



EKONOMI- HÖGSKOLAN

Företagsekonomiska institutionen

FEKH89

Företagsekonomi: Examensarbete i finansiering på kandidatnivå

Vårterminen 2021

Den som söker skall finna

En studie om investerarsentiment baserad på Google-data och dess påverkan på svenska reglerade och oreglerade aktiemarknader

Författare:

Marcus Josefsson

William Ramshage

Alexander Serrano

Handledare:

Anamaria Cociorva

Sammanfattning

Examensarbetets titel: Den som söker skall finna

Seminariedatum: 3 juni 2021

Kurs: FEKH89, Examensarbete i finansiering på kandidatnivå, 15 högskolepoäng

Författare: Marcus Josefsson, William Ramshage, Alexander Serrano

Handledare: Anamaria Cociorva

Fem nyckelord: Beteendekonomi, Investerarsentiment, FEARS-index, Google Trender, oreglerad aktiemarknad

Forskningsfrågor: Är det skillnad på positiva och negativa ord i investerarsentiment på Stockholmsbörsen och First North Sweden? Påverkar det konstruerade FEARS-indexet aktieavkastning, volatilitet och handelsvolym och skiljer sig dess potentiella inverkan åt på reglerade och oreglerade marknader?

Syfte: Att skapa ett FEARS-index som mäter investerarsentiment på den svenska aktiemarknaden samt undersöka om det är skillnad på positivt och negativt sentiment både på Stockholmsbörsen och First North Sweden.

Metod: Studien är av kvantitativ karaktär och antar en deduktiv ansats och använder sig av tidsserieanalys. Multipel linjär regression genomförs för att undersöka sambandet mellan variablerna. Studien använder sig även av Google Trender för att konstruera ett index som mäter investerarnas sinnesstämning. Data är huvudsakligen hämtad från Google Trender och Bloomberg.

Teoretiska perspektiv: Det teoretiska ramverket utgörs av finansiella teorier om effektiva marknadshypotesen, beteendekonomi och investerarsentiment samt ett fokus på tidigare studier som undersökt investerarsentiment baserat på just Google-sökningar.

Resultat: FEARS-indexet påverkar avkastning och förändringar i handelsvolym både på Stockholmsbörsen och First North Sweden. Dessa förändringar är starkare på Stockholmsbörsen än på First North Sweden. FEARS-indexen påverkar inte förändringar i volatilitet på av nämnda marknader. Resultaten visar sig robusta även vid användning av andra metoder för att skapa FEARS-indexen.

Slutsats: FEARS-index kan användas för att undersöka nutida och framtida, kortsiktiga samband på den svenska aktiemarknaden. Studien bidrar även med ökad kunskap om vilken sorts sentiment som påverkar den svenska aktiemarknaden i högst utsträckning, samt för vilken sorts marknad som sentimenten har störst påverkan.

Summary

Title: Search and you Shall Find

Seminar date: 3rd of June 2021

Course: FEKH89, Business Administration: Bachelor Degree Project in Financial Management, Undergraduate Level, 15 ECTS

Authors: Marcus Josefsson, William Ramshage, Alexander Serrano

Advisor: Anamaria Cociorva

Key words: Behavioral finance, investor sentiment, FEARS-index, Google trends, Multilateral Trading Facility.

Research questions: Is there a difference between positive and negative words in investor sentiment on OMX Stockholm and First North Sweden? Does the constructed FEARS-index influence returns, volatility and trading volumes and is there a difference in influence between regulated and unregulated stock markets?

Purpose: To construct a FEARS-index that measures investor sentiment in the Swedish stock market, in addition to examining whether there is a difference in positive and negative sentiment on OMX Stockholm and First North Sweden.

Methodology: The report is quantitative with a deductive approach, by applying time series analysis. Multiple linear regression is conducted to examine the relationship between the variables. The study utilizes Google Trends to construct an index that measures investor sentiment. Data is mainly retrieved from Google Trends and Bloomberg.

Theoretical perspectives: Financial theories on effective market hypothesis, behavioral finance and investor sentiment, with a focus on previous studies on investor sentiment based on Google search data.

Results: The FEARS-index influences stock returns and changes in trading volumes on both OMX Stockholm and First North Sweden. The influence is stronger on OMX Stockholm than First North Sweden. The FEARS-indexes do not influence volatility in either of the markets. The results are robust according to other methods of constructing the FEARS-indexes.

Conclusions: FEARS-indexes can be applied to examine contemporaneous and future short-term relationships on the Swedish stock market. The contributions of the study are which type of sentiment that affects the Swedish stock market and how it differs between regulated and unregulated markets.

Förord

Vi riktar ett stort tack till vår handledare Anamaria Cociorva som har följt denna uppsatsen och återkopplat med stort intresse och värdefulla synpunkter. Vi hoppas att uppsatsen bidrar med nya intressanta insikter för läsaren och att den kan inspirera till framtida studier.

Lund den 28 maj 2021

Marcus Josefsson, William Ramshage och Alexander Serrano

Begreppslista

FEARS-index: *Financial and Economic Attitudes Revealed by Search* är ett index för investerarsentiment baserad på sökmotordata.

First North Sweden: De svenska aktierna på den oreglerade aktiemarknaden Nasdaq First North.

First North Sweden All-Share: Aktiekursindex som innehåller alla aktier som är listade på First North Sweden.

Google Trender: Ett verktyg skapat av Google som låter användaren se hur sökintresset för ett givet sökord förändrats över tid.

