search query with the reasoning that this captures the attention of investors and not non-investors. In this study we apply many of the methods presented by Da et al. (2011) but we apply them to a smaller region where the stock's ticker symbols have not grabbed enough attention to generate SVIs. We use the Nordic region as our landscape of research and in contradiction to Da et al. (2011), we employ the use of the actual firm names as search queries to map the retail investors' attention. We believe that by conducting this study we map retail investors' attention rather than institutional investors' attention, since retail investors are less likely to use ticker symbols and advanced tools such as Bloomberg than institutional investors are. Hence, it is more probable that the proportion of investors that use Google when searching for information on stocks in fact are retail investors and not institutional investors.

3 Theory & Hypothesis

Motivation for the Predictive Power of SVI

From the literature review we identify three plausible reasons as to why Google Search Volume Indices (SVI) can be used to predict positive stock price movements: (1) search query data is a feasible measure of investor attention, (2) the Google search engine is a good proxy for the total amount of searches made online, and (3) attention-grabbing stocks are subject to net buying and not net selling which induces a positive price pressure. These reasons are more elaborated on below.

Firstly, following the results from Mondria et al. (2010) and Da et al. (2011), we reason that Internet search query data is feasible as a measure of attention and that this data can be used as a direct proxy since it is based on people's activity on the Internet. In addition, according to Internet World Stats (2012), more than 60 percent of all people in Europe and approximately 80 percent of all people in North America are Internet users. Also, as stated by Pew Internet & American Life Project, 90 percent of all Internet users use search engines to find information (Pew Research Center, 2012). Hence, by looking at search query data one can estimate what people currently are interested in and consequently what they pay attention to. More specifically, by looking at search query data that is generated from firm names one can estimate people's interest in certain firms and consequently determine what firms people pay attention to. As discussed by Da et al. (2011), firm specific search queries can be seen as indicators of people's intentions to invest. Before deciding on whether or not to invest in a specific firm an investor often investigates the firm properly beforehand. This type of investigation could consist of finding the latest news about the company, reading the latest financial statements and/or visiting other forums where the firm might be discussed. We assume that the majority of these types of firm specific information are to be looked for by using search engines. Also, Da et al. (2011) proved that it is more likely that retail investors search for these types of firm specific information using Google than institutional investors. Furthermore, we expect that there is a time delay between the time of searching for information and the time of the stock purchase to potentially be made. Hence, by

Table 1. Market share for the largest search engines 2007-2011

Search Engine	2007	2008	2009	2010	2011
Google	79.00%	81.27%	81.92%	84.88%	82.99%
Yahoo	10.78%	9.43%	7.22%	6.16%	6.12%
Baidu	1.45%	3.12%	5.33%	3.38%	4.89%
Bing	-	-	1.89%	3.30%	3.92%
Ask	0.84%	0.58%	0.55%	0.68%	0.56%
AOL	1.84%	0.66%	0.59%	0.46%	0.40%
Excite	0.06%	0.05%	0.02%	0.03%	0.03%
Lycos	0.02%	0.02%	0.01%	0.01%	0.01%
AltaVista	0.19%	0.14%	0.09%	0.06%	0.00%
MSN	2.13%	1.78%	0.75%	0.05%	-
All the Web	0.03%	0.02%	0.02%	0.01%	-
Microsoft Live Search	2.43%	1.84%	0.65%	-	-

Table 1 shows market share per search engine between 2007 and 2011. The market share for each search engine represents its fraction of the total searches made online during a specific year. The data is provided by Net Applications Inc. and compiled from approximately 160 million online users per month, worldwide.
Source: Net Applications Inc, 2012-05-20, www.netmarketshare.com

looking at firm specific attention gathered from search engines, we expect that retail investor attention can help predict future movements in both stock prices and other stock measures. Yet, people whose intentions are not to invest certainly make up a fraction of these searches. This is covered in section five.

Secondly, we reason that the Google search engine is a good proxy for the total amount of searches that are made on different search engines. Table 1 shows that Google has largest market share among search engines. During 2004 through to 2006 Google's market share was constantly over 50 percent and growing. Since 2007, Google's market share has been laying steadily around 80 percent, (Net Applications, 2012). As Google is the most popular online search engine it should be considered the most appropriate proxy for the search behaviour of all Internet users.

Thirdly, following Barber and Odean's (2008) findings that attention-grabbing stocks are subject to net buying and not net selling, we pose that there exist more potential buyers than potential sellers. This implies that an increase in attention should indicate a positive price pressure, i.e. an increase in future stock price rather than a decrease. Barber and Odean (2008) reason that net buying results from the buy side retail investors' limitation to research each of the thousands of stocks that they can potentially buy. In contrast, the sell side retail investors do only have to research a smaller set of stocks, i.e. the stocks that they already own. Consequently, shocks in retail investor attention should on average lead to net buying rather than net selling. Incorporating the results of Barber and Odean (2008), we pose that stocks of firms with high attention are to be seen as attention-grabbing stocks. These attention-grabbing stocks are then subject to the case of net buying which in turn should be followed by an increase in stock price. Additionally, following the discussions of Gervais et al. (2001), attention shocks should imply higher trading volume, which can be used as a proxy for liquidity (Datar, Naik, & Radcliffe, 1998). Hence we expect that an attention shock can predict not only higher future returns but also higher future liquidity. Furthermore, we believe that the higher trading pressure will lead to a more volatile price.

Like in almost all cases of analyses that try to establish active trading strategies, compelling arguments exist in favour of the view that it is impossible to gain abnormal returns whilst remaining at the same level of risk. Also, there exist compelling arguments in favour of the view that there are no market inefficiencies to exploit. For example, the concept of market efficiency means that price should be adjusted for all public and private information available, (Fama, 1998). However, as Barber and Odean (2008) and Kahneman (1973) argue, this idea assumes that investors' scope of attention is equally distributed over all stocks. Due to the scarcity of investors' attention, any smoothness of allocated attention is not likely in reality. Some stocks grab more attention than others and investors choose where to place their attention. Barber and Odean (2008) argue that this leads to an inefficient mispricing on the market. From this perspective markets cannot be seen as perfect. On the other hand, an attention shock might be due to launches of new products, releases of financial data or even media publicity, which is discussed by Da et al. (2011). If this is the case, the attention shock helps prices to adjust quicker, which supports the arguments for the theory of efficient markets. Thus by studying how investors allocate their attention one can argue that markets either are inefficient in the sense that attention is not equally spread all the time or that markets are efficient in the sense that attention is moved to where it is needed. No matter which of the two perspectives we use, the conclusion of both is that prices will adjust. This implies that one should be able to predict positive stock measure movements by studying the allocation of attention.

In sum, we expect stock prices to increase after an attention shock. Following the above argumentation we see three reasons as to why this should be the case. Firstly, we expect search query data to be a feasible proxy for retail investor attention. Secondly, we consider Google as representative for all searches made online. Thirdly, there is a stronger pressure from potential buyers than sellers, which indicates that an attention shock should lead to net buying, inducing a positive price pressure. Furthermore we expect the liquidity and volatility of the stock to increase shortly after an attention shock due to the increased trading pressure.

Measuring Abnormality

Before stating our hypothesis we want to highlight that this study focuses on the relationship between abnormality in the variables, and not the relationship between the absolute magnitudes of the variables. Thus we aim to briefly touch upon the explanation of how we measure abnormality. Following Da et al. (2011), a variable's abnormal value is defined as the actual value subtracted by the expected value. To robustness check our results, we calculate abnormal values by defining different versions of expected values. For stock returns, the expected value is defined by using the CAPM with two different types of market benchmarks: (1) the country specific market benchmark[4] and (2) the Nordic market benchmark[5]. For search volume, the expected value is defined in two different ways: (a) by using the median of the eight prior weeks[6] and (b) by using the mean of the eight prior weeks[7]. The above leaves us with a set of four combinations of analytical frameworks when studying the relation between abnormal return and abnormal SVI. The ex-

[4] From here on referred to as the country specific approach.

[5] From here on referred to as the Nordic approach.

[6] From here on referred to as the median approach.

[7] From here on referred to as the mean approach.

pected value of liquidity and volatility is defined as the mean of the prior eight weeks. The definitions are more elaborated on in section five.

Hypotheses

Considering our reasoning as to why Google SVI can be used to predict positive stock measure movements, we hypothesise that by studying search volume data provided by Google Insights abnormal values in stock measures can be predicted. We hypothesise that there is a positive relation between abnormal search volume observed in the SVI and abnormal stock return. We also hypothesise that abnormal search volumes can be used to predict abnormal liquidity and abnormal volatility. Furthermore, we hypothesise that one can predict larger stock measure movements for smaller firms.

4 Data

To test the hypothesis two sets of data are needed. Firstly, we need historical stock measure data from the stocks in our sample. Secondly, we need historical search volume data to see how attention has developed over time. In this section we aim to describe the data sets used in this thesis.

Stock Measures Data

We use stock data for NASDAQ OMX Nordic listed firms, i.e. all stocks listed on the NASDAQ OMX exchanges in Stockholm, Copenhagen, Helsinki and Reykjavik. The list of stocks included in this index is obtained directly from the NASDAQ OMX Nordic web site[8]. In total, there are 571 firms listed on the OMX Nordic exchanges, of which 273 are listed in Stockholm, 167 are listed in Copenhagen, 127 are listed in Finland and 4 are listed in Iceland[9]. The stock data set contains information on returns, prices (close, bid, ask, high and low), turnover by volume and market value of company. The returns are adjusted for re-invested dividends and other changes to equity capital. Table 2 presents descriptive statistics for returns and market value of company, including descriptive statistics for SVI. A detailed and more precise description of the variables is presented in Table A2 in appendix A. Table A3 in appendix A shows all descriptive statistics, including more stock variables.

[8] See www.nasdaqomxnordic.com.

[9] The list was obtained in March 2012 and is not adjusted for historical changes to it. This might induce a survivorship bias that one should be aware of. However, if we were to include delisted firms in our sample the results might have been skewed since the attention of these firms most likely behave differently from surviving firms.

Table 2: Descriptive statistics of the sample properties

		Total sample	Small Cap	Mid Cap	Large Cap
No. of companies	Total sample	386	181	110	95
	Sweden	161	59	55	47
	Denmark	119	80	19	20
	Finland	103	41	34	28
	Iceland	3	1	2	0
No. of observations	Total sample	164 010	77 106	46 434	40 470
	Sweden	68 586	25 134	23 430	20 022
	Denmark	50 268	34 080	7 668	8 520
	Finland	43 878	17 466	14 484	11 928
	Iceland	1 278	426	852	0
Search volume index (SVI)	Mean	35.378	28.810	38.759	44.013
	Std. Dev.	24.862	23.964	24.561	23.427
	Min	0	0	0	0
	Median	37.0	28.0	40.0	46.0
	Max	100	100	100	100
Return index	Mean	0.223%	0.111%	0.294%	0.352%
	Std. Dev.	6.278%	7.214%	5.486%	5.132%
	Min	-75.000%	-75.000%	-64.193%	-41.128%
	Median	0%	0%	0%	0.314%
	Max	600.170%	600.170%	73.879%	61.795%
Market value of company (SEK)	Mean	22 025	1 028	4 340	81 946
	Std. Dev.	131 144	3 882	4 955	253 804
	Min	2	2	53	62
	Median	1 718	408	3 201	26 633
	Max	3 259 216	80 426	88 565	3 259 216

Table 2 summarizes descriptive statistics for the sample in total and for each of the capital sizes (small cap, mid cap and large cap) over the sample period 2004-2011. *No. of companies* describes the total number of firms per capital size and country, with valid SVI. *No. of observations* describes the total number of observed (and validated) SVI values for the total sample as well as per capital size and country. *Search volume index (SVI)* describes the observed SVI values for the total sample and for each of the capital sizes. *Return index* describes the return index values for the total sample and for each of the capital sizes. *Market value of company* describes the consolidated market values for the total sample and for each of the capital sizes.

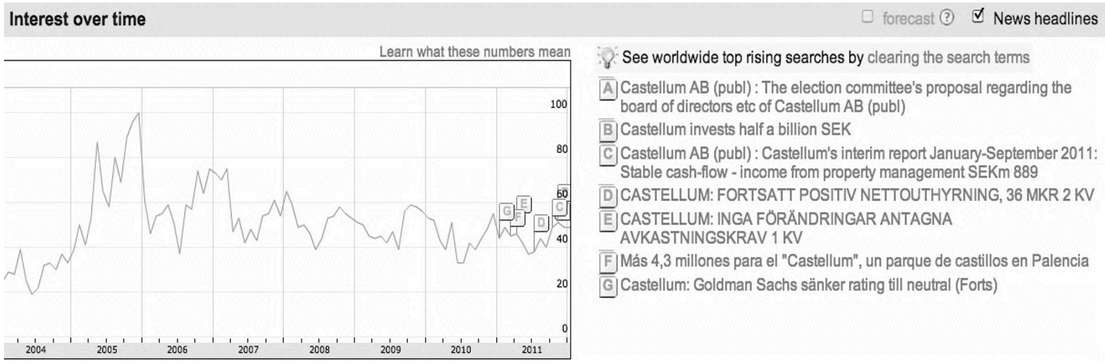
Google Search Volume Index

Each day, Google presents data on how much attention a specific search query got on the Google search engine two days ago. They present information about where in the world the searches were performed, which news is related to the peaks in attention[10] and what other search queries are related to the search query of interest. Through the tool Google Insights, a search volume index (SVI) is presented together with a downloadable file, showing how people throughout the world take interest in a certain search query.[11] The search volume data is available

[10] Since 2011

[11] Google calculates the SVI from a random subset of the actual historical search data to be able to increase the response speed. This is why SVIs generated from the same search query might be slightly different when they are downloaded at different points in time. As stated in Da et al. (2011), however, the impact of such a sampling error is small. Da et al. (2011) downloaded several SVI files per firm to compute their correlation, which they calculated to be above 97 percent on average. Thus, we reason that the findings of this study shall be applicable to new sets of downloaded SVI although they will not be identical to the particular downloads of SVI used in this sample.

Graph 1: Illustration of Google Insights Search Volume Index



Graph 1 displays a Google Insights Search Volume Index (SVI) between January 2004 and December 2011 for the search query "Castellum". Castellum Aktiebolag is a real estate company in Sweden. On the left hand side, the interest over time is presented in a graph where the values can vary between 0 and 100. On the right hand side, a random selection of news headlines that are related to certain peaks in the SVI are presented.
Source: Google Insights, 2012-03-26, www.google.com/insights/search

on a daily basis up to 90 days back in time and thereafter on a weekly basis.[12] For each day or week, a number on a scale from 0 to 100 represents the interest for that search query during that day or week, where 100 represent the highest attention observed during the time period of interest. Google scales the data by dividing each value of the index by the highest value for each specific search query and then multiply it by 100. Thus, plots for different search queries cannot be compared unless they have their respective highest value in common.[13]

Since the number of people that have access to the Internet increases, an increase in SVI could be expected to depend on either of two reasons: (1) the interest in the company is increasing or (2) the overall number of Internet users is increasing. However, for each search query entered, the corresponding index is adjusted for the fact that the total number of Google searches is changing with time. By doing that, Google Insights indirectly controls for the growth in number of computers and access to Internet. This means that the index uses the total number of Google searches as reference point. Thus, if one sees a downward slope of the SVI, it does not necessarily mean that the absolute traffic for a search query is decreasing, but that its fraction of the total Google traffic is decreasing. This implies that the index is appropriate to use as a proxy for attention as it implicitly shows how fractions of total attention on Google develop over time for different search queries.