Investerarsentiment: Investerares sinnesstämning eller humör för aktiemarknader.

OMXS: Aktiekursindex som innehåller alla aktier noterade på den Stockholmsbörsen.

Stockholmsbörsen: Samtliga aktier på den svenska, reglerade aktiemarknaden Nasdaq OMX Stockholm.

INNEHÅLLSFÖRTECKNING

1. INLEDNING	7
1.1. Bakgrund	8
1.2. Problemdiskussion	9
1.3. Syfte	12
1.4. Studiens relevans	12
1.5. Avgränsning	13
1.6. Disposition	13
2. TEORI OCH UTVECKLANDE AV HYPOTESER	15
2.1. Teoretiskt ramverk	15
2.1.1. Effektiva marknadshypotesen	15
2.1.2. Beteendekonomi och investerarsentiment	16
2.1.3. Positivt och negativt sentiment	17
2.1.4. Vilka investerare påverkas av sin sinnesstämning?	18
2.2. Tidigare studier	20
2.2.1. Mått på investerarsentiment	21
2.2.2. Google Trender och aktieavkastning	22
2.2.3. Google Trender, volatilitet och handelsvolym	23
2.3. Sammanställning av teorier och tidigare forskning	24
2.4. Litteraturkritik	26
2.5. Utvecklande av hypoteser	27
2.5.1. FEARS-indexets kraft på olika marknader	27
2.5.2. Negativt och positivt sentiment	28
2.5.3. Investor sentiment med Google Trender	29
3. METOD OCH DATAINSAMLING	31
3.1. Val av metod	31
3.2. Datainsamling och databehandling	31
3.2.1. Google sökvolymsindex	32
3.2.2. Databehandling av Google sökvolymsindex	34
3.2.3. Tidsperiod och geografisk region	36
3.2.4. Ordlistan	37
3.2.5. Konstruktion av FEARS-indexet	38
3.2.6. Nasdaq OMX Stockholm och First North Sweden	39
3.2.7. Kontrollvariabler	40
3.3. Metod för hypotesprövning	42
3.3.1. Wilcoxon rank sum test	42
3.3.2. Regressionsanalyser med minstakvadratmetoden	44
3.3.3. Variabelanalys	45
3.3.3.1. Multikollinearitet	45
3.3.3.2. Linjäritet	46
3.3.4. Residualanalys	46

3.3.4.1. Homoskedacitet	47
3.3.4.2. Stationäritet	47
3.3.4.3. Ingen autokorrelation	48
3.3.4.4. Normalfördelade felterm	49
3.4. Akaike Information Criterion	49
3.5. FEARS-index och avkastning	50
3.6. FEARS-index, volatilitet och handelsvolym	51
3.7. Metod- och datakritik	53
3.8. Sammanställning av hypoteser och metodik	56
4. RESULTAT OCH ANALYS	59
4.1. Hypotes 1 - Jämförelse av positivt och negativt sentiment	59
4.1.1. Resultat för hypotes H1a-H1d	59
4.1.2. Analys av hypotes H1a-H1d	61
4.2. Hypotes 2 - FEARS-index och avkastning	62
4.2.1. FEARS-index och avkastning i tidpunkt t	63
4.2.2. FEARS-index och avkastning i tidpunkt t+1	65
4.2.3. Skillnaden mellan FEARS-index och avkastning på olika marknader	66
4.3. Hypotes 3 - FEARS-index och volatilitet	67
4.3.1. FEARS-index och volatilitet i tidpunkt t och t+1	68
4.3.2. Analys av FEARS-index och volatilitet	70
4.3.3. Skillnaden mellan FEARS-index och volatilitet på olika marknader	71
4.4. Hypotes 4 - FEARS-index och handelsvolym	71
4.4.1. FEARS-index och handelsvolym i tidpunkt t	72
4.4.2. FEARS-index och handelsvolym i tidpunkt t+1	74
4.4.3. Skillnaden mellan FEARS-index och volatilitet på olika marknader	75
4.5. FEARS-indexets robusthet - AIC-modellen	75
4.6. Diagnostiska tester	77
5. SLUTSATSER OCH DISKUSSION	79
5.1. Slutsatser	79
5.2. Diskussion	79
5.3. Vidare forskning	81
KÄLLFÖRTECKNING	83
Appendix 1. Baslista med sökord	91
Appendix 2. Testresultat och statistik	93

1. INLEDNING

Uppsatsens inledning ger en överblick av ämnesområdet och argumenterar varför ämnesområdet är intressanta att undersöka. Syftet med undersökningen presenteras, tillsammans med vad studien bidrar med, praktiska avgränsningar och strukturen för resten av uppsatsen.

1.1. Bakgrund

Grundläggande finansiella teorier förutsätter att priserna på en effektiv marknad återspeglar all tillgänglig information. En effektiv prissättning förutsätter därmed att investerarna agerar fullt rationellt både i sin informationsinsamling, uppfattning av informationen och i sitt beslutsfattande. Denna process är dock i högsta grad subjektiv. Investerare är bara människor. I sin inflytelserika bok "The General Theory of Employment, Interest and Money" reflekterar ekonomen John Maynard Keynes (1936) bland annat över att människan inte alltid tar noga beräknade beslut. Det ligger i människans natur att ibland låta sig påverkas av sporadiska insikter och optimism. Ibland kan människan plötsligt agera, vare sig det är logiskt grundat eller inte. Keynes benämner dessa inslag "*animal spirits*". I en ekonomisk kontext är idén en tidig förlaga till vad som senare kom att kallas beteendekonomi.