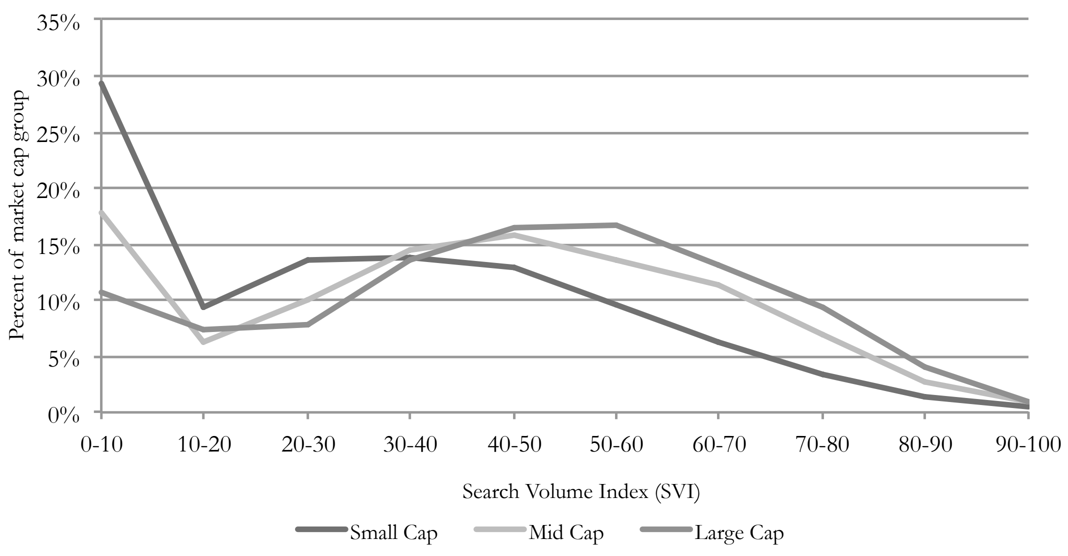
Graph 1 represents the SVI between January 2004 and December 2011 for the search query "Castellum", a real estate company in Sweden. The highest relative interest for the query "Castellum" was observed in the end of 2005. This observation is given an index value of 100. Also,

---

[12] Since this analysis covers a long time period, longer than 90 days, weekly data is used. However, studying daily data might reveal other interesting findings.

[13] However, there is a feature in Google Insights that allows a comparison between several search queries. One can group up to five search queries and then compare up to five groups. Within each group of 25 search queries, the queries will all be connected to the one that has the highest absolute search pressure. Thus it is only possible to compare the groups of 25 search queries if the groups have their respective highest value in common. This again limits the ability to compare the different SVI values unless the maximum value is shared. Nevertheless, comparing firms is of less interest to test our hypothesis since what matters is the SVI per firm and not relative SVI across firms.

Graph 2: Distribution of Search Volume Index (SVI) values per market capitalisation group



Graph 2 shows the distribution of Search Volume Index (SVI) values per market capitalisation group (small cap, mid cap and large cap) for the sample over the sample period 2004-2011. A SVI value shows how much relative attention a specific search query had during one week relative to all other weeks in the period of interest and can take any value between 0 and 100, where 100 represent the highest attention observed per firm. The graph is based on all valid SVI observations between January 2004 and December 2011 for 386 firms listed on the Nordic OMX exchanges These are grouped in ten groups based on the SVI values. The graph shows how large fraction of the total number of observations within each market cap group that is in each SVI interval. Total number of observations with valid SVI values are 76 733 for Small Cap, 46 211 for Mid Cap and 40 275 for Large Cap. Table A1 in appendix A contains more detailed information about the observation distribution within the market capitalisation groups.

Google Insights identifies the highest peaks in search volume and presents a random selection of related news headlines to the right of the index.

For this thesis, we have downloaded files with weekly search data showing the worldwide interest in the firms listed on the OMX Nordic exchanges from January 2004 to December 2011.[14] The SVIs are generated by using firm names as search queries. For Google to be able to generate these data sets, a minimum amount of search volume must exist throughout the chosen time period. For some search queries not enough relative interest was registered during the sample period for Google Insight to generate weekly data. Instead, monthly search volume data was generated. Since this longer time period complicates our ability to accurately explore the correlation between abnormal SVI and abnormal values of stock measures, we decided to drop the firms whose name only generated monthly data. Additionally, we dropped the firms that had too low search volume to generate any data. Dropping firms with monthly data or no data reduced the sample to 453 firms (118 firms were dropped)[15].

[14] The files were downloaded 21st March 2012.

[15] Since part of our hypothesis is that there exists stronger price pressure among the smaller firms, we realise that Google's truncation becomes an issue that works against us. This is because the firms that typically are the ones with relatively small amounts of data and attention also are the ones that are the smallest firms. Due to Google's truncation, we cannot explicitly explore the correlation with abnormal SVI among the smaller firms' stock returns. However, we presume that our data set comprises enough observations to test our hypothesis.

Since people most likely use the name communicated by the firm as search query (instead of the explicit firm name), we selected SVI files based on the communicated firm name instead of the explicit firm name. An example that illustrates this is the explicit firm name Tiimari Oyj Abp that was replaced by the shorter search query Tiimari.[16]

Moreover, for some firm names such as Tivoli, Sigma, Solar and Orion, the search term could be seen as ambiguous, implying a difficulty in distinguishing whether the majority of the searches really were connected to the actual firms. By looking at the news headlines that Google Insights provides for each search query, one can control for whether the entered search query seems to reflect the company or whether it seems to reflect something else. After dropping firms with ambiguous names, the sample was reduced to 386 firms (67 firms were dropped).

The sample of 386 firms is distributed over three market cap groups: Small Cap, Mid Cap and Large Cap. Graph 2 shows how large fraction of the total number of observations within each market capitalisation group that has SVI values within certain intervals. Since the maximum SVI value is fixed at 100, we know that an extremely high peak observed in search volume will lower all other observations' index values because of the scaling. This means that if a large fraction of observations has SVI values below 30, we know the data consists of high peaks in search volume, forcing all other observations to take on a lower index value. The attention curve seems to be flatter for larger firms than for smaller firms. This indicates that the peaks in attention for smaller firms are relatively higher compared to their normal level than those for larger firms are. Hence, more extreme attention shocks seem to have occurred for smaller firms. Descriptive statistics for the SVI data is presented in Table 2, together with descriptive statistics for stock measures data.

5 Analytical Framework

To investigate our hypothesis, we study characteristics of Nordic stocks listed on OMX Stockholm, OMX Copenhagen, OMX Reykjavik and OMX Helsinki stock exchanges. The Nordic setting is interesting since the ticker symbol of these stocks do not grab enough relative interest throughout not only the region in question but also throughout the rest of the world for Google to be able to generate SVIs.[17] This allows us to apply the method conducted by Da et al. (2011) to a smaller region, however modifying it to generate SVIs based on search queries derived from firm names instead of stock ticker symbols. In this study we are not interested in the direct relation between the magnitudes of search volume and returns, liquidity or volatility, but the relation between the abnormal values of search volume and abnormal values of the mentioned stock measures.

Using firm name as search query implies that the data will include searches made by non-investors since there are several purposes for searching for company names on Google. For example, the interest in a product or the interest finding the closest store. These reasons for enter-

[16] To decide whether we should use the explicit firm name or the communicated firm name as search query we downloaded 100 SVI files for 100 firms where we used (1) the explicit firm name as search query and (2) the communicated firm name as search query. Using method (1) 48 percent of the files had usable data. Using method (2) 73 percent of the files had usable data. Hence, method (2) was selected.

[17] In an attempt to use ticker symbols as search queries for 100 firms, SVI data was only available for 27 firms. The reason to this was that too few searches had been made on the ticker symbols during the selected time period for 73 of the firms.

ing search queries will cause noisy data, an issue that is discussed by Da et al. (2011). However, since we are interested in seeing how stock measures correlate with SVI, we shall not control for who is making the searches. Additionally, Google does not categorise the objectives people have when searching for a specific search query, implying that controlling for it is hard. We are aware of the fact that the data set might become noisy which can explain any potential insignificant results. In Da et al. (2011) a deeper discussion about what drives SVI is performed.[18]

We start this section by elaborating on the definitions of abnormality that was briefly touched upon in section three. Thereafter we define the analytical framework used to study the relationship between the abnormal values of the variables. This framework includes a Fama-MacBeth (1973) regression model method and two event study methods.

Abnormality

In accordance with Da et al. (2011), the abnormal value of a variable at time t is defined as the actual value at time t subtracted by the expected value at time t. To robustness check our results we have different approaches to determine the expected values. The expected value of stock returns is either defined by using the country specific approach or the Nordic approach. The expected value for search volume is either defined by using the median approach or the mean approach. The expected value for liquidity and volatility is defined as the mean of the prior eight weeks. In what follows we will describe each of the definitions for abnormal search volume, abnormal return, abnormal liquidity and abnormal volatility, respectively.

Abnormal Search Volume

If the attention paid to a firm had been equal over time, the SVI could have been illustrated as a horizontal line with a slope of zero. What is interesting to study is what happens with stock measures close to the events when search volume deviates from what is expected, hence when abnormal search volumes are observed. Abnormal SVI for firm i at time t, $ASVI_{i,t}$, is defined as

$$ASVI_{i,t} = log\ (SVI_{i,t}) - log\ [E(SVI_{i,t})] \quad (1)$$

Hence, abnormal SVI for firm i at time t is defined as the log of SVI for firm i at time t minus the log of expected SVI for firm i at time t.[19] Expected SVI is defined by using two different approaches. In the first approach, the median approach, expected SVI is defined as the median value of SVI during the prior eight weeks

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

In the second approach, the mean approach, expected SVI is defined as the mean value of SVI during the prior eight weeks

$$E(SVI_{i,t}) = mean(SVI_{i,t-1}, SVI_{i,t-2}, \dots, SVI_{i,t-8}) \quad (3)$$

We argue that the median approach is more appropriate to use since one-week peaks in search volume that might have appeared during the eight-week period will have less affect on the ex-

[18] Da et al. (2011) conclude that SVI is positively correlated to the size of the stock, extreme stock returns and abnormal turnover. Also, they conclude that the occurrence of news matters, suggesting that a stock with lots of recent news coverage is less likely to receive "unexpected" attention. However, because of low R2-values, they also conclude that the existing proxies of investor attention only explain a small fraction of variations in SVI.

[19] This implies that observations with SVI value zero disappear from the study since the logarithm of zero returns a missing value. Zero values mean that Google could not generate SVI for a specific week. It must not represent zero attention and including it in the analysis as zero attention would be misleading.

Table 3: Descriptive statistics of the variable properties

		Total sample	Small Cap	Mid Cap	Large Cap
No. of observations	Total sample	164 010	77 106	46 434	40 470
	Sweden	68 586	25 134	23 430	20 022
	Denmark	50 268	34 080	7 668	8 520
	Finland	43 878	17 466	14 484	11 928
	Iceland	1 278	426	852	0
Abnormal SVI (Median approach)	Mean	34.744	31.335	35.487	40.479
	Std. Dev.	23.899	23.273	23.977	23.869
	Min	0	0	0	0
	Median	36.5	32.5	37.0	42.5
	Max	99.0	97.5	99.0	99.0
Abnormal SVI (Mean approach)	Mean	35.387	28.809	38.790	44.014
	Std. Dev.	23.346	22.073	23.129	22.285
	Min	0	0	0	0
	Median	36.875	28.625	40.875	46.375
	Max	96.625	96.625	94.625	95.875

Table 3 summarizes descriptive statistics for the sample in total and for each of the capital sizes (small cap, mid cap and large cap) over the sample period 2004-2011. *No. of observations* describes the total number of observed (and validated) SVI values for the total sample as well as per country and capital size. *Abnormal search volume (Median approach)* describes the abnormal SVI values, defined by using the median approach, for the total sample and for each of the capital sizes. *Abnormal search volume (Mean approach)* describes the abnormal SVI values, defined by using the mean approach, for the total sample and for each of the capital sizes.

pected value. If expected SVI is defined using the mean approach, a one-week peak (drop) in SVI will potentially have a large impact on the normal value and skew it upwards (downwards). Additionally, a shorter time span than eight weeks might be more affected by few one-week peaks and a larger time span than eight weeks could become non-representative for what is "normal" at the time of interest, since it could imply effects from time trends and seasonality.[20]

Since we are interested in the dynamics of each firm's specific abnormal SVI and since SVI is scaled to always stay between 0 and 100, abnormal SVI can be compared across firms in a cross-section analysis, regardless the firm size etc. Combining (1) and (2) as well as (1) and (3) lead to the respective explicit expressions for abnormal SVI

$$ASVI_{i,t} = log\ (SVI_{i,t}) - \ log\ [median(SVI_{i,t-1}, SVI_{i,t-2}, \dots, SVI_{i,t-8})] \quad (4)$$

$$ASVI_{i,t} = log\ (SVI_{i,t}) - \ log\ [mean(SVI_{i,t-1}, SVI_{i,t-2}, \dots, SVI_{i,t-8})] \quad (5)$$

Summary statistics of abnormal SVI defined by using the two different approaches is presented in Table 3. The table highlights that the mean approach generates, on average, higher abnormal SVI even though the median approach results in the most extreme values of abnormal SVI.

Abnormal Return

Since we are interested in investigating whether abnormal search volume has any predictable power over abnormal weekly stock returns, we define abnormal weekly stock return as

$$AR_{i,t} = r_{i,t} - E\ (r_{i,t}) \quad (6)$$

[20] Our main results are robust to other choices of periods for obtaining the expected value (4 weeks, 6 weeks, 10 weeks etc.).

Hence, abnormal weekly stock return is calculated by subtracting the expected weekly return of firm *i* at time *t* from the actual weekly return of firm *i* at time *t*. The expected weekly return should be comprehended as the market risk adjusted weekly return of firm *i* at time *t*, following the CAPM expression

$$E(r_{i,t}) = r_{f_i,t} + \beta_{i,t} * (r_{m_i,t} + r_{f_i,t}) \tag{7}$$

We define market return, $r_{m_i,t}$, by using two approaches: (1) the country specific approach where country specific indices[21] are used and (2) the Nordic approach where the OMX Nordic Index is used. The risk free rate of return, $r_{f_i,t}$, is defined depending on what stock exchange each firm *i* is listed on by using the price development of 1 month STIBOR[22], CIBOR[23], HELIBOR[24] and REIBOR[25], respectively. The market betas, $\beta_{i,t}$, are firm specific and are estimated by regressing each stock's past three-year excess returns on the market benchmark's excess return at each time *t*.

Abnormal Liquidity

As stated above we are also interested in studying the relation between abnormal SVI and abnormal liquidity. We use two proxies for liquidity in this study: bid-ask spread (Demsetz, 1968) and trading volume (Datar, Naik, & Radcliffe, 1998). Firstly, we define abnormal liquidity based on bid-ask spread, $ALBA_{i,t}$, as the actual bid-ask spread, $ba_{i,t}$, minus the expected bid-ask spread, $E(ba_{i,t})$, where the expected value is defined as the mean of bid-ask spread during the prior eight weeks at time *t* for firm *i*, resulting in

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

Secondly, abnormal liquidity based on trading volume, $ALTV_{i,t}$, is defined in a similar way, subtracting the mean of trading volume, $tv_{i,t}$, during the prior eight weeks from the actual trading volume at time *t* for firm *i*, resulting in

$$ALTV_{i,t} = tv_{i,t} - mean(tv_{i,t-1}, tv_{i,t-2}, \dots, tv_{i,t-8}) \tag{9}$$

Abnormal Volatility

To be able to study the dynamics of volatility in the case of an attention shock, we need a measure of volatility that depends on present stock prices and not historical prices. We believe that a volatility measure that takes into account too many historical values may underestimate any new peak in volatility due to the smoothing effect of measuring volatility over a long time period. Since we do not want to complicate the identification of any distinctive changes in volatility, we define our proxy for weekly volatility for firm *i* at time *t* as the average of intra-day high-minus-low values per week

$$\tilde{\sigma}_{i,t} = mean(ph_{i,t,1} - pl_{i,t,1}, ph_{i,t,2} - pl_{i,t,2}, \dots, ph_{i,t,5} - pl_{i,t,5}) \tag{10}$$

---

[21] OMX Stockholm Index, OMX Copenhagen Index, OMX Helsinki Index or OMX Reykjavik Index depending on where firm *i* is listed.

[22] Stockholm Interbank Offered Rate, available at www.riksbank.se.

[23] Copenhagen Interbank Offered Rate, available at www.nationalbanken.se.

[24] Helsinki Interbank Offered Rate, available at www.soumenpankki.fi.

[25] Reykjavik Interbank Offered Rate, available at www.sedlabanki.is.

where $ph_{i,t,n}$ and $pl_{i,t,n}$ represent the intra-day highest and lowest price of the stock for firm *i* at time *t*, where *n* ranges from 1 (Monday) to 5 (Friday). The high-minus-low measure was discussed as proxy for volatility by Parkinson (1980) already during the early 1980s and is nowadays still commonly used as a fair proxy for volatility (Goyenko, Holden, & Trzcinka, 2009). We define abnormal volatility, $AV_{i,t}$, as the actual volatility proxy minus the expected volatility proxy, where the expected value is defined as the mean of the volatility proxy during the prior eight weeks at time *t* for firm *i*, resulting in

$$AV_{i,t} = \tilde{\sigma}_{i,t} - mean(\tilde{\sigma}_{i,t-1}, \tilde{\sigma}_{i,t-2}, \dots, \tilde{\sigma}_{i,t-8}) \qquad (11)$$

The variables defined in this section are of great importance for the framework presented below and will be the dependent and independent variables of interest in the Fama-MacBeth regression model and in the event study models.

Analytical Framework to Study the Relation Between the Variables

The hypothesis will be tested by using two methodologies. Firstly, we will test our hypothesis by using a Fama-MacBeth (1973) cross-sectional regression model. The Fama-MacBeth setting enables us to empirically study the relation between abnormal SVI and abnormal values of stock measures for the stocks in our sample whilst simultaneously and automatically controlling for time-fixed effects such as large macro-economic shocks. Secondly, we will perform two event studies in which we focus on the most extreme observations of growth in SVI to see if the most extreme events can predict even higher abnormal stock returns. The findings from the event studies will be used to simulate a trading strategy, exploring whether one could gain from studying search volume data or not.

The Fama-MacBeth Regression Model Specification

The Fama-MacBeth (1973) regression method has been used in previous literature to study the relation between retail investor attention (proxied by abnormal SVI) and stock measures (returns, volatility, trading volume, bid-ask spread) (Chemmanur & Yan, 2009; Da, Engelberg, & Gao, 2011).

Skoulakis (2006) proved that Fama-MacBeth cross-sectional regression is more efficient than OLS regression for panel data that both have a large cross-section and a long time series of data, in the sense that it estimated more accurate t-statistics. By using the Fama-MacBeth regression method, all regression variables are cross-sectionally demeaned and all independent variables are also standardised. This implies that the regression coefficients will represent the effect on the dependent variable from a one-standard-deviation change in the corresponding independent variables. The use of Fama-MacBeth cross-sectional regression accounts for time-specific shocks that affect the entire data set. The standard errors are computed using the Newey-West (1987) formula with eight lags, to correct for any effects of autocorrelation and heteroskedasticity in the error terms, which is something that might occur in a cross-sectional regression.

By conducting this study we expect to see a positive relation between abnormal SVI and abnormal returns since an abnormal value of SVI indicates an increase in investor attention. Furthermore, we expect to see a positive relation between abnormal SVI and abnormal liquidity, which is proxied both by abnormal bid-ask spread and by abnormal trading volume. We also expect to see a positive relation between abnormal SVI and abnormal volatility.

In our main regression we respectively regress future abnormal returns for ten subsequent weeks on abnormal SVI, log of market value of company and an interaction term between the two. Thus, the regression specification is as follows

$$AR_{i,t} = \beta_0 + \beta_1 * ASVI_{i,t} + \beta_2 * LMVC_{i,t} + \beta_3 * ASVI_{i,t} * LMVC_{i,t} + a_i + u_t + \varepsilon_{i,t} \quad (12)$$

where $AR_{i,t}$ is the abnormal return of firm *i* at time *t*, $ASVI_{i,t}$ is the abnormal SVI of firm *i* at time *t*, $LMVC_{i,t}$, is the log of market value of company for firm *i* at time *t* (included as a control variable), a_i represents the firm fixed effects, u_t represents the time fixed effects and $\varepsilon_{i,t}$ represents the residuals. The main coefficients of interest are β_1 and β_3, since they describe the relation between abnormal SVI and abnormal returns. The regression results for different definitions of abnormal SVI (median and mean approach) and for different calculation of abnormal return (country specific and Nordic approach) are presented in Table B1 in appendix B.

We also perform alternative regressions where we control for other variables. Firstly, we control for one week lagged abnormal return to eliminate pure momentum effects. This results in the first alternative regression specification

$$AR_{i,t} = \beta_0 + \beta_1 * ASVI_{i,t} + \beta_2 * LMVC_{i,t} + \beta_3 * ASVI_{i,t} * LMVC_{i,t} + \beta_4 * AR_{i,t-1} + a_i + u_t + \varepsilon_{i,t} \quad (13)$$

where $AR_{i,t-1}$ is the previous week's abnormal return. All other variables are defined above. The regression results are presented in Table B2 in appendix B, using both the median and mean approach as well as the country specific and Nordic approach. Secondly, we add squared abnormal SVI to the main regression model in equation (12) to study if we could expect diminishing returns instead of linear returns. The second alternative regression specification is

$$AR_{i,t} = \beta_0 + \beta_1 * ASVI_{i,t} + \beta_2 * LMVC_{i,t} + \beta_3 * ASVI_{i,t} * LMVC_{i,t} + \beta_4 * ASVI_{i,t}^2 + a_i + u_t + \varepsilon_{i,t} \quad (14)$$

where $ASVI_{i,t}^2$ is the squared abnormal SVI for firm *i* at time *t*. All other variables are defined above. The regression results are presented in appendix Table B3 in appendix B, also this for all approaches to define abnormality.

To study the relation between abnormal liquidity and abnormal SVI we use the same independent variables as in the main regression model in equation (12). Instead of abnormal return as dependent variable, we use abnormal bid-ask spread and abnormal trading volume. The regression specification is as follows

$$\boldsymbol{AL_{i,t}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1} * ASVI_{i,t} + \boldsymbol{\beta_2} * LMVC_{i,t} + \boldsymbol{\beta_3} * ASVI_{i,t} * LMVC_{i,t} + a_i + u_t + \boldsymbol{\varepsilon_{i,t}} \quad (15)$$

where $\boldsymbol{AL_{i,t}}$ is a variable vector for the variables $ALBA_{i,t}$ and $ALTV_{i,t}$ defined in equation (8) and (9).

Using a similar methodology the relation between abnormal volatility and abnormal SVI is studied, resulting in the regression specification

$$AV_{i,t} = \beta_0 + \beta_1 * ASVI_{i,t} + \beta_2 * LMVC_{i,t} + \beta_3 * ASVI_{i,t} * LMVC_{i,t} + a_i + u_t + \varepsilon_{i,t} \quad (16)$$

where $AV_{i,t}$ is the abnormal volatility proxy for firm *i* at time *t*. The regression results from equations (15) and (16) are presented in Table B4 in appendix B. For these regressions results are presented for the country specific approach combined with the median and mean approach.[26]

[26] Note that since we are interested in investigating the relation between abnormal SVI and abnormal values of stock measures we are concerned about correlation and not causality and we should not worry about omitted variable bias or what drives SVI. Recall footnote 18.

Table 4: Descriptive statistics of the passive event study groups

SVI growth	i > p50	i > p60	i > p70	i > p80	i > p90
No. of observations	56 914	41 596	31 574	20 975	10 404
Mean	15.48%	20.91%	25.98%	34.12%	49.96%
Std. Dev.	24.96%	27.25%	29.51%	33.35%	41.62%
Min	0%	2.99%	7.14%	13.33%	25.64%
Median	8.57%	13.33%	18.00%	25.40%	40.00%
Max	1566.67%	1566.67%	1566.67%	1566.67%	1566.67%

Table 4 summarizes descriptive statistics for the passive event study groups over the sample period 2004-2011. Each group consists of all observations in the sample over the entire sample period with a SVI growth over the Xth percentile, where X can take the values 50, 60, 70, 80 and 90.

The Event Study Model Specification

The event study model aids the study of how certain variables develop after certain events. This model is not widely used in previous literature, but we find it useful for explaining how stock returns develop after attention shocks of different magnitudes. In this analysis, we will use the event study framework to explain how abnormal stock returns develop after attention events where the events are identified by looking at the magnitude of growth in SVI.[27] Growth in SVI, $SVIG_{i,t}$ is defined as net growth

$$SVIG_{i,t} = \frac{SVI_{i,t}}{SVI_{i,t-1}} - 1 \quad (17)$$

where $SVI_{i,t}$ represents SVI for firm *i* at time *t* and $SVI_{i,t-1}$ represents SVI for firm *i* at time *t-1*. Descriptive statistics for SVI growth is presented in Table A3 in appendix A.

Since we look at the growth in SVI instead of the abnormal SVI not all observations with high abnormal SVI will be included in the analysis. More specifically, if several weeks' observed SVI are defined as abnormal, we will only include the first observed abnormal SVI of these weeks in the event study, since that is the observation with highest observed growth in SVI. Hence, an event is defined when the SVI growth is above a certain threshold. By altering the threshold we can examine what magnitude of growth in SVI is expected to constitute an event that can predict highest abnormal returns.

Passive Event Study: In the passive event study, the thresholds are defined by looking at SVI growth for all observations in the sample, meaning all historical observations for all firms over the entire sample period (from January 2004 to December 2011). All observations from all weeks and firms are ranked based on SVI growth. The thresholds are defined as the 50th, 60th, 70th, 80th and 90th percentile of growth in SVI for these observations. The thresholds sort the observations into five groups: those above the 50th percentile, those above the 60th percentile, those above the 70th percentile, those above the 80th percentile and those above the 90th percentile. Summary statistics for the groups are presented in Table 4.

Using this definition of the thresholds, we run our main Fama-MacBeth regression stated in equation (12) for all observations with SVI growth above each threshold as well as for all obser-

[27] To limit the scope of this thesis we have chosen not to perform an event study on other stock measures than stock returns.

vations that lie in-between two thresholds (i.e. that are within a decile group[28]). In the passive event study we use the median approach to define abnormal SVI. Additionally, we conduct the event analysis by using both the country specific approach and Nordic approach. The dependent variable used in this study is the abnormal return during the first, second and third weeks subsequent to the event.[29]

This study is conducted to test if (1) the predictable power is stronger for the most extreme values of abnormal SVI than for the less extreme values of abnormal SVI, (2) the opposite relation exist, or (3) if the predictable power is equal for all levels of SVI growth. Additionally, we test whether the predictable power is stronger for observations within certain decile groups.

The benefits of this study are that each group will contain a large set of observations and that the thresholds will be set to one specific magnitude of SVI growth, i.e. not varying with time. The potential drawback of this study is that the thresholds cannot be determined before all data is gathered since it is based on all observations between 2004 and 2011. Neither, it cannot be used to make a weekly evaluation of whether an observed SVI is above a certain threshold or not, solidly based on that week's observations. This is because the observed SVI data point cannot be allocated into a group unless all data is collected.

In the following section, we present a second event study method in which we use an active definition of the threshold to be able to evaluate observations on a weekly basis and determine which ones could be defined as an event.

Active Event Study: To solve the potential drawback stated above, we develop a way to study the relationship between SVI growth and abnormal returns where SVI growth can be evaluated actively and continuously. In the active event study, we define new thresholds each week that are based on all firms' most recently observed SVI value. This differs from the passive event study in the sense that we continuously define new thresholds each week based on new observations instead of defining a fixed set of thresholds that are based on the entire set of data. Instead of running a Fama-MacBeth regression like the main regression presented in equation (12), we perform Student's t-tests to see if the average abnormal returns occurring one to two weeks subsequent to the event for the specific thresholds are significantly different from zero. The abnormal return is defined using the country specific approach.

The potential drawback of the active event study is that the lowest value in each decile group, i.e. the threshold, can vary notably between weeks. For example, the value of the 80th percentile that defines the threshold might represent SVI growth of 20 percent one week and 60 percent the next. However, this might also add extra value to this study. If we find that the observations with top growth in SVI are expected to earn abnormal returns 1-2 weeks after the event, disregarding the magnitude of the growth, the potential relation between abnormal returns and abnormal SVI found in the prior analyses can have complementary findings showing that it is not only a large abnormal SVI value but a large growth in SVI relative to other firms that is necessary to predict a high abnormal return.

---

[28] The decile groups contain the observations between the 50th-60th percentiles, the 60th-70th percentiles, the 70th-80th percentiles, the 80th-90th percentiles and the 90th-100th percentiles.