Beteendekonomi utgår från att investerarens beslutsfattande inte alltid är fullt rationellt. Beslutsfattandet kan påverkas av känslor, sinnesstämning, brist på tid eller förmåga att tillskansa sig information eller att investeraren saknar självkontroll. Beteendekonomi kan förklara vissa av de händelser på aktiemarknaden som annars inte är förenliga med klassiska finansiella teorier baserade på investerarnas rationalitet. Den Svarta måndagen 1987 och IT-bubblan på 90-talet är två exempel på sådana händelser (Baker & Wurgler, 2007). Beteendekonomiska faktorer kan alltså få reella konsekvenser på prissättningen på kort sikt.

Idén om att investerare påverkas av sin sinnesstämning, så kallat *investerarsentiment*, är en faktor som kan tas i beaktning. Investerarsentiment är investerarens humör eller mentalitet, och som med många människor kan detta påverka individens agerande (Baker & Wurgler, 2007). Till exempel kan en optimistisk investerare ha en övertro på positiva nyheter och underskatta negativa nyheter, och följaktligen välja att köpa mer av en aktie. Det motsatta gäller för pessimistiska investerare som undervärderar positiva nyheter och övervärderar

negativa nyheter. På senare tid har idén om investerarsentiment fått fäste och mer forskning har syftat till att undersöka sambandet mellan de finansiella marknaderna och investerarsentiment. Baker och Wurgler (2007) beskriver områdets potential, men även en av de stora svårigheterna som forskningen om investerarsentiment står inför: “Now, the question is no longer, as it was a few decades ago, whether investor sentiment affects stock prices, but rather how to measure investor sentiment and quantify its effects.” (Baker & Wurgler, 2007, sid. 130).

Google Trender har visat sig vara ett användbart verktyg för att mäta människors beteende och tankegångar. Google Trender är ett verktyg som låter användaren se hur populärt sökintresset är för ett givet sökord på Google. För drygt tio år sedan lanserade Google det så kallade *Google Flu Trends*. Sökmotorföretaget såg ett samband mellan när influensasäsongerna inträffade och antalet sökningar på ord som är relaterade till influensan, såsom olika symptom. Med denna informationen lyckades företaget i förväg prediktera när influensasäsonger skulle inträffa. Resultatet blev intressant för hälsomyndigheter och sjukhus världen över eftersom sjukvården lättare kunde förbereda sig för fler patienter inom en snart framtid (Ginsberg, Mohebbi, Patel, Brammer, Smolinski & Brilliant, 2009). Med detta resultat visade de alltså att det går att förutsäga mänskligt beteende utifrån vad människor söker efter på internet.

1.2. Problemdiskussion

En ökning av icke-professionella investerare har skett på senare år, drivet av bland annat lägre courtageavgifter och lättillgängliga handelsplattformar (Weber Shandwick, 2021). Denna grupp investerare får allt mer påverkan på aktiemarknader. Till skillnad från traditionella investerare som framför allt tar hänsyn till finansiella nyheter, årsredovisningar och rapporter från professionella aktieanalytiker, föredrar icke-professionella investerare ofta onlineforum och sociala medier (Weber Shandwick, 2021). Ett aktuellt och välkänt exempel är aktien Gamestop. I början av 2021 identifierade en grupp icke-professionella investerare att finansinstitut hade tagit en kort position och organiserade sig på internet för att dra nytta av situationen (Weber Shandwick, 2021).

Ökningen får ytterligare komplikationer då icke-professionella investerare tenderar att påverkas mer av sin sinnesstämning (Baker & Wurgler, 2007; Barber, Odean & Zhu, 2009;

Foucault, Sraer & Thesmar, 2011). Detta märks av i exemplet Gamestop där de icke-professionella investerarna verkade ha en övertro till hur mycket finansinstitutens korta positioner gick att utnyttja. Aktiekursen drevs i sin tur till nivåer som var långt ifrån motiverade av fundamentala analyser. Detta tyder på att de icke-professionella investerarna påverkades av sin sinnesstämning, samtidigt som deras agerande hade en faktisk påverkan på prissättningen, åtminstone på kort sikt. Händelsen motiverar varför det är intressant att hitta ett tillvägagångssätt för att mäta investerares sinnesstämning eftersom sinnesstämningen kan ge en indikation om i vilken riktning marknaden är på väg.

På den svenska marknaden gör Konjunkturinstitutet sådana mätningar på ett makroekonomiskt plan. Konjunkturinstitutet publicerar ekonomiska förväntningar på framtiden och upplevelser av nuläget varje månad genom enkäter där 5500 företag och 1500 hushåll deltar (Konjunkturinstitutet, 2021). Trots att denna metod omfattar förväntningar på den svenska ekonomin på ett välstrukturerat vis kan undersökningarna antas vara kostsamma och ett större urval hade kunnat bidra till bättre precision. I enkätbaserade undersökningar löper man även risken att få oärliga och subjektiva svar, särskilt om frågorna berör ett känsligt ämne. Med tanke på att allt fler investerare baserar sina beslut på information från internet kan det dessutom finnas potential i alternativa metoder. Konjunkturinstitutets mätningar på hushåll sker månatligen (Konjunkturinstitutet, 2021), vilket föranleder ett behov av högre svarsfrekvens då börsrörelser blivit allt snabbare och extrema (Oldstrand, 2020).

I sin artikel "The Sum of All FEARS Investor Sentiment and Asset Prices" skapade Da, Engelberg och Gao (2015) ett så kallat FEARS-index (*Financial and Economic Attitudes Revealed by Search*) med hjälp av data över sökaktivitet på Google. Författarna studerade negativt laddade ord relaterade till ekonomi och fann att när sökaktiviteten för dessa ord ökade tenderade aktiekurser att falla de nästkommande dagarna och stiga därefter. Detta förklaras med att när investerare är oroliga och pessimistiska inför framtiden söker de på fler negativt laddade ord än vanligt, och därefter agerar de utifrån sitt humör. FEARS-indexet blev således ett mått på investerarnas sinnesstämning. Sökdata från Google är både gratis och tillgänglig i realtid. Detta gör FEARS-indexet till ett verktyg som är enklare och mer tillgängligt på begäran jämfört med andra mått på investerarnas sinnesstämning. Da, Engelberg och Gao (2015) konstruerade dock sitt FEARS-index med den amerikanska aktiemarknaden som utgångspunkt, och det finns endast ett fåtal forskare som har genomfört liknande studier utanför USA. Den svenska aktiemarknaden är inte en av dem som har

undersökts. Vidare använde sig författarna endast av negativt laddade ord eftersom dessa visade sig ha starkast samband med aktiekursutvecklingen just på den amerikanska marknaden under den studerade tidsperioden. Till skillnad från deras studie kommer vi jämföra såväl positivt som negativt sentiment. Dessutom genomfördes studien på en reglerad aktiemarknad. Deras resultat behöver därför inte nödvändigtvis hålla på oreglerade marknader eftersom kraven på företagen är lägre och sammansättningen av investerare kan vara annorlunda på dessa marknader.

I Sverige är en oreglerad aktiemarknad en handelsplattform där mindre företag kan vara listade. Syftet med att ha en oreglerad aktiemarknad är att skapa en offentlig handelsplattform där kraven på att vara listad är lägre jämfört med på en reglerad börs. Detta är något som ofta gynnar mindre företag eftersom det är kostsamt att vara börsnoterad på en reglerad marknad och följa alla dess regler (Nasdaq OMX Nordic, 2021). På den oreglerade marknaden får det mindre företaget fortfarande tillgång till externt kapital. I december 2019 stod OMX Stockholm för 96,51% av marknadsvärdet på den svenska aktiemarknaden. First North Sweden stod för 2,88% (SCB, 2021a).

Den amerikanska motsvarigheten till oreglerade aktiemarknader kallas för *Alternative Trading Systems* (ATS) och fungerar inte likadant som de oreglerade svenska marknaderna. Till skillnad mot de oreglerade aktiemarknaderna i Sverige fungerar ett ATS som en mellanhand som parar ihop köpare och säljare utanför de offentliga aktiemarknaderna. Dock är det ingen skillnad på vilka aktier som handlas på ett ATS och en reglerad handelsplattform. ATS:er har samtidigt lägre krav på att offentliggöra information om en transaktion och dess parter innan transaktionen har genomförts (Goldman Sachs, 2009). Det går alltså inte att vara listad på ett ATS på samma sätt som på den svenska oreglerade marknaden First North Sweden. Ett ATS är snarare ett sätt att handla med aktier över disk.

Det är intressant att jämföra reglerade och oreglerade aktiemarknader i Sverige av flera anledningar. För det första har de svenska hushållen en större andel av sitt kapital placerat på oreglerade marknader jämfört med andra samhällssektorer (SCB, 2021a), och som nämnt tenderar denna typen av investerare att påverkas lättare av sin sinnesstämning. Dessutom är företagen på oreglerade marknader ofta mindre och denna typen av aktier tenderar att vara mer volatila (Baker & Wurgler, 2007). Detta brukar innebära att aktierna blir mer känsliga för förändringar i investerarsentiment jämfört med större företag (Baker & Wurgler, 2006). Det

är alltså intressant att jämföra reglerade och oreglerade marknader dels på grund av vilka investerare som handlar, dels på grund av vilka sorters aktier som handlas på de olika marknaderna. Slutligen är det särskilt intressant att jämföra reglerade och oreglerade aktiemarknader i just Sverige eftersom det inte går att göra denna jämförelsen i USA. Studier om FEARS-index har aldrig jämfört reglerade och oreglerade aktiemarknader, varken i eller utanför USA.

1.3. Syfte

Syftet är att skapa ett FEARS-index som mäter investerarsentiment på den svenska aktiemarknaden samt undersöka om det är skillnad på positivt och negativt sentiment både på Stockholmsbörsen och First North Sweden.

1.4. Studiens relevans

Genom att skapa det första FEARS-indexet på den svenska aktiemarknaden bidrar vår studie till en ökad förståelse för hur svenska investerare handlar utifrån sin sinnesstämning, vilket är ett område med begränsad kunskap sedan tidigare. FEARS-indexet bidrar även med ett nytt enkelt verktyg för att mäta investerarsentiment i realtid och på begäran. Andra mått är släpande, kostsamma och har sämre dataupplösning jämfört med FEARS-indexet. Studien belyser även skillnaden mellan hur stor påverkan investerares sinnesstämning har på handeln på reglerade aktiemarknader jämfört med oreglerade aktiemarknader. Detta är något som inte har gjorts i Sverige tidigare och det är även en ny tillämpning av FEARS-indexet att jämföra reglerade och oreglerade marknader.