[29] Based on the findings from the Fama-MacBeth regression method the event study analysis is only performed for the three weeks subsequent to the event.

Based on the findings of the active study we develop an interpretation of a simple trading strategy to see whether it is possible to gain from trading based on information from Google Insights SVI. If we find a significant relation between SVI and returns we will simulate how a market neutral strategy can gain from studying search volume data. We will use the findings of the active event study to get an estimation of how much one can gain per week and scale the result to yearly return.

6 Empirical Results

In line with our hypothesis, our findings suggest that SVIs provided by Google Insights can be used to predict positive abnormal returns. Additionally, our findings indicate that SVI can help predict liquidity. However, in contradiction to our hypothesis, we find no significant evidence that SVI can be used to predict abnormal volatility. Furthermore, holding abnormal SVI fixed we find that predicted abnormal values of stock measures are higher for small firms, as hypothesised. Finally, by simulating a trading strategy based on the analysis of the predictive power of SVIs we find that a potential return up to 11.5 percent can be earned per year.

Structure

The results of the Fama-MacBeth regression method are presented in appendix B. In this analysis we use ten dependent variables: abnormal return for each of the ten subsequent weeks. These results are presented for both definitions of normal SVI level: (1) the median approach and (2) the mean approach. The results are also presented for the two definitions of expected return: (a) the country specific approach and (b) the Nordic approach.[30]

We firstly present the main regression results from equation (12) in Table B1, where abnormal search volume, log of market value of company and an interaction term between the two are used as independent variables. Secondly, we present the results from equation (13) in Table B2, where we control for one week lagged abnormal return. This control variable might reduce the significance of our results since it implies that we control for historical performance. Thirdly, in Table B3, we included the squared abnormal SVI presented in equation (14) to study whether the relationship between abnormal returns and abnormal SVI is linear or squared. Fourthly, Table B4 shows the relation between abnormal SVI and abnormal liquidity, and between abnormal SVI and abnormal volatility, based on the Fama-MacBeth regressions presented in equation (15) and (16).

The results of the event studies are presented in appendix C. These results are presented for the median approach of defining abnormal SVI. The results of the passive event study are presented for the two definitions of expected return, while the results of the active event study are presented for the country specific benchmark approach of defining abnormal return. Depending on how the expected values are defined, the abnormal values will differ which will give different results.

[30] This is applied to Table B1, B2 and B3 which present the results from the regressions that use abnormal return as dependent variable. For Table B4, i.e. the table that presents results from the regression where abnormal trading volume, abnormal bid-ask spread and abnormal volatility respectively are used as the dependent variable, the results are only presented for the country specific approach, but for both the median and mean approach.

In appendix C, all results are divided into two panels depending on whether the analysis is performed over (i) all observations over each specific threshold or over (ii) all observations between two thresholds (within decile groups). Table C1.1 and C1.2 present the results from the passive event study, where the main regression presented in equation (12) is performed for each definition of event threshold and dependent variable. In this analysis we use three dependent variables: abnormal return for each of the first, second and third week subsequent to the event. These results are presented for the two definitions of expected return: (a) the country specific approach and (b) the Nordic approach. Table C2 presents the results from the Student's *t*-tests performed in the active event study, where the event thresholds are continuously and actively defined. The dependent variables are the abnormal return for each of the first and second subsequent week and the cumulative abnormal return over the two.

According to the above structure the results can be summarized as follows. If abnormal SVI is defined according to the median approach, we find a positive correlation between abnormal SVI and abnormal returns for the first subsequent week and the second subsequent week at a statistically significant level (Table B1). Adding one week lagged abnormal return (Table B2) as control variable does not affect the results significantly. Additionally, the results in Table B3 show that the relation between abnormal SVI and abnormal returns is more likely linear than squared. Table B4 shows significant evidence of a positive relation between the trading volume proxy for abnormal liquidity and abnormal SVI. However, it does not show any significant evidence of a relation between abnormal volatility and abnormal SVI, but an indication of a positive relation.

From the passive event study (Table C1) we find significant evidence that the highest attention shocks between 2004-2011 could predict the highest abnormal returns. However, the active event study (Table C2) finds that when continuously defining the event thresholds the highest abnormal returns could be predicted by the observations in the 80^{th}-90^{th} decile group.

Before going deeper into the results from the different sub studies, we briefly discuss how well the different approaches to define abnormality performed.

Evaluation of the Approaches to Define Abnormality

When studying the relation between abnormal SVI and abnormal return the median approach to define abnormal SVI yields the most significant results (Table B1, B2 and B3). However, when studying the relation between abnormal SVI and abnormal trading volume the mean approach yields the most significant results (Table B4). Hence, it is not clear how abnormal SVI should be defined. To us, the median approach should be more appropriate than the mean approach following the discussion in section five saying that if the mean approach is used, attention peaks that occurred during the prior eight weeks period will have too large effect on the normal level of attention. However, arguments for using the mean approach exist as well. If abnormal SVI is estimated using the mean approach, one-week attention peaks and attention drops affect the estimated normal level of attention. Given that the magnitudes of the peaks are equal to the magnitude of the drops, this would leave the mean close to the median value. For stocks that tend to have more peaks than drops, the mean approach will incorporate this when estimating the normal level of search attention. However, the reason as to why the median approach yields more significant results when studying stock returns and why the mean approach yields more significant results when studying trading volume is not clear.

The discussion above emphasises the importance of knowing how to define abnormal SVI to be able to predict stock measure movements. Neither the median nor the mean approach yields significant results when studying the relation between abnormal SVI and abnormal bid-ask spread, and between abnormal SVI and abnormal volatility. However, this must not necessarily mean that no relations exist, but that the two approaches to define abnormal SVI might not be the most appropriate ones to use for the purpose of studying these relations. Since the mean approach yields no significant results when abnormal return is studied, the median approach is used in the passive and active event study analysis.

The results are hardly affected by which of the two market indices we use when defining expected return. This suggests that our findings are robust to the use of both benchmark approaches. Henceforth we will present the results in parallel: the results from the Nordic benchmark approach will be presented in brackets and the results from the country specific benchmark approach will be presented without brackets. This holds for all results except for those covering liquidity, volatility and the active event study. For those, only results from the country specific approach are presented.

Results from the Fama-MacBeth Regression Model

The results from the Fama-MacBeth main regression are presented in Table B1. Given this setting, we find evidence that a one standard deviation increase in abnormal SVI predicts a statistically significant positive abnormal return of 21.628 [19.299] basis points for the first subsequent week and 17.919 [15.003][31] basis points for the second subsequent week, at a 10 percent significance level. Additionally we find a significantly negative coefficient on the interaction term between abnormal SVI and market value of company for the first subsequent week and for the second subsequent week (-21.239 [-18.531] and -18.003 [-14.903] basis points). This suggests that an increase in abnormal SVI is followed by a relatively larger price increase for smaller firms than larger firms, which is in support of our hypothesis. The magnitude of the interaction coefficient indicates that the positive price pressure is much stronger among smaller firms than large firms. Furthermore, if SVI is considered to be a proxy for retail investor activity and not institutional investor activity, following the results of Da et al. (2011), the coefficient of the interaction-term shows that retail investor activity has a larger price impact on smaller firms than on larger firms.

We see no significant results when defining abnormal SVI using the mean approach. Neither do we find any evidence of significant positive or negative abnormal returns for any other subsequent week on any conventional level. This indicates that the price pressure lasts for no more than two weeks and since we do not see any significant price decline thereafter this could be supportive of the idea that attention is allocated to where it is needed and that attention shocks make prices adjust to appropriate levels. However, this also supports the argumentation that markets are not perfect since there on average exists a two-week price adjustment period. The findings from this analysis are in line with, and extend, the results presented by Da et al. (2011), i.e. that SVI data can be used to predict abnormal stock returns on markets where ticker symbols attract relatively low interest, such as the Nordic region.

The results from the Fama-MacBeth regression with one week lagged abnormal return as control variable are presented in Table B2. Controlling for one week lagged abnormal return does

[31] Not significant at a 10 percent significance level.

not substantially alter our results. The coefficient of abnormal SVI for the second subsequent week is still significant at a 10 percent significance level and of the same sign, magnitude and t-statistic as in Table B1. The second subsequent week an abnormal return of 18.189 [14.934][32] basis points is expected, given a one standard deviation increase in abnormal SVI. This suggests that our findings are robust even after controlling for one week lagged abnormal return.

The results of including squared abnormal SVI to the main regression are presented in Table B3. The inclusion of this variable does not alter our results, since the coefficient of the included variable is not significant at any conventional level and leaves the coefficient for the first level variable relatively unchanged. The first and second subsequent week an abnormal return of 21.947 [20.194] and 18.233 [15.815] basis points, respectively, is expected given a one standard deviation increase in abnormal SVI. Thus, we cannot show that the relation between abnormal SVI and future abnormal returns is squared. This implies that we do not see any diminishing abnormal returns as abnormal SVI increases. Instead we expect to see a linear relation between abnormal returns and abnormal SVI.

Table B4 shows the results of studying the relation between abnormal liquidity and abnormal SVI, and between abnormal volatility and abnormal SVI. This table has three important implications. Firstly, we see that the choice of proxies has a larger impact than expected. As previously discussed, both trading volume and bid-ask spread are stated to be feasible proxies for liquidity (Datar, Naik, & Radcliffe, 1998; Demsetz, 1968). However, when bid-ask spread is used as a proxy for liquidity, we see no significant evidence of any predictable power in SVI, whereas the trading volume proxy does leave us with significant evidence. Consequently, this implies that the analysis is sensitive to the skill of selecting proxies. This might also explain why we do not find any significant correlation between abnormal volatility proxy and abnormal SVI. Secondly, when it comes to the proxy for volatility, the coefficient of interest is positive, however not significant[33], indicating that an increase in abnormal SVI could lead to an increase in volatility. In conjunction with the results from the main regression presented in equation (12), this supports the logic that the expected increases in stock price during each of the two subsequent weeks are volatile. Following the arguments above, using a different proxy for volatility might yield a more significant result.[34] Thirdly, if trading volume is used as a proxy for liquidity, a one standard deviation increase in abnormal SVI is expected to predict an increase in trading volume by 2.341 million SEK the first subsequent week and 1.709 million SEK the second subsequent week. In accordance with Ding and Hou's (2011) findings, this indicates that stocks are expected to become more liquid after an attention shock. Furthermore, this is both in line with Barber and Odean's (2008) price pressure hypothesis and with our hypothesis that liquidity will increase after attention shocks. Additionally, by looking at the coefficient of the interaction term we see that the effect is larger for smaller firms, indicating that the trading volume of smaller firms is much more attention driven than that of larger firms. Most likely, large firms are already frequently traded and attention shocks will barely affect their trading volume. Additionally, for the bid-ask proxy

[32] Not significant at a 10 percent significance level.

[33] Not significant at any conventional significance level

[34] Recall from the first implication of Table B4 that the analysis is sensitive to the choice of proxy and if a different volatility proxy is used, the result might change and might be significant. Goyenko, Holden and Trzcinka (2009) discusses different proxies for volatility that would be possible to use instead. However, we leave the analysis of other proxies for volatility to future research.

for liquidity, the coefficient of abnormal SVI is negative, indicating that bid-ask spread decreases after an attention shock. This can be interpreted as an increase in liquidity, which is in line with the findings from the trading volume analysis. But the low t-statistic shows that no significant conclusion can be drawn from the results of the bid-ask study.

In sum, the Fama-MacBeth regression method finds evidence in support of our hypothesis posing that the SVIs provided by Google Insights can be used to predict positive abnormal returns. As hypothesised, we find that the predictable power of the SVI is higher for smaller firms than for larger firms. The main results are robust, both after controlling for market value of company and one week lagged abnormal return. Additionally, the relation between abnormal SVI and abnormal returns is expected to be linear. Finally, the Fama-MacBeth regression method indicates a positive relation between abnormal SVI and liquidity and between abnormal SVI and volatility, however not significant. In almost all regressions the R-squared value is low (below 0.1) which we suspect is due to the fact that abnormal SVI does not explicitly explain abnormal return, abnormal liquidity and abnormal volatility, but rather gives an indication of future movements in these stock measures.

Results from the Event Study Model

The results from the passive event study method are presented in Table C1.1 and Table C1.2 in appendix C. We find significant evidence that higher SVI growth predicts higher abnormal returns, especially for the second subsequent week. This is in consensus with the findings from the Fama-MacBeth regression method.

In the passive event study method we grouped all observations (for all firms) between 2004-2011 based on SVI growth and studied how abnormal returns developed for the firms within these groups after attention shocks. As presented in Table C1.1, for the observations with SVI growth above the 50^{th} percentile, a one standard deviation increase in abnormal SVI is expected to predict a 9.209 [13.733] basis points abnormal return during the second subsequent week. Similarly, for the observations with SVI growth above the 60^{th}, 70^{th}, 80^{th} and 90^{th} percentile threshold values, a one standard deviation increase in abnormal SVI is expected to predict a 13.056 [12.479], 16.562 [15.812], 23.169 [23.015] and 24.051 [17.760] basis points abnormal return for the second subsequent week. These results are not significant for all groups, but indicate that the price pressure is higher for those observations with highest SVI growth. Additionally, if the observations above the 50^{th} percentile are divided into decile groups based on SVI growth, we find that the results for the two subgroups with lowest SVI growth[35] become insignificant. This indicates that the highest predictable power of SVI emerges in the observations that have the highest SVI growth: the 80^{th}-90^{th} decile group and 90^{th}-100^{th} decile group. Again we find that the difference between using the Nordic and country specific benchmark approaches is small, which indicates that our results are robust to the choice of benchmark index. The results from the passive event study both validate the results from the Fama-MacBeth regression model and add to our findings that the most extreme events of the time period predict higher abnormal returns. However, the passive event study cannot be used to continuously evaluate what firms have extreme abnormal attention shocks.

[35] The two decile subgroups containing the observations between the 50th and 70th percentiles.

In the active event study thresholds were continuously redefined at each time *t*. Each week, all observations of that week were ranked based on the growth in SVI and if an observation is above a certain threshold we define it as an event. For example, any firm's observation at time *t* with SVI growth above the Xth percentile would constitute an event. We then studied the abnormal returns for all event observations of all weeks, to see what an event observation on average could predict in abnormal returns. We simulated the results for different thresholds, starting with the observations above the 50th percentile each week, followed by those above the 60th, 70th, 80th and 90th percentile respectively. The results from the Student's *t*-tests are presented in Table C2, showing that the active thresholds can be used to predict in what stocks cumulative abnormal returns over the two subsequent weeks can be found. For the observations with SVI growth above the 70th percentile (of all observations during the same week), the average cumulative return over the two subsequent weeks is 21.382 basis points. Similarly, for the observations with SVI growth above the 80th and 90th percentile (of all observations during the same week), the average cumulative abnormal return over the two subsequent weeks is 21.701 and 17.117 basis points. The results are only significant for observations above these thresholds, supporting the results from the passive event study. Furthermore, we find significant evidence that the average cumulative abnormal return over the two subsequent weeks is significant for the 70th-80th, 80th-90th and 90th-100th decile groups. The highest cumulative abnormal return (22.151 basis points) is expected to be found within the 80th-90th decile group, when the event thresholds are continuously redefined.