I denna studien undersöker vi perioden från den 1a januari 2015 till den 31 december 2019. Vår valda tidsperiod är särskilt intressant att undersöka eftersom den har annorlunda egenskaper jämfört med den som Da, Engelberg och Gao (2015) använde när de skapade det ursprungliga FEARS-indexet. De studerade en längre tidsperiod som bland annat innefattade finanskrisen 2008. Deras valda tidsperiod innehöll både perioder av tydliga nedgångar och uppgångar på aktiemarknaden, medan vår valda tidsperiod snarare präglas av uppgångar.

Slutligen vidgar vi kunskapsområdet genom att jämföra positivt och negativt sentiment. Till skillnad från tidigare studier utför vi tester för att undersöka om positivt eller negativt laddade ord har starkast samband med avkastningen på den svenska aktiemarknaden.

1.5. Avgränsning

Studien är avgränsad till den svenska aktiemarknaden eftersom det inte finns ett FEARS-index som är anpassat för den svenska aktiemarknaden sedan tidigare. Den svenska aktiemarknaden skiljer sig dessutom mot den amerikanska aktiemarknaden. I Sverige finns oreglerade marknadsplatser för aktiehandel, som exempelvis First North.

Vi har avgränsat oss till Sverige när det gäller den geografiska regionen för sökaktiviteten på Google eftersom vi är intresserade av svenskarnas sökaktivitet. Vi använder oss enbart av svenska översättningar av söktermerna på Google för att förenkla skapandet av FEARS-indexet. Det går att argumentera för att även de engelska översättningarna av söktermerna ska inkluderas eftersom svenskar kan tänkas googla på både svenska och engelska. Detta skulle dock komplicera skapandet av FEARS-indexet. En stor fördel med FEARS-indexet är just att det är tillgängligt och enkelt att skapa, och denna fördelen går förlorad om alltför många bearbetningar behöver göras. Vi går in på detta djupare i metodavsnittet.

Vi har medvetet avgränsat oss till tiden innan covid-19-pandemin eftersom denna turbulenta tid kan ses som en anomali på aktiemarknaden med stora prisfluktuationer. Om tidsperioden hade täckt krissituationer får det konsekvenser för studiens generaliserbarhet under mer normala marknadsförhållanden. Anledningen till varför vi inte använder en tidsperiod som sträcker sig längre bakåt i tiden är för att detta får negativa konsekvenser för datainsamlingen och databehandlingen, vilket diskuteras vidare i metodavsnittet. Avslutningsvis har vi valt att enbart undersöka Stockholmsbörsen som reglerad marknad och First North Sweden som oreglerad marknad, även om det finns fler marknader i Sverige. Anledningen är att First North Sweden och Stockholmsbörsen är de i särklass största reglerade och oreglerade marknaderna i Sverige (SCB, 2021a).

1.6. Disposition

I nästkommande avsnitt presenteras de grundläggande teorier som uppsatsen bygger på och tidigare forskning inom samma ämnesområde. Vi gör även en kritisk granskning av litteraturen. Avsnitt 2 mynnar ut i de hypoteser som vi avser att pröva för att uppnå uppsatsens syfte. I avsnitt 3 beskrivs tillvägagångssättet för att uppnå syftet. Vi presenterar datainsamling, databehandling, praktiska tester som prövar våra hypoteser samt argumenterar för varför vårt tillvägagångssätt är lämpligt. I avsnitt 4 presenteras och analyseras resultaten av våra tester utifrån studiens syfte och de teorier och tidigare forskning som presenterades i avsnitt 2. I avsnitt 5 formulerar vi slutsatser och besvarar syftet med studien. Avslutningsvis diskuteras studien ur ett bredare perspektiv och vi ger förslag till framtida forskning.

2. TEORI OCH UTVECKLANDE AV HYPOTESER

Teoriavsnittet ämnar att fastställa det teoretiska ramverket som studien utgår från och redogöra för tidigare studier inom ämnesområdet. Litteraturgenomgången börjar med att fastställa breda och grundläggande teorier för ämnesområdet och presenterar därefter forskningsresultat kring investerarsentiment i mer detalj. Därefter görs en kritisk granskning av tidigare studier. Till sist mynnar avsnittet ut i fyra hypoteser.

2.1. Teoretiskt ramverk

Det teoretiska ramverket är en genomgång av de grundläggande teorier som denna studien och tidigare forskning bygger på.

2.1.1. Effektiva marknadshypotesen

Den effektiva marknadshypotesen säger att priserna på en effektiv marknad återspeglar all tillgänglig information (Fama, 1970). En marknad kan vara starkt, semi-starkt eller svagt effektiv beroende på hur mycket av den tillgängliga informationen som återspeglas i priserna på marknaden. En marknad är svagt effektiv om det inte går att förutspå framtida aktiekursförändringar med tidigare prisförändringar och trender. Aktiekursutvecklingen brukar då jämföras med en *random walk* som beter sig helt slumpmässigt och oberoende av tidigare värden (Lo & Mackinlay, 1988). På en starkt effektiv marknad, till skillnad från en svagt effektiv marknad, går det inte att förutspå framtida aktiekursförändringar med någon information över huvud taget, oavsett om den är tillgänglig för allmänheten eller inte (Fama, 1970). Den effektiva marknadshypotesen används som grund för mycket inom finansiell teori, bland annat för beräkningen av onormal avkastning. Det finns dock invändningar mot att denna teorin håller. Studier i beteendekonometri tar istället ett psykologiskt perspektiv och argumenterar för att investerares beslut påverkas av sociala och psykologiska faktorer. Besluten tas därmed inte alltid på de rationella grunder som den effektiva marknadshypotesen föreslår (Malkiel, 2003).

Enligt Shaker (2013) uppfyller inte OMXS30, vilket är ett kapitalviktat aktieindex med de 30 mest omsatta aktierna på Stockholmsbörsen, den svaga formen av effektiva marknadshypotesen. Den dagliga avkastningen visade sig vara beroende av historiska värden och var således ingen *random walk*. Det indikerar att det finns störningar på den svenska

aktiemarknaden som gör det möjligt att prediktera framtida aktieavkastning. Resultatet av Shaker (2013) föranleder att beteendekonomi är ett intressant område med en eventuell legitim påverkan på svenska aktiemarknader.

Förutom bevisen att de reglerade svenska aktiemarknaderna inte kan anses vara effektiva finns det anledning att tro att beteendekonomiska faktorer har olika påverkan på olika aktiemarknader i Sverige. En skillnad mellan olika aktier som handlas i Sverige är uppdelningen mellan reglerade och oreglerade handelsplatser. I Sverige driver Nasdaq både reglerade och oreglerade aktiemarknader. Kraven på företag som är noterade på en reglerad börs, som exempelvis Stockholmsbörsen, är högre jämfört med dem på First North Sweden, som är en av Nasdaqs oreglerade svenska handelsplattformar. De högre kraven gäller framförallt redovisning och rapportering (Nasdaq, 2021a; Nasdaq, 2021b; Nasdaq OMX Nordic, 2021). Exempelvis krävs det färre finansiella rapporter innan börsnotering och färre finansiella rapporter per år för företag på First North Sweden. Dessutom behöver företag på First North Sweden ha anlitat en certifierad rådgivare som kontrollerar att företaget följer handelsplattformens regler (Nasdaq, 2021b). Detta antyder att det kan finnas en större risk för informationsasymmetri och spekulation om företagsspecifik information inte sprids eller är lika tillgänglig på de olika marknaderna. Detta skulle i sådana fall göra de oreglerade marknaderna mindre effektiva jämfört med de reglerade marknaderna. Även på Nasdaqs egna hemsida får läsaren en försiktig upplysning om att de oreglerade marknaderna kan ha större inslag av beteendekonomi: "Ju färre regler, desto större ansvar faller på dig som investerare." (Nasdaq OMX Nordic, 2021).

2.1.2. Beteendekonomi och investerarsentiment

Till skillnad från forskning av Fama (1970) utgår beteendekonomisk teori från att investerare påverkas av en övertygelse om framtida utveckling på aktiemarknader (Brown & Cliff, 2005; De Long et al., 1990). Förväntningar på framtida kassaflöden och risk som inte kan hänvisas till rationella grunder betecknas som investerarsentiment. Baker och Wurgler (2007) anser att de höga värderingar på teknikaktier innan IT-kraschen i början av 2000-talet exemplifierar ett högt investerarsentiment. Detta har på senare tid blivit ett forskningsområde med syftet att mäta investerarsentiment och förutse dess påverkan på finansiella marknader.

Brown och Cliff (2005) vidgar perspektivet från tidigare forskning som koncentrerat sig på felaktig prissättning på enstaka aktier eller inom särskilda sektorer, och undersöker istället påverkan på en stor mängd finansiella tillgångar, med slutsatsen att investerarsentiment kan förutsäga avkastning under de närmaste tre åren. Baker och Wurgler (2006) presenterar resultatet att framtida avkastning är beroende av investerarsentiment vid periodens början. Ett högt sentiment innebär att investerare är positivt inställda och optimistiska inför framtiden, och ett negativt sentiment tvärt om. Ett inledningsvis högt sentiment förutspår en låg avkastning de kommande månaderna. Om investerarsentimentet istället skulle vara lågt inledningsvis förutspås en hög avkastning framöver. Dessa resultaten är starkast framförallt på spekulativa aktier, det vill säga aktier som ofta karakteriseras av hög volatilitet, hög tillväxt och ingen utdelning. Resultaten brukar även komma tillsammans med ett mönster som kallas *prisivändning (price reversal)*. Ett prisvändningsmönster innebär att priserna tillfälligt går upp eller ner som en konsekvens av förändringar i sentiment, men att de därefter återgår till de ursprungliga värdena när sinnesstämningen återhämtar sig. Detta mönster stämmer överens med teorin om att investerarsentiment kan driva prissättningen bort från de fundamentala priserna, men att de så småningom återvänder. Resultaten ligger även i linje med tidigare teorier om investerarsentiment som förklarar varför sinnesstämningen får investerare att agera irrationellt. Optimistiska investerare tenderar nämligen att övervärdera positiva nyheter och undervärdera negativa nyheter, medan negativa investerare tenderar att göra tvärt om (Barber & Odean, 2008).

Baker och Wurgler (2007) menar att det som gör spekulativa aktier särskilt känsliga för investerarsentiment är att de är svåra att prissätta och prissättningen är subjektiv. Författarna ger som exempel ett ungt företag med hög tillväxtpotential men som för tillfället inte är lönsamt och som har en osäker framtid. Denna beskrivningen passar bra in på företag som handlas på First North Sweden.