Based on these findings, it would be possible to construct a simplified trading strategy that is triggered to invest as actively defined events occur. For the sake of illustration, we quantified possible gains of such a trading strategy by simulating a simple market neutral portfolio. The concept of a market neutral portfolio is to exploit market inefficiencies by being simultaneously long and short in matched equity portfolios (Jobman, 2002). In doing this, one can construct a portfolio that is close to beta neutral. Thus, the idea is to create a long portfolio each week that consists of the stocks whose observations constitute an event and simultaneously short the market portfolio. To maintain the beta neutrality of the market neutral portfolio, rebalancing of the weights is required each week since events are redefined on a continuous basis. Nevertheless, this example of a trading strategy is simplified since it strongly assumes that the market portfolio has the same risk profile as the long portfolio that is to be invested in each week. However, this might not be the case and the trading strategy should indeed be elaborated on.[36]

Within the frame of this simplified trading strategy, one can on average expect to earn the highest abnormal returns by investing in the 80th-90th decile group each week. This simulation suggests that the strategy may perform a return of approximately 11.5 percent per year, if (1) taking a long position in the 80th-90th decile group stocks each week, (2) taking a short position in the market portfolio and (3) holding these positions for two weeks.[37] At first glance, a yearly return of 11.5 percent might seem overestimated. However, it may seem more reasonable consider-

[36] A weighted portfolio consisting of OMX Nordic Small Cap Index, OMX Nordic Mid Cap Index, OMX Nordic Large Cap might be even better matched with the long portfolio. As presented in Graph 2 small firms are more likely to have attention shocks than large firms, which implies that the long portfolio will most likely have a larger proportion of small firms than the proportion small firms in the OMX Nordic Index.

[37] The annual return is calculated by multiplying the average two-week cumulative abnormal return earned by investing in the 80th-90th decile group stocks (22.151 basis points) by the number of weeks per year, since the findings of this thesis only allow the investment decisions to be made once a week. Reinvestments of earnings have not been taken into consideration when calculating the yearly return.

ing that this analysis is performed with no regards to transaction costs and taxes. Undeniably, a perfectly market neutral portfolio is difficult to construct since the beta values are derived from historical performance and do not necessarily represent the present and future risk profile of the assets. Moreover, Da et al. (2011) do unfortunately not provide any example of a trading strategy in their paper. Thus, the simulated trading strategy cannot be compared to any alike previous research for validation.

In sum, the results from the passive event study model find evidence that a higher value of growth in SVI can be seen as a predictor of higher abnormal returns. This holds especially for the second subsequent week, which is in consensus with and thus validates our findings from the Fama-MacBeth regression model method. More specifically, this method adds to our findings that the highest predictable power in SVI emerges when looking at the observations with the highest level of growth in SVI. The results from the active event study model shows that actively and continuously defined event thresholds can be used to predict in what particular stocks we can find cumulative abnormal returns over the two subsequent weeks. Within the frame of an illustrative and simplified trading strategy, we use these findings to see that one can potentially gain a yearly return of approximately 11.5 percent per year.

Both the Fama-MacBeth regression method and the event study find evidence in support of our hypothesis that abnormal search volume data can be used to predict positive abnormal values of stock measures. Hence, our results are robust to the selection of method. Additionally our results remain robust when including controls for market value of company, one week lagged abnormal return and second level abnormal SVI. Furthermore we found that our results are robust for different benchmark approaches to define expected return. However, they are sensitive to the definition of expected SVI, which emphasises the importance of accurately defining expected SVI.

7 Conclusion

The results from our empirical tests suggest that there exists a positive relation between stock returns and values of search volume. Verifying our hypothesis, we find evidence of a statistically and economically significant positively linear relation between abnormal stock returns and abnormal values of the Search Volume Index (SVI) provided by Google Inc. This presupposes that the search volume data is generated from using the firm name as search query. In line with this we find that such a positive relation especially holds for smaller firms, which also was hypothesised. Furthermore, we find that the predicted abnormal return is highest when observing the highest growth in search volume. The results are robust both to methodology and to the inclusion of the control variables market value of company and one week lagged abnormal return. We further seek to find a relation between search volume and liquidity and volatility. Whilst finding an evidence of a positive relation between abnormal search volume and abnormal liquidity, we only find an indication of a positive relation between abnormal search volume and abnormal volatility. Finally, we see that there are possible gains of studying the SVI, with up to 11.5 percent in yearly return before transaction costs and taxes.

The paper has sought to fill the gap in the research of relations between stock returns and search volume, i.e. a direct proxy for investor attention, by presenting a method that is applicable on smaller markets. The paper has also sought to contribute to the discussion on investor attention by applying the use of this direct measure of investor attention to a market that provide new sets of data in an attempt to reinforce the view that search volume is a feasible proxy for investor attention.

A potential caveat with this study is that SVIs could not be provided for a significant number of firms listed on the OMX Nordic exchanges. Hence, our findings are only valid for firms that have sufficient relative interest on Google to generate an index at Google Insights. Additionally, all returns in this study are before transaction costs and taxes. This implies that the findings of this study might not hold after considering transaction costs and taxes. We therefore suggest future research to apply the methodology of this thesis to regions where one can include all firms in the sample to see any there is any effect on the results. Moreover, we recommend future research to investigate if any institutional investors trade based on SVI data to see if real cases support the findings of this thesis.

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.

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

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.

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

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting Influenza Epidemics Using Search Engine Query Data. *Nature , 457* (7232), pp. 1012-1014.

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

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, Breath of Ownership, and Liquidity. *Review of Financial Studies , 17* (2), pp. 439-461.

Internet World Stats. (2012, 04 28). *Internet World Stats - Usage and Population Statistics.* Retrieved 05 02, 2012, from Internet Usage Statistics: www.internetworldstats.com

Hirshleifer, D., & Teoh, S. (2003). Limited Attention, Information Disclosure and Financial Reporting. *Journal of Accounting and Finance , 36* (1-3), pp. 337-386.

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.

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

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

Net Applications. (2012, 05 20). *Search Engine Market Share*. Retrieved 05 20, 2012, from Search Engine Market Share: http://marketshare.hitslink.com/

Mondria, J., Wu, T., & Zhang, Y. (2010). The Determinants of International Investment and Attention Allocation: Using Internet Search Query Data. *Journal of International Economics , 82* (1), pp. 85-95.

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.

Pew Research Center. (2012, 01 20). *Pew Internet*. Retrieved 03 14, 2012, from Pew Internet - A Project of the Pew Research Center: http://www.pewinternet.org/

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.

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

Appendix A

Table A1: Distribution of Search Volume Index (SVI) values per market capitalisation group

Search Volume Index (SVI)	Small Cap	Mid Cap	Large Cap	Total
0-10	22 959	8 430	4 533	35 922
	29.78%	18.15%	11.20%	21.90%
10-20	7 160	2 926	2 950	13 036
	9.29%	6.30%	7.29%	7.95%
20-30	10 354	4 659	3 169	18 182
	13.43%	10.03%	7.83%	11.09%
30-40	10 588	6 723	5 432	22 743
	13.73%	14.48%	13.42%	13.87%
40-50	9 940	7 334	6 665	23 939
	12.89%	15.79%	16.47%	14.60%
50-60	7 391	6 293	6 742	20 426
	9.59%	13.55%	16.66%	12.45%
60-70	4 743	5 221	5 260	15 224
	6.15%	11.24%	13.00%	9.28%
70-80	2 580	3 173	3 743	9 496
	3.35%	6.83%	9.25%	5.79%
80-90	1 055	1 272	1 608	3 935
	1.37%	2.74%	3.97%	2.40%
90-100	336	403	368	1 107
	0.44%	0.87%	0.91%	0.67%
Total	77 106	46 434	40 470	164 010

Table A1 shows the distribution of Search Volume Index (SVI) values per market capitalisation group (small cap, mid cap and large cap) for the sample over the sample period 2004-2011. A SVI value shows how much relative attention a specific search query had during one week relative to all other weeks in the period of interest and can take any value between 0 and 100, where 100 represent the highest attention observed per firm. The graph is based on all valid SVI observations between January 2004 and December 2011 for 386 firms listed on the Nordic OMX exchanges These are grouped in ten groups based on the SVI values. The graph shows how large fraction of the total number of observations within each market cap group that is in each SVI interval. Total number of observations with valid SVI values are 76 733 for Small Cap, 46 211 for Mid Cap and 40 275 for Large Cap.

Table A2: Definition of variables

Variable	Components	Definitions
Abnormal return (AR) (percent)	AR	Calculated by subtracting the expected weekly return at time t (defined as the market risk adjusted return) of a firm from the actual weekly return at time t.
	Stock return (r)	The development of an index of total return. Assuming that dividends are re-invested for the purchase of extra units of the equity (at the closing price applicable on the ex-dividend date), this total return index shows a theoretical growth in value of a share holding over a specific time period. The index is adjusted for capital events.
	Benchmark return (rm)	The country specific benchmarks used are NASDAQ OMX Country Index: Stockholm, Copenhagen, Helsinki and Reykjavik. The Nordic benchmark used is the NASDAQ OMX Nordic.
	Risk free return (rf)	For each country, we use 1-month Stockholm (STIBOR), Copenhagen (CIBOR), Helsinki (HELIBOR) and Reykjavik (REIBOR) Interbank Offered Rate, respectively.
Abnormal search volume (ASVI) (no unit)	ASVI	Defined as the log of SVI minus the log of the median/mean SVI of the prior eight weeks.
	Search Volume Index (SVI)	Aggregate weekly search volume data obtained from Google Insights, generated from firm name as search query.
Abnormal liquidity (AL) (SEK)	Bid-Ask AL (ALBA)	Defined as the observed bid-ask spread minus the expected bid-ask spread, which is determined by the mean of the bid-ask spread during the prior eight weeks.
	Trading volume AL (ALTV)	Defined as the observed trading volume minus the expected trading volume, which is determined by the mean of the trading volume during the prior eight weeks.
	Bid price	The bid price offered at close of market.
	Ask price	The ask price quoted at close of market.
	Bid-ask spread (bas)	Proxy for liquidity which values are calculated by subtracting bid price from ask price.
	Trading volume (tv)	Represents the total number of shares that are traded for a stock on each day multiplied by the stock price. This figure is expressed in thousands. This variable is adjusted for capital events.
Abnormal volatility (AV) (SEK)	AV	Defined as the observed volatility proxy minus the expected volatility proxy, which is determined by the mean of the volatility proxy during the prior eight weeks.
	Volatility proxy	The volatility proxy for each week is defined as the average of the week's five intra-day high-minus-low values.
	High price (ph)	The highest price observed each day.
	Low price (pl)	The lowest price observed each day.
	Price (p)	The official closing price.
Control variables	Market value of company (MVC) (SEK)	For companies with a single listed equity security, MVC is the share price multiplied by the number of ordinary shares in issue. Each time new tranches of stock are issued or whether there occurs any other capital change, the amount of issue is updated. For firms with more than one listed or unlisted equity security, MVC represents: Equity A(MV) + Equity B(MV) + Equity C(MV) and so forth. This is the consolidated market value of a company.
	Log of MVC (LMVC)	Defined as the log of MVC.
	LMVC * ASVI	Interaction variable between log of market value of company and abnormal SVI.
	$ASVI^2$	Squared abnormal SVI.
	AR(n-1)	One week lagged abnormal return.

Table A2 summarizes the variables used in the studies of this thesis. It also presents definitions of the components that complete the construction of these particular regression variables used in the framework section. First, the regression variables are defined and below their definition follow definitions for the components of the regression variables. The last variables explained are the control variables. These are defined on a stand alone basis.

Table A3: Descriptive statistics of the variable properties

		Total sample	Small Cap	Mid Cap	Large Cap
No. of companies	Total sample	386	181	110	95
	Sweden	161	59	55	47
	Denmark	119	80	19	20
	Finland	103	41	34	28
	Iceland	3	1	2	0
No. of observations	Total sample	164 010	77 106	46 434	40 470
	Sweden	68 586	25 134	23 430	20 022
	Denmark	50 268	34 080	7 668	8 520
	Finland	43 878	17 466	14 484	11 928
	Iceland	1 278	426	852	0
Search volume index (SVI)	Mean	35.378	28.810	38.759	44.013
(no unit)	Std. Dev.	24.862	23.964	24.561	23.427
	Min	0	0	0	0
	Median	37.0	28.0	40.0	46.0
	Max	100	100	100	100
Return index	Mean	0.223%	0.111%	0.294%	0.352%
(percent)	Std. Dev.	6.278%	7.214%	5.486%	5.132%
	Min	-75.000%	-75.000%	-64.193%	-41.128%
	Median	0%	0%	0%	0.314%
	Max	600.170%	600.170%	73.879%	61.795%
Market value of company	Mean	22 025	1 028	4 340	81 946
(SEK)	Std. Dev.	131 144	3 882	4 955	253 804
	Min	2	2	53	62
	Median	1 718	408	3 201	26 633
	Max	3 259 216	80 426	88 565	3 259 216
Abnormal SVI (Median approach)	Mean	34.744	31.335	35.487	40.479
(no unit)	Std. Dev.	23.899	23.273	23.977	23.869
	Min	0	0	0	0
	Median	36.5	32.5	37.0	42.5
	Max	99.0	97.5	99.0	99.0
Abnormal SVI (Mean approach)	Mean	35.387	28.809	38.790	44.014
(no unit)	Std. Dev.	23.346	22.073	23.129	22.285
	Min	0	0	0	0
	Median	36.875	28.625	40.875	46.375
	Max	96.625	96.625	94.625	95.875
Growth in SVI	Mean	2.235%	2.788%	2.196%	1.683%
(percent)	Std. Dev.	24.728%	27.014%	24.765%	21.975%
	Min	-	-	-	-
	Median	0%	0%	0%	0%
	Max	-	-	-	-
Bid-ask spread	Mean	4.464	7.806	1.477	1.678
(SEK)	Std. Dev.	36.964	52.449	7.543	15.693
	Min	-3.0	0	0	-3.0
	Median	0.250	0.500	0.250	0.220
	Max	2750	2750	275	1300
Trading volume	Mean	92 376.960	4 775.604	30 007.140	302 239.400
(SEK)	Std. Dev.	848 944.800	52 669.650	142 991.800	1 602 287.000
	Min	0	0	0	0.200
	Median	686.950	74.595	707.800	8 634.723
	Max	38 500 000	8 623 560	7 276 502	38 500 000

Table A3 continues on next page

Table A3 continued from previous page

		Total sample	Small Cap	Mid Cap	Large Cap
Volatility proxy	Mean	15.198	7.279	6.327	38.915
(SEK)	Std. Dev.	168.237	39.101	16.531	325.063
	Min	0	0	0	0.000
	Median	2.383	1.300	2.421	4.461
	Max	10 909.800	2 943.375	806.400	10 909.800

Table A3 summarizes descriptive statistics for the sample in total and for each of the capital sizes (small cap, mid cap and large cap) over the sample period 2004-2011. *No. of companies* describes the total number of firms per capital size and country. *No. of observations* describes the total number of observed (and validated) SVI values for the total sample as well as per capital size and country. *Search volume index (SVI)* describes the observed SVI values for the total sample and for each of the capital sizes. *Return index* describes the return index values for the total sample and for each of the capital sizes. *Market value of company* describes the consolidated market values for the total sample and for each of the capital sizes. *Abnormal search volume (Median approach)* describes the abnormal search volume values, defined using the median approach, for the total sample and for each of the capital sizes. *Abnormal search volume (Mean approach)* describes the abnormal search volume values, defined using the mean approach, for the total sample and for each of the capital sizes. *Growth in search volume* describes the observed values of growth in search volume between two time periods for the total sample and for each of the capital sizes. *Bid-ask spread* describes the bid-ask spread for the total sample and for each of the capital sizes. *Trading volume* describes the trading volume, expressed in thousands, for the total sample and for each of the capital sizes. *Volatility proxy* describes the volatility proxy for the total sample and for each of the capital sizes.

Appendix B

Table B1: Fama-MacBeth regression using subsequent weeks' abnormal return as dependent variable

(1) Defining ASVI using the Median Approach

	(a) Defining AR using the Country Specific Approach					(b) Defining AR using the Nordic Approach				
Week *n*	ASVI	LMVC	ASVI * LMVC	Cons.	*R2 (Obs)*	ASVI	LMVC	ASVI * LMVC	Cons.	*R2 (Obs)*
1	21.628**	3.642	-21.239**	1.045	*0.033*	19.299*	3.701	-18.531**	1.849	*0.033*
	(9.912)	(4.566)	(9.022)	(0.119)	*(104 892)*	(9.952)	(4.735)	(8.961)	(1.143)	*(104 892)*
2	17.919*	3.695	-18.003**	1.123	*0.034*	15.003	3.754	-14.903*	1.899	*0.034*
	(9.627)	(4.630)	(8.551)	(0.123)	*(104 892)*	(9.457)	(4.827)	(8.411)	(1.159)	*(104 892)*
3	11.889	3.731	-11.963	1.061	*0.034*	12.076	3.789	-11.978	1.858	*0.033*
	(8.962)	(4.653)	(8.098)	(0.124)	*(104 892)*	(8.966)	(4.852)	(8.019)	(1.183)	*(104 892)*
4	12.932	3.693	-11.676	1.126	*0.033*	13.254	3.755	-12.772	1.815	*0.033*
	(10.575)	(4.695)	(8.668)	(1.323)	*(104 892)*	(9.559)	(4.902)	(8.579)	(1.242)	*(104 892)*
5	11.093	3.688	-11.861	1.467	*0.034*	8.326	3.748	-8.694	1.812	*0.034*
	(9.853)	(4.636)	(8.934)	(1.344)	*(104 892)*	(10.057)	(4.850)	(9.012)	(1.269)	*(104 892)*
6	11.781	3.674	-11.672	0.693	*0.034*	9.828	3.749	-9.436	1.446	*0.034*
	(9.520)	(4.713)	(8.512)	(1.269)	*(104 892)*	(9.847)	(4.932)	(8.798)	(1.248)	*(104 892)*
7	9.037	3.672	-9.749	0.478	*0.034*	6.748	3.745	-7.578	1.308	*0.034*
	(10.741)	(4.876)	(9.753)	(1.260)	*(104 892)*	(10.955)	(5.082)	(9.918)	(1.241)	*(104 892)*
8	13.924	3.658	-13.316	0.450	*0.034*	10.841	3.738	-10.463	1.283	*0.035*
	(9.973)	(4.849)	(8.955)	(1.338)	*(104 536)*	(10.129)	(5.071)	(9.138)	(1.320)	*(104 536)*
9	12.407	3.706	-11.528	0.033	*0.034*	11.500	3.789	-10.817	0.890	*0.035*
	(10.405)	(4.895)	(9.223)	(1.230)	*(104 177)*	(10.649)	(5.100)	(9.450)	(1.283)	*(104 177)*
10	11.984	3.688	-12.417	-0.531	*0.034*	11.687	3.767	-10.320	0.352	*0.035*
	(9.828)	(4.957)	(8.816)	(1.344)	*(103 821)*	(10.017)	(5.170)	(8.975)	(1.314)	*(103 821)*

(2) Defining ASVI using the Mean Approach

	(a) Defining AR using the Country Specific Approach					(b) Defining AR using the Nordic Approach				
Week *n*	ASVI	LMVC	ASVI * LMVC	Cons.	*R2 (Obs)*	ASVI	LMVC	ASVI * LMVC	Cons.	*R2 (Obs)*
1	1.219	3.583	-1.178	1.347	*0.029*	1.217	3.641	-0.484	1.396	*0.029*
	(2.335)	(4.600)	(1.454)	(1.200)	*(104 516)*	(2.267)	(4.763)	(1.451)	(1.131)	*(104 516)*
2	-0.538	3.647	-1.006	1.505	*0.029*	-0.952	3.708	-0.455	1.458	*0.029*
	(2.315)	(4.664)	(1.572)	(1.234)	*(104 516)*	(2.288)	(4.858)	(1.536)	(1.170)	*(104 516)*
3	-1.102	3.699	-0.972	1.203	*0.030*	-1.464	3.754	-0.410	1.989	*0.030*
	(2.401)	(4.700)	(1.644)	(1.286)	*(104 516)*	(2.401)	(4.887)	(1.621)	(1.221)	*(104 516)*
4	1.471	3.649	-1.486	1.568	*0.030*	2.662	3.709	-0.787	1.276	*0.030*
	(1.948)	(4.753)	(1.562)	(1.289)	*(104 516)*	(1.945)	(4.956)	(1.558)	(1.199)	*(104 516)*
5	0.920	3.643	-1.071	1.578	*0.030*	-0.125	3.699	-0.267	1.246	*0.030*
	(2.604)	(4.766)	(1.681)	(1.342)	*(104 516)*	(2.544)	(4.981)	(1.636)	(1.248)	*(104 516)*
6	0.113	3.655	-0.756	1.139	*0.030*	-0.618	3.723	-0.089	1.958	*0.031*
	(1.862)	(4.837)	(1.739)	(1.297)	*(104 516)*	(1.903)	(5.055)	(1.726)	(1.271)	*(104 516)*
7	1.486	3.648	-1.110	0.586	*0.031*	0.580	3.721	-0.859	1.479	*0.032*
	(2.347)	(4.880)	(1.691)	(1.286)	*(104 516)*	(2.391)	(5.084)	(1.707)	(1.257)	*(104 516)*
8	2.428	3.651	-1.093	0.557	*0.031*	1.564	3.725	-0.826	1.501	*0.031*
	(1.946)	(4.945)	(1.753)	(1.367)	*(104 160)*	(2.020)	(5.173)	(1.744)	(1.341)	*(104 160)*
9	2.110	3.656	-0.034	0.105	*0.031*	3.286	3.733	0.192	1.069	*0.032*
	(2.350)	(4.967)	(1.627)	(1.390)	*(103 803)*	(2.481)	(5.179)	(1.653)	(1.370)	*(103 803)*
10	2.883	3.656	-0.159	-0.250	*0.030*	3.190	3.732	0.178	0.681	*0.031*
	(2.113)	(4.953)	(1.660)	(1.277)	*(103 447)*	(2.216)	(5.175)	(1.669)	(1.239)	*(103 447)*

Table B1 reports results of Fama-MacBeth (1973) cross-sectional regressions. The sample consists of 386 firms listed on NASDAQ OMX Nordic exchanges. The sample period is from January 2004 to December 2011. The dependent variable is week(t+n) abnormal return (where n goes from 1 to 10), denoted in basis points. For regressions in panels 1a and 1b, the abnormal SVI value is determined by using an expected SVI value defined as the median of the eight prior weeks. For regressions in panels 2a and 2b, the abnormal SVI value is determined by using an expected SVI value defined as the mean of the eight prior weeks. For regressions in panels 1a and 2a, the benchmark used as market return is country specific. The indices used as benchmarks are OMX Copenhagen, OMX Helsinki, OMX Stockholm and OMX Reykjavik. For regressions in panels 1b and 2b, the benchmark used as market return is the OMX Nordic index. The independent variables are abnormal SVI (ASVI), log of market value of company (LMVC) and an interaction term between ASVI and LMVC. The independent variables are defined in Table A2 in appendix A. All variables are cross-sectionally demeaned to have a regression intercept close to zero. The independent variables are standardized. This means that one can interpret the regression coefficients as the impact of a one-standard-deviation change in the independent variables. 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 percent level. ** represents significance at the 5 percent level. *** represents significance at the 1 percent level.

Table B2: Fama-MacBeth regression using subsequent weeks' abnormal return as dependent variable, controlling for prior week's abnormal return

(1) Defining ASVI using the Median Approach

Week	(a) Defining AR using the Country Specific Approach						(b) Defining AR using the Nordic Market Approach					
n	ASVI	LMVC	ASVI * LMVC	AR (*n*-1)	Cons.	*R2 (Obs)*	ASVI	LMVC	ASVI * LMVC	AR (*n*-1)	Cons.	*R2 (Obs)*
2	18.189*	3.821	-18.336**	-5.341	1.174	*0.051*	14.934	3.884	-15.093*	-5.084	1.933	*0.051*
	(9.525)	(4.936)	(8.425)	(8.899)	(1.249)	*(104 892)*	(9.324)	(5.112)	(8.268)	(9.092)	(1.197)	*(104 892)*
3	15.342	3.868	-15.456	-5.361	1.154	*0.051*	12.333	3.930	-12.394	-5.115	1.928	*0.051*
	(17.862)	(4.974)	(17.950)	(8.931)	(1.256)	*(104 892)*	(8.833)	(5.157)	(7.860)	(9.068)	(1.214)	*(104 892)*
4	12.983	3.822	-13.533	-5.281	1.222	*0.051*	14.178	3.888	-13.650	-5.047	2.033	*0.051*
	(9.615)	(5.000)	(12.665)	(9.035)	(1.328)	*(104 892)*	(9.544)	(5.192)	(8.548)	(9.201)	(1.270)	*(104 892)*
5	11.657	3.827	-12.140	-5.286	1.536	*0.052*	8.867	3.891	-9.092	-5.054	1.885	*0.051*
	(9.793)	(4.935)	(8.854)	(9.379)	(1.361)	*(104 892)*	(9.969)	(5.132)	(8.908)	(9.582)	(1.312)	*(104 892)*
6	11.744	3.827	-11.584	-5.208	0.708	*0.052*	9.673	3.909	-9.369	-5.009	1.527	*0.051*
	(9.452)	(5.010)	(8.446)	(9.271)	(1.300)	*(104 892)*	(9.736)	(5.219)	(8.703)	(9.512)	(1.290)	*(104 892)*
7	9.409	3.816	-10.191	-5.129	0.641	*0.051*	6.914	3.890	-7.769	-4.889	1.491	*0.052*
	(10.629)	(5.161)	(9.682)	(9.121)	(1.315)	*(104 892)*	(10.842)	(5.359)	(9.845)	(9.386)	(1.298)	*(104 892)*
8	13.707	3.809	-13.262	-5.336	0.601	*0.052*	10.504	3.890	-10.283	-5.118	1.470	*0.053*
	(9.926)	(5.145)	(14.946)	(9.263)	(1.385)	*(104 536)*	(10.032)	(5.357)	(9.085)	(9.479)	(1.363)	*(104 536)*
9	11.661	3.845	-11.001	-5.285	0.153	*0.052*	15.614	3.929	-15.161	-5.067	1.022	*0.053*
	(10.285)	(5.203)	(9.126)	(9.308)	(1.341)	*(104 177)*	(10.510)	(5.406)	(9.325)	(9.522)	(1.320)	*(104 177)*
10	10.538	3.827	-10.181	-5.387	-0.440	*0.052*	14.317	3.907	-14.100	-5.168	0.445	*0.053*
	(9.670)	(5.264)	(8.692)	(9.309)	(1.399)	*(103 821)*	(9.821)	(5.468)	(8.813)	(9.498)	(1.368)	*(103821)*

(2) Defining ASVI using the Mean Approach

Week	(a) Defining AR using the Country Specific Approach						(b) Defining AR using the Nordic Market Approach					
n	ASVI	LMVC	ASVI * LMVC	AR (*n*-1)	Cons.	*R2 (Obs)*	ASVI	LMVC	ASVI * LMVC	AR (*n*-1)	Cons.	*R2 (Obs)*
2	-0.538	3.783	-0.998	-5.505	1.577	*0.048*	-1.075	3.843	-0.577	-5.218	1.407	*0.048*
	(2.348)	(4.993)	(1.666)	(9.155)	(1.259)	*(104 516)*	(2.285)	(5.162)	(1.637)	(9.324)	(1.210)	*(104 516)*
3	-0.987	3.854	-0.985	-5.579	1.261	*0.048*	-1.421	3.909	-0.514	-5.279	1.136	*0.048*
	(2.468)	(5.052)	(1.729)	(9.199)	(1.313)	*(104 516)*	(2.464)	(5.220)	(1.700)	(9.314)	(1.260)	*(104 516)*
4	3.874	3.791	-1.430	-5.399	1.623	*0.048*	3.116	3.851	-0.863	-5.117	1.436	*0.049*
	(1.949)	(5.089)	(1.634)	(9.426)	(1.299)	*(104 516)*	(1.944)	(5.271)	(1.624)	(9.584)	(1.227)	*(104 516)*
5	1.524	3.789	-0.953	-5.426	1.623	*0.049*	0.511	3.849	-0.269	-5.165	1.381	*0.049*
	(2.622)	(5.087)	(1.746)	(9.422)	(1.357)	*(104 516)*	(2.564)	(5.286)	(1.715)	(9.621)	(1.287)	*(104 516)*
6	0.295	3.813	-0.676	-5.417	1.115	*0.049*	-0.453	3.887	-0.121	-5.208	1.981	*0.049*
	(1.888)	(5.144)	(1.804)	(9.366)	(1.333)	*(104 516)*	(1.931)	(5.350)	(1.783)	(9.575)	(1.315)	*(104 516)*
7	1.616	3.802	-1.196	-5.424	0.789	*0.049*	0.899	3.877	-0.924	-5.174	1.703	*0.050*
	(2.409)	(5.194)	(1.754)	(9.162)	(1.343)	*(104 516)*	(2.460)	(5.396)	(1.770)	(9.416)	(1.313)	*(104 516)*
8	2.606	3.803	-1.202	-5.485	0.673	*0.049*	1.837	3.878	-0.936	-5.241	1.648	*0.050*
	(1.878)	(5.262)	(1.827)	(9.403)	(1.438)	*(104 160)*	(1.978)	(5.481)	(1.813)	(9.623)	(1.407)	*(104 160)*
9	2.178	3.809	-0.071	-5.446	0.178	*0.049*	3.357	3.886	0.184	-5.179	1.143	*0.051*
	(2.347)	(5.300)	(1.697)	(9.411)	(1.457)	*(103 803)*	(2.480)	(5.506)	(1.717)	(9.622)	(1.430)	*(103 803)*
10	2.059	3.798	-0.190	-5.435	-0.142	*0.049*	3.296	3.873	0.164	-5.166	0.797	*0.050*
	(2.182)	(5.273)	(1.718)	(9.637)	(1.300)	*(103 447)*	(2.283)	(5.486)	(1.722)	(9.857)	(1.253)	*(103 447)*