2.1.3. Positivt och negativt sentiment

Att något som uppfattas som negativt har starkare effekt än dess positiva motsvarighet är en grundsats som fått ett så stort stöd att det ofta ses som en av psykologins mest universella sanningarna (Baumeister, Bratslavsky, Finkenauer & Vohs, 2001). Baumeister et al. studerar detta fenomen i "Bad is Stronger than Good", där de slår fast uppfattningen av de två orden genom en omfattande sammanställning av dess förekomst utifrån flertalet ämnesområden. En

av de grundläggande anledningarna till denna förekomst menar de beror på hur dåliga händelser orsakar större emotionella upplevelser och mer långtgående effekter. Av denna anledning tenderar människan att lägga oproportionerligt stor vikt vid att söka och bearbeta information för att undvika dessa händelser, gentemot den tid som de lägger ner för att finna positiva händelser.

I en ekonomisk kontext går detta att koppla till ett framträdande fenomen inom beteendekonomi, nämligen dispositionseffekten (Barberis & Thaler, 2003). Dispositionseffekten innebär att investerare avstår från att sälja aktier som presterar dåligt och säljer istället de aktier som går bra. I vissa situationer hade det varit rationellt att realisera sina förluster och istället investera i något annat. Dispositionseffekten gör dock att investeraren avstår från att sälja eftersom det skulle innebära att investeraren ger upp möjligheten att aktien kommer ge positiv avkastning senare. Om investeraren aldrig säljer kan den aldrig gå med förlust. Detta är ytterligare ett beteendekonomiskt fenomen som kan förklara irrationellt beteende (Barberis & Thaler, 2003).

2.1.4. Vilka investerare påverkas av sin sinnesstämning?

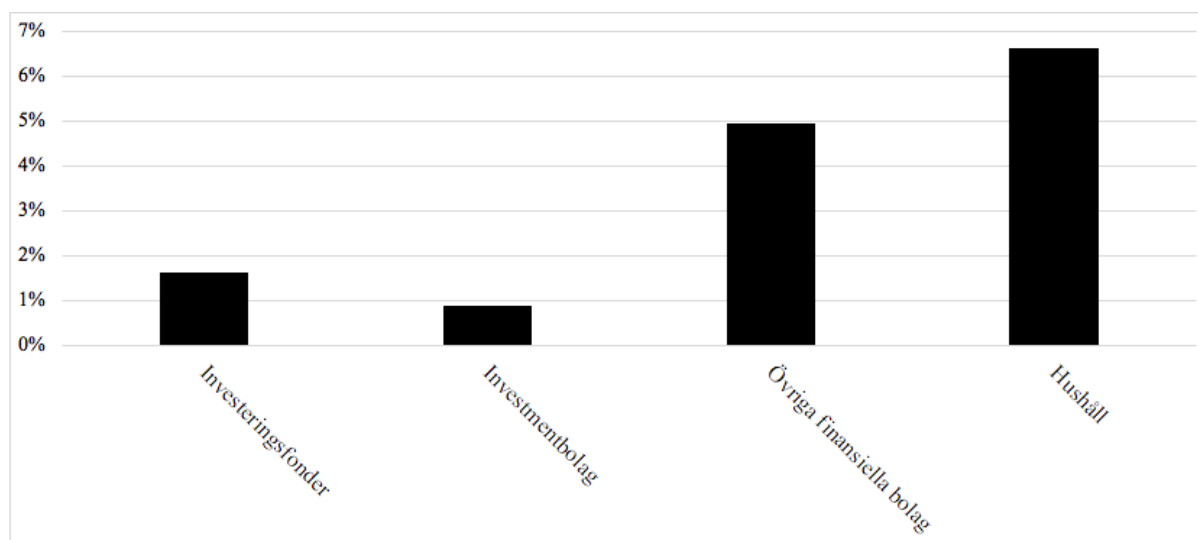
Det är viktigt att veta under vilka förutsättningar som beteendekonomiska forskningsresultat är applicerbara. Forskningsresultat inom beteendekonomi går nämligen inte att applicera på alla typer av investerare. Som bekant hävdar beteendekonomin att vissa fenomen på de finansiella marknaderna delvis kan förklaras av att alla parter inte agerar fullt så rationellt som andra teorier förutsätter (Barberis & Thaler, 2003; Ritter, 2003). Samtidigt betyder inte detta att det inte finns aktörer som agerar fullt rationellt utifrån tekniska och fundamentala analyser. Denna typen av investerare kallas *fundamentalister* eller *arbitrageinvesterare* enligt Baker och Wurgler (2006). Fundamentalister är de investerare som utför finansiella beräkningar för att värdera finansiella instrument, utifrån exempelvis årsredovisningar, jämförelser med liknande företag på marknaden och annan finansiell information. Investeraren tar därefter beslut utifrån om den nuvarande prissättningen är motiverad av den fundamentala analysen eller inte. Forskningen om investerarsentiment behandlar dock irrationella investerare, de som Baker och Wurgler (2006) kallar för *spekulativa investerare* och som i andra sammanhang benämns *noise traders* (Barberis & Thaler, 2003).