Table B2 reports results of Fama-MacBeth (1973) cross-sectional regressions. These regressions are controlled for week(t+n-1) abnormal return, i.e. the abnormal return the week before the future week of interest (where n goes from 2 to 10). The sample consists of 386 firms listed on NASDAQ OMX Nordic exchanges. The sample period is from January 2004 to December 2011. The dependent variable is abnormal return for week(t+n) denoted in basis points. For regressions in panels 1a and 1b, the abnormal SVI value is determined by using an expected SVI value defined as the median of the eight prior weeks. For regressions in panels 2a and 2b, the abnormal SVI value is determined by using an expected SVI value defined as the mean of the eight prior weeks. For regressions in panels 1a and 2a, the benchmark used as market return is country specific. The indices used as benchmarks are OMX Copenhagen, OMX Helsinki, OMX Stockholm and OMX Reykjavik. For regressions in panels 1b and 2b, the benchmark used as market return is the OMX Nordic index. The independent variables are abnormal SVI (ASVI), log of market value of company (LMVC), an interaction term between ASVI and LMVC and one week lagged abnormal return. The independent variables are defined in Table A2 in appendix A. All variables are cross-sectionally demeaned to have a regression intercept close to zero. The independent variables are standardized. This means that one can interpret the regression coefficients as the impact of a one-standard-deviation change in the independent variables. 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 percent level. ** represents significance at the 5 percent level. *** represents significance at the 1 percent level.

Table B3: Fama-MacBeth regression using subsequent weeks' abnormal return as dependent variable, controlling for squared ASVI

(1) Defining ASVI using the Median Approach

	(a) Defining AR using the Country Specific Approach						(b) Defining AR using the Nordic Market Approach					
Week *n*	ASVI	LMVC	ASVI * LMVC	$ASVI^2$	Cons.	*R2 (Obs)*	ASVI	LMVC	ASVI * LMVC	$ASVI^2$	Cons.	*R2 (Obs)*
1	21.947**	3.645	-21.052**	-0.547	1.488	*0.040*	20.194**	3.699	-18.835**	-2.336	1.737	*0.040*
	(9.386)	(4.552)	(8.585)	(1.479)	(1.692)	*(104 892)*	(9.581)	(4.727)	(8.656)	(1.481)	(1.697)	*(104 892)*
2	18.233**	3.698	-17.752**	-0.062	1.361	*0.040*	15.815*	3.752	-15.033*	-1.891	1.592	*0.041*
	(9.106)	(4.630)	(8.098)	(1.374)	(1.637)	*(104 892)*	(9.050)	(4.830)	(8.073)	(1.364)	(1.613)	*(104 892)*
3	12.034	3.734	-12.740	-1.037	1.702	*0.040*	13.219	3.787	-12.740	-1.687	1.918	*0.041*
	(9.044)	(4.653)	(8.258)	(1.439)	(1.781)	*(104 892)*	(9.149)	(4.859)	(8.254)	(1.461)	(1.752)	*(104 892)*
4	13.154	3.693	-10.876	-1.164	1.765	*0.040*	15.596	3.747	-12.533	-1.890	1.937	*0.040*
	(9.503)	(4.689)	(8.618)	(1.418)	(1.746)	*(104 892)*	(9.612)	(4.897)	(8.614)	(1.437)	(1.701)	*(104 892)*
5	11.751	3.688	-12.367	0.136	1.691	*0.041*	9.184	3.742	-9.390	-1.426	1.724	*0.041*
	(9.450)	(4.637)	(8.557)	(1.681)	(1.780)	*(104 892)*	(9.703)	(4.855)	(8.687)	(1.688)	(1.740)	*(104 892)*
6	11.602	3.671	-10.933	0.332	0.885	*0.041*	9.653	3.739	-8.763	-1.221	1.018	*0.041*
	(9.251)	(4.721)	(8.307)	(1.641)	(1.631)	*(104 892)*	(9.643)	(4.941)	(8.624)	(1.596)	(1.594)	*(104 892)*
7	8.369	3.664	-9.121	1.062	-0.004	*0.041*	6.337	3.732	-7.143	-0.606	2.280	*0.042*
	(10.616)	(4.868)	(9.750)	(1.609)	(1.634)	*(104 892)*	(10.807)	(5.081)	(9.898)	(1.601)	(1.599)	*(104 892)*
8	13.187	3.658	-12.710	1.445	-0.241	*0.041*	10.125	3.733	-9.879	-0.041	1.900	*0.042*
	(9.727)	(4.837)	(8.784)	(1.605)	(1.673)	*(104 536)*	(9.996)	(5.059)	(9.038)	(1.585)	(1.637)	*(104 536)*
9	13.225	3.701	-10.908	1.944	-1.083	*0.041*	13.373	3.779	-10.402	0.311	1.157	*0.042*
	(9.906)	(4.888)	(8.869)	(1.605)	(1.695)	*(104 177)*	(10.231)	(5.094)	(9.156)	(1.565)	(1.669)	*(104 177)*
10	12.753	3.686	-10.581	1.268	-0.973	*0.041*	15.441	3.759	-10.437	-0.461	1.297	*0.042*
	(9.422)	(4.956)	(8.504)	(1.639)	(1.764)	*(103 821)*	(9.715)	(5.168)	(8.741)	(1.494)	(1.704)	*(103 821)*

(2) Defining ASVI using the Mean Approach

	(a) Defining AR using the Country Specific Approach						(b) Defining AR using the Nordic Market Approach					
Week *n*	ASVI	LMVC	ASVI * LMVC	$ASVI^2$	Cons.	*R2 (Obs)*	ASVI	LMVC	ASVI * LMVC	$ASVI^2$	Cons.	*R2 (Obs)*
1	3.433	3.561	-1.731	-0.890	2.463	*0.035*	3.929*	3.616	-1.107	-1.206	3.613	*0.036*
	(2.420)	(4.602)	(1.505)	(1.153)	(1.642)	*(104 516)*	(2.377)	(4.763)	(1.467)	(1.155)	(1.540)	*(104 516)*
2	-0.452	3.636	-0.981	-1.091	1.946	*0.036*	-0.637	3.694	-0.535	-1.250	3.235	*0.036*
	(2.331)	(4.683)	(1.667)	(1.012)	(1.368)	*(104 516)*	(2.377)	(4.863)	(1.637)	(1.028)	(1.326)	*(104 516)*
3	0.893	3.699	-1.487	-0.718	1.998	*0.036*	0.586	3.757	-0.952	-0.602	1.749	*0.036*
	(2.531)	(4.712)	(1.651)	(0.983)	(1.519)	*(104 516)*	(2.626)	(4.908)	(1.596)	(0.992)	(1.464)	*(104 516)*
4	4.720	3.621	-1.900	-0.652	2.103	*0.036*	1.309	3.674	-1.357	-0.863	1.125	*0.036*
	(2.259)	(4.797)	(1.602)	(1.147)	(1.480)	*(104 516)*	(2.332)	(4.993)	(1.607)	(1.141)	(1.433)	*(104 516)*
5	1.827	3.628	-1.405	-0.627	1.818	*0.036*	1.262	3.672	-0.777	-1.161	1.994	*0.036*
	(2.842)	(4.755)	(1.625)	(0.956)	(1.515)	*(104 516)*	(2.756)	(4.961)	(1.590)	(0.955)	(1.455)	*(104 516)*
6	1.345	3.606	-0.967	-2.835	1.719	*0.036*	1.514	3.663	-0.493	-1.160	2.075	*0.036*
	(2.140)	(4.867)	(1.716)	(1.090)	(1.593)	*(104 516)*	(2.244)	(5.074)	(1.711)	(1.060)	(1.514)	*(104 516)*
7	2.527	3.641	-1.035	-0.830	0.933	*0.037*	2.338	3.707	-0.977	-1.169	2.209	*0.037*
	(2.553)	(4.901)	(1.688)	(1.291)	(1.656)	*(104 516)*	(2.620)	(5.100)	(1.672)	(1.290)	(1.613)	*(104 516)*
8	1.144	3.647	-0.733	-0.539	0.359	*0.037*	0.697	3.724	-0.635	-0.802	1.519	*0.037*
	(1.915)	(4.942)	(1.686)	(1.140)	(1.700)	*(104 160)*	(2.020)	(5.178)	(1.664)	(1.111)	(1.629)	*(104 160)*
9	4.278	3.647	-0.274	-0.697	0.365	*0.038*	3.981	3.719	-0.206	-0.994	1.590	*0.039*
	(2.693)	(4.986)	(1.633)	(1.092)	(1.707)	*(103 803)*	(2.813)	(5.189)	(1.613)	(1.121)	(1.629)	*(103 803)*
10	3.081	3.662	0.109	0.378	-0.853	*0.036*	3.039	3.727	0.273	-0.103	0.585	*0.037*
	(2.575)	(4.999)	(1.672)	(0.969)	(1.595)	*(103 447)*	(2.550)	(5.216)	(1.653)	(0.958)	(1.475)	*(103 447)*

Table B3 reports results of Fama-MacBeth (1973) cross-sectional regressions. These regressions are controlled for squared abnormal SVI. The sample consists of 386 firms listed on NASDAQ OMX Nordic exchanges. The sample period is from January 2004 to December 2011. The dependent variable is abnormal return for week(t+n) (where n goes from 2 to 10), denoted in basis points. For regressions in panels 1a and 1b, the abnormal SVI value is determined by using an expected SVI value defined as the median of the eight prior weeks. For regressions in panels 2a and 2b, the abnormal SVI value is determined by using an expected SVI value defined as the mean of the eight prior weeks. For regressions in panels 1a and 2a, the benchmark used as market return is country specific. The indices used as benchmarks are OMX Copenhagen, OMX Helsinki, OMX Stockholm and OMX Reykjavik. For regressions in panels 1b and 2b, the benchmark used as market return is the OMX Nordic index. The independent variables are abnormal SVI (ASVI), log of market value of company (LMVC), an interaction term between ASVI and LMVC and squared abnormal SVI. The independent variables are defined in Table A2 in appendix A. All variables are cross-sectionally demeaned to have a regression intercept close to zero. The independent variables are standardized. This means that one can interpret the regression coefficients as the impact of a one-standard-deviation change in the independent variables. 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 percent level. ** represents significance at the 5 percent level. *** represents significance at the 1 percent level.

Table B4: Fama-MacBeth regression with three stock measures as dependent variables

(1) Defining ASVI using the Median Approach

Dependent Variable		Independent Variables				
Measure	Week	ASVI	LMVC	ASVI * LMVC	Cons.	*R2 (Obs)*
Abnormal Volume (ALTV)	1	263.898	-91.094	-254.914	28.171	*0.055*
		(1 365.568)	(682.216)	(1 471.620)	(86.835)	*(94 008)*
	2	-254.515	-25.825	262.421	7.530	*0.055*
		(1407.852)	(751.882)	(1 521.174)	(92.950)	*(94 015)*
Abnormal Bid-Ask Spread (ALBA)	1	-0.183	0.058	0.138	-0.067	*0.022*
		(0.285)	(0.116)	(0.249)	(0.055)	*(92 223)*
	2	-0.067	0.084	0.059	-0.089	*0.023*
		(0.295)	(0.121)	(0.256)	(0.059)	*(92 217)*
Abnormal Volatility (AV)	1	0.391	-0.019	-0.425	-0.081	*0.025*
		(0.993)	(0.846)	(1.103)	(0.076)	*(104 159)*
	2	0.935	-0.071	-1.059	-0.163	*0.028*
		(1.502)	(1.248)	(1.723)	(0.148)	*(104 159)*

(2) Defining ASVI using the Mean Approach

Dependent Variable		Independent Variables				
Measure	Week	ASVI	LMVC	ASVI * LMVC	Cons.	*R2 (Obs)*
Abnormal Volume (ALTV)	1	2 341.448***	-80.087	-2 309.12**	5.996	*0.045*
		(868.523)	(614.618)	(966.248)	(83.337)	*(105 960)*
	2	1 709.350**	-18.289	-1 767.498*	-28.057	*0.045*
		(841.072)	(669.879)	(924.135)	(90.043)	*(105 967)*
Abnormal Bid-Ask Spread (ALBA)	1	-0.399	0.016	0.281	-0.074	*0.020*
		(0.392)	(0.100)	(0.333)	(0.050)	*(104 256)*
	2	-0.153	0.039	0.105	-0.065	*0.020*
		(0.375)	(0.102)	(0.317)	(0.051)	*(104 262)*
Abnormal Volatility (AV)	1	0.823	-0.099	-1.023	-0.034	*0.024*
		(1.414)	(0.784)	(1.546)	(0.067)	*(117 885)*
	2	2.841	-0.223	-3.274	-0.101	*0.027*
		(2.208)	(1.424)	(2.586)	(0.131)	*(117 885)*

Table B4 reports results from twelve Fama-MacBeth (1973) cross-sectional regressions. The sample consists of 386 firms listed on NASDAQ OMX Nordic exchanges. The sample period is from January 2004 to December 2011. The dependent variable is either abnormal trading volume, abnormal bid-ask spread or abnormal volatility, for week 1 and week 2 forward in time. Trading volume is denoted in thousands of units (SEK). Bid-ask spread and volatility are denoted in single units (SEK). For the six top regressions the abnormal SVI value is determined by using an expected SVI value defined as the median of the eight prior weeks. For the six bottom regressions the abnormal SVI value is determined by using an expected SVI value defined as the mean of the eight prior weeks. The independent variables are abnormal SVI (ASVI), log of market value of company (LMVC) and an interaction term between ASVI and LMVC. The independent variables are defined in Table A2 in appendix A. All variables are cross-sectionally demeaned to have a regression intercept close to zero. The independent variables are standardized. This means that one can interpret the regression coefficients as the impact of a one-standard-deviation change. 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 percent level. ** represents significance at the 5 percent level. *** represents significance at the 1 percent level.

Appendix C

Table C1.1: Passive event study based on historical weekly SVI growth, for all observations above percentiles

		(i) Defining percentiles based on relative growth in SVI over all observations 2004-2011									
		(a) Defining AR using the Country Specific Approach					(b) Defining AR using the Nordic Approach				
Percentile	Week *n*	ASVI	LMVC	ASVI * LMVC	Cons.	*Rsqr (Obs)*	ASVI	LMVC	ASVI * LMVC	Cons.	*Rsqr (Obs)*
i > p50	1	-8.407	3.121	5.113	4.050	*0.053*	-5.911	3.069	3.555	4.395	*0.053*
		(13.989)	(3.966)	(12.516)	(3.114)	*(36 079)*	(13.782)	(3.976)	(12.469)	(3.085)	*(36 079)*
	2	9.209	3.787	-14.361	-2.491	*0.055*	13.733*	3.747	-11.816*	0.818	*0.055*
		(12.817)	(3.807)	(11.498)	(2.869)	*(36 079)*	(7.641)	(3.835)	(6.082)	(2.801)	*(36 079)*
	3	7.894	3.614	-6.376	0.184	*0.056*	13.435	3.559	-10.241	3.021	*0.056*
		(13.258)	(3.499)	(11.778)	(2.937)	*(36 079)*	(13.439)	(3.537)	(12.041)	(2.952)	*(36 079)*
i > p60	1	-4.744	3.238	1.126	4.749	*0.059*	-4.373	3.262	1.805	3.032	*0.059*
		(15.24)	(4.131)	(13.638)	(3.363)	*(31 846)*	(14.959)	(4.076)	(13.488)	(3.322)	*(31 846)*
	2	13.056*	3.778	-18.721	-3.310	*0.061*	12.479*	3.780	-10.178	-0.0306	*0.061*
		(6.816)	(3.965)	(13.706)	(3.068)	*(31 846)*	(7.102)	(3.971)	(13.810)	(3.005)	*(31 846)*
	3	10.498	3.523	-6.749	-0.649	*0.060*	16.129	3.490	-11.039	1.983	*0.061*
		(14.282)	(3.652)	(12.789)	(3.061)	*(31 846)*	(14.410)	(3.664)	(12.986)	(3.081)	*(31 846)*
i > p70	1	-0.397	3.325	-1.513	2.196	*0.072*	-1.282	3.364	0.091	4.983	*0.072*
		(17.484)	(4.357)	(16.178)	(3.706)	*(23 782)*	(17.540)	(4.319)	(16.398)	(3.636)	*(23 782)*
	2	16.562	3.758	-20.714	-5.577	*0.078*	15.812	3.818	-19.318	-2.634	*0.078*
		(18.157)	(4.677)	(17.179)	(3.512)	*(23 782)*	(18.295)	(4.666)	(17.298)	(3.523)	*(23 782)*
	3	6.957	3.957	-4.990	-1.294	*0.079*	11.435	3.967	-8.004	0.830	*0.080*
		(17.777)	(4.152)	(16.818)	(3.586)	*(23 782)*	(17.785)	(4.158)	(16.893)	(3.614)	*(23 782)*
i > p80	1	8.165*	4.461	-5.702	0.461	*0.106*	6.273	3.434	-1.840	3.220	*0.106*
		(4.687)	(5.652)	(22.948)	(4.588)	*(15 713)*	(23.938)	(5.637)	(22.965)	(4.552)	*(15 713)*
	2	23.169	4.572	-19.902	-2.994	*0.115*	23.015	3.732	-18.962	-4.362	*0.115*
		(24.862)	(6.51)	(24.243)	(4.514)	*(15 713)*	(25.327)	(6.487)	(24.759)	(4.435)	*(15 713)*
	3	-18.591	4.683	24.482	-1.676	*0.112*	-12.561	3.903	20.438	0.315	*0.113*
		(23.372)	(5.467)	(22.91)	(4.525)	*(15 713)*	(22.995)	(5.517)	(22.582)	(4.601)	*(15 713)*
i > p90	1	78.266	8.247	-15.362*	-1.367	*0.225*	-38.964	3.181	52.899	0.031	*0.229*
		(51.786)	(9.504)	(8.476)	(8.055)	*(7 803)*	(47.699)	(9.669)	(49.694)	(8.151)	*(7 803)*
	2	24.051*	9.019	-73.114	-7.961	*0.228*	17.760*	5.348	-71.522	-5.537	*0.230*
		(13.781)	(10.355)	(55.877)	(8.360)	*(7 803)*	(9.428)	(10.291)	(55.050)	(8.449)	*(7 803)*
	3	1.183	9.595	4.092	-1.258	*0.232*	15.323	4.892	-7.032	-11.449	*0.235*
		(55.559)	(9.452)	(58.539)	(7.151)	*(7 803)*	(56.161)	(9.626)	(59.092)	(7.209)	*(7 803)*

Table C1.1 reports results from a passive event study that is based on all values of weekly growth in SVI from the period of January 2004 to December 2011. The sample consists of 386 firms listed on NASDAQ OMX Nordic exchanges. The results are obtained by running Fama-MacBeth cross-sectional regressions in each interval group. Each percentile interval collects values of historical growth in all observations, *i*, of SVI from one week to the next. The intervall i > p50 coversall observations between 2004-2011 with SVI growth above median of all observatios 2004-2011, while the group i > p60 only covers the top 40% of all observations, etc. The dependent variable is weekly abnormal return denoted in basis points for week 1, week 2 and week 3 forward in time. In the left part of the table, indices used as benchmarks are OMX Copenhagen, OMX Helsinki, OMX Stockholm and OMX Reykjavik. In the right part of the table, the benchmark used as market return is the OMX Nordic index. The independent variables are abnormal SVI (ASVI), log of market value of company (LMVC) and an interaction term between ASVI and LMVC. The independent variables are defined in Table A2 in appendix A The abnormal SVI value is defined using the median approach. Every variable is cross-sectionally demeaned to have a regression intercept close to zero. The independent variables are standardized. This means that one can interpret the regression coefficients as the impact of a one-standard-deviation change. 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 percent level. ** represents significance at the 5 percent level. *** represents significance at the 1 percent level.

Table C1.2: Passive event study based on historical weekly SVI growth, for all observations between percentiles

		(ii) Defining percentile intervals based on relative growth in SVI over all observations 2004-2011									
		(a) Defining AR using the Country Specific Approach					(b) Defining AR using the Nordic Approach				
Percentile Interval	Week *n*	ASVI	LMVC	ASVI * LMVC	Cons.	*Rsqr (Obs)*	ASVI	LMVC	ASVI * LMVC	Cons.	*Rsqr (Obs)*
p50 < i < p60	1	8.084	2.4429	-7.830	9.216	*0.156*	9.294	2.2149	-8.778	2.303	*0.160*
		(6.089)	(6.618)	(5.623)	(6.130)	*(11 751)*	(61.167)	(6.769)	(18.354)	(6.049)	*(11 751)*
	2	-7.154	4.3169	8.854	-0.529	*0.158*	15.723	4.1359	-11.702	3.147	*0.158*
		(33.288)	(6.588)	(31.859)	(5.730)	*(11 751)*	(33.258)	(6.554)	(31.894)	(5.617)	*(11 751)*
	3	1.268	3.4878	-0.455	4.704	*0.162*	9.014	3.2774	-6.049	8.100	*0.162*
		(29.428)	(6.287)	(27.883)	(5.854)	*(11 751)*	(29.726)	(6.396)	(28.291)	(5.815)	*(11 751)*
p60 < i < p70	1	-91.943	4.1858	75.837	11.014	*0.225*	-80.305	4.2731	64.193	6.037	*0.224*
		(62.460)	(8.723)	(56.846)	(8.712)	*(7 793)*	(61.773)	(8.579)	(56.304)	(8.593)	*(7 793)*
	2	68.398	3.7759	-59.147	-3.417	*0.217*	76.628	3.6088	-65.636	0.058	*0.214*
		(58.730)	(8.392)	(53.104)	(8.018)	*(7 793)*	(55.752)	(8.402)	(50.102)	(8.061)	*(7 793)*
	3	40.232	2.7125	-17.012	5.835	*0.229*	68.380	2.3388	-43.233	12.529	*0.233*
		(58.538)	(8.724)	(54.304)	(8.134)	*(7 793)*	(61.045)	(8.707)	(57.104)	(8.188)	*(7 793)*
p70 < i < p80	1	10.498	3.2378	51.115	-3.525	*0.210*	12.504	3.2275	-13.563	0.932	*0.208*
		(17.866)	(8.476)	(50.364)	(8.960)	*(8 199)*	(14.948)	(8.104)	(18.716)	(8.686)	*(8 199)*
	2	11.190*	3.894	-89.969	3.551	*0.204*	8.281*	3.8321	-7.567*	7.457	*0.202*
		(6.819)	(8.953)	(59.860)	(7.317)	*(8 199)*	(5.289)	(8.961)	(3.689)	(7.413)	*(8 199)*
	3	73.369	4.3607	-77.336	-4.998	*0.212*	74.100	4.1271	-77.125	-1.849	*0.217*
		(49.812)	(7.621)	(84.303)	(7.744)	*(8 199)*	(50.724)	(7.529)	(55.208)	(7.698)	*(8 199)*
p80 < i < p90	1	11.115*	3.3207	51.785	0.079	*0.212*	14.069*	3.4963	-16.389*	5.036	*0.209*
		(6.565)	(8.616)	(49.242)	(7.329)	*(8 065)*	(8.643)	(8.508)	(9.957)	(7.338)	*(8 065)*
	2	11.745*	4.3118	-10.313	-9.987	*0.207*	11.498*	4.0788	-12.035*	-3.919	*0.207*
		(6.214)	(8.187)	(42.524)	(6.870)	*(8 065)*	(6.219)	(8.036)	(6.783)	(6.753)	*(8 065)*
	3	12.348	3.0646	-18.735	8.327	*0.213*	19.607	3.2822	-30.416	12.588	*0.217*
		(53.565)	(8.967)	(51.974)	(7.024)	*(8 065)*	(54.382)	(9.141)	(54.232)	(7.131)	*(8 065)*
i > p90	1	15.034*	3.2473	-15.377*	-1.367	*0.225*	18.964*	3.1814	-22.899*	0.031	*0.229*
		(8.451)	(9.504)	(7.807)	(8.055)	*(7 803)*	(10.699)	(9.669)	(12.694)	(8.151)	*(7 803)*
	2	18.266*	4.0196	-17.114*	-7.961	*0.228*	17.760*	4.348	-16.522*	-5.537	*0.230*
		(10.786)	(10.355)	(10.877)	(8.360)	*(7 803)*	(10.428)	(10.291)	(9.051)	(8.449)	*(7 803)*
	3	1.183	4.5955	4.092	-12.581	*0.232*	15.323	4.8923	-7.032	-11.449	*0.235*
		(55.559)	(9.452)	(58.539)	(7.151)	*(7 803)*	(56.161)	(9.626)	(59.092)	(7.209)	*(7 803)*

Table C1.2 reports results from a passive event study that is based on all values of weekly growth in SVI from the period of January 2004 to December 2011. The sample consists of 386 firms listed on NASDAQ OMX Nordic exchanges. The sample consists of 386 firms listed on NASDAQ OMX Nordic exchanges. The results are obtained by running Fama-MacBeth cross-sectional regressions in each decile group. Each decile group (percentile interval) collects values of historical growth in all observations, i, of SVI from one week to the next. The observations with the highest SVI growth are grouped in the interval i > p90 for example. The dependent variable is weekly abnormal return denoted in basis points for week 1, week 2 and week 3 forward in time. In the left part of the table, indices used as benchmarks are OMX Copenhagen, OMX Helsinki, OMX Stockholm and OMX Reykjavik. In the right part of the table, the benchmark used as market return is the OMX Nordic index. The independent variables are abnormal SVI (ASVI), log of market value of company (LMVC) and an interaction term between ASVI and LMVC. The independent variables are defined in Table A2 in appendix A The abnormal SVI value is defined by using the median approach. Every variable is cross-sectionally demeaned to have a regression intercept close to zero. The independent variables are standardized. This means that one can interpret the regression coefficients as the impact of a one-standard-deviation change. 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 percent level. ** represents significance at the 5 percent level. *** represents significance at the 1 percent level.

Table C2: Active event study based on weekly SVI growth, continously updated percentiles

(i) Defining percentiles based on relative growth in SVI, for all observations each specific week

		Abnormal Return, Percentile intervals				
Week	Stats	i > p50	i > p60	i > p70	i > p80	i > p90
1	mean	15.841	15.782	14.799**	15.808***	9.143**
	se	(27.338)	(12.597)	(6.014)	(3.774)	(5.475)
	obs	46 008	38 412	29 198	19 541	9 856
2	mean	11.287	11.441	6.686*	5.982**	8.354*
	se	(22.292)	(14.506)	(3.884)	(3.197)	(5.199)
	obs	45 908	38 334	29 142	19 505	9 838
Cum.	mean	27.129	27.175	21.382**	21.701***	17.117***
1-2	se	(30.196)	(35.140)	(11.061)	(5.068)	(7.262)
	obs	45 877	38 307	29 119	19 487	9 829

(ii) Defining percentile intervals based on relative growth in SVI, for all observations each specific week

		Abnormal Return, Percentile intervals				
Week	Stats	p50 < i < p60	p60 < i < p70	p70 < i < p80	p80 < i < p90	p90 < i
1	mean	10.185	7.891	8.268**	18.583***	9.143**
	se	(15.068)	(24.708)	(4.845)	(5.284)	(5.175)
	obs	11 072	10 460	9 855	9 750	9 856
2	mean	6.668	13.021	13.821**	4.522	8.354*
	se	(14.577)	(21.864)	(6.746)	(4.937)	(5.199)
	obs	11 053	10 439	9 835	9 733	9 838
Cum.	mean	15.107	20.006	22.029**	22.151***	17.117***
1-2	se	(26.757)	(26.593)	(9.614)	(6.112)	(7.262)
	obs	11 047	10 434	9 830	9 725	9 829

Table C2 reports results from several Student's *t*-tests. The sample consists of 386 firms listed on NASDAQ OMX Nordic exchanges. The sample period is from January 2004 to December 2011. Each percentile selects observations, *i*, depending on their relative value of growth in SVI compared to all other firms during the same week, i.e. the interval i > p60 contains all observations that have top 40% SVI growth of all observations during their weeks, for example. The dependent variable is weekly abnormal return denoted in basis points for week 1 and week 2 forward in time. In the top part of the table, each interval contains all observations with SVI growth above the specific percentiles. In the bottom part of the table, each interval is limited both downwards and upwards and constitute deciles. The expected SVI value is defined by using the median approach. The indices used as benchmarks are OMX Copenhagen, OMX Helsinki, OMX Stockholm and OMX Reykjavik. The standard errors are reported within parentheses below the regression coefficients. * represents significance at the 10 percent level. ** represents significance at the 5 percent level. *** represents significance at the 1 percent level.