Vad som kännetecknar spekulativa investerare är att de agerar utifrån marknadsbrus (*noise*) som om det vore information (Black, 1986). Marknadsbrus, till skillnad från information, är händelser, faktorer, omständigheter och osäkerhet om framtiden som inte bör, eller som redan har, inkorporerats i priserna på marknaden. Det kan vara att en investerare får en övertro och köper en viss aktie efter en rekommendation från någon annan, och håller följaktligen en portfölj som inte är tillräckligt diversifierad. En spekulativ investerare kan ta till sig gamla nyheter om en försämrade framtidsutsikt för en industri, och sedan agera utifrån denna informationen som om priserna på marknaden inte redan har inkorporerat informationen (De Long, Shleifer, Summers & Waldmann, 1990). Denna beskrivningen får en spekulativ investerare att framstå som en oinformerad investerare, och mycket riktigt är denna investerartypen vanligare bland icke-professionella investerare (Baker & Wurgler, 2007; Barber, Odean & Zhu, 2009; Foucault, Sraer & Thesmar, 2011).

Tidigare studier har avfärdat att en spekulativ investeringsstrategi skulle ha en påtaglig inverkan på prissättning med motiveringen att fundamentalister utnyttjar spekulativa investerares oinformerade beslut emot dem, och driver således tillbaka priserna till rationellt motiverade värden (Fama, 1965). Senare litteratur erkänner däremot att till exempel investerarsentiment kan påverka prissättningen på finansiella marknader (Baker & Wurgler, 2007; Brown & Cliff, 2005; Da, Engelberg & Gao, 2015). Anledningen till detta är att det tycks finnas en så kallad gräns för arbitragemöjligheter. Till exempel kan investeringsföretag ha gränser för hur stor en kort position får vara. En professionell investerare kan i sådana fall inte fullt utnyttja att spekulativa investerare har drivit upp priserna långt ifrån deras fundamentala värden (De Long et al., 1990).

Icke-professionella investerare tenderar att vara spekulativa investerare av flera anledningar enligt Barber och Odean (2008). För det första påverkas de mer av sitt sentiment. För det andra saknar de tillgång till information och verktyg för informationssökning i samma utsträckning som professionella investerare. Slutligen har icke-professionella investerare ofta inte tillräckligt med tid för att uppmärksamma all information som är tillgänglig på marknaden. Icke-professionella investerare agerar därför endast utifrån den information som de redan har uppmärksammat. Sentiment kan avgöra vilket beslut en investerare gör, medan den begränsade uppmärksamheten avgör från vilket urval av information som en investerare kan ta beslut (Barber & Odean, 2008). Detta beskriver väl hur tätt förknippade investerares uppmärksamhet och investerarsentiment är för icke-professionella investerare.

Det finns anledning att tro att det finns en större andel oinformerade investerare på de oreglerade marknaderna i Sverige jämfört med de reglerade marknaderna. Figur 1 visar hur stor andel av olika svenska sektors aktieäggande som var placerat på en oreglerad svensk marknad i genomsnitt under perioden december 2014 till december 2019 (SCB, 2021a). Där framgår det att de svenska hushållen i genomsnitt hade drygt 6,5 % av sitt investerade kapital placerat på en oreglerad svensk marknad. Sektorerna för investmentbolag, investeringsfonder och övriga finansiella bolag (vilket inkluderar bland annat banker) hade placerat 0,9 %, 1,6 % respektive 5 % av kapitalet på oreglerade marknader. Dessa tre sektorer kan antas vara mer informerade investerare, och alla tre har mindre andel kapital placerat på oreglerade marknader jämfört med de svenska hushållen.



Figur 1: Andel av investerat kapital placerat på oreglerade svenska aktiemarknader. Fördelat på olika svenska sektorer. Genomsnitt över perioden december 2014 till december 2019 (SCB, 2021a).

2.2. Tidigare studier

I detta avsnittet gör vi en litteraturgenomgång som är specialiserad inom ämnesområdena investerarsentiment och Google Trender. Vi redovisar och diskuterar resultat från tidigare studier som kommer komma till användning när vi utformar våra egna hypoteser och vårt tillvägagångssätt.

2.2.1. Mått på investerarsentiment

Förutom de redan nämnda enkäterna har tidigare studier använt flera andra klassiska mått på investerarsentiment (Baker & Wurgler, 2007). Eftersom icke-professionella investerare är de som är mest benägna att påverkas av sin sinnesstämning kan icke-professionella investerares orderböcker användas som mått på investerarsentiment. Logiken är att investerare som investerar är optimistiska, och investerare som väljer att sälja är pessimistiska. Ibland används även handelsvolymen på hela marknaden och inte bara från icke-professionella investerare (Baker & Wurgler, 2007). Sentiment kan också mätas genom att undersöka vilka instrument investerare väljer att investera i, och framförallt vilka fonder. Baker och Wurgler argumenterar att investerare som placerar sitt kapital i riskfyllda aktiefonder är optimistiska, medan sinnesstämningen på marknaden är negativ när investerare istället väljer att placera kapital i "säkrare" räntefonder. Antalet börsnoteringar och förstadagsavkastningen vid börsnoteringar har också använts som mått på investerarsentiment. Slutligen, eftersom värdet på optioner ökar när den förväntade volatiliteten på den underliggande tillgången ökar, kan optionspriser användas för att härleda den underförstådda förväntade volatiliteten. Den förväntade volatiliteten blir ett tecken för investerares framtidstro. Det amerikanska VIX-indexet är baserat på just den underförstådda volatiliteten på optioner på S&P 100-indexet, och är ett typiskt mått på investerarsentiment (Baker & Wurgler, 2007). Alla de ovannämnda metoderna är klassiska mått på investerarsentiment, men de senaste åren har fler kreativa metoder börjat användas.

Liknande studier inom ämnesområdet har ofta tagit avsats i att undersöka negativt laddade ord eftersom tidigare forskning tyder på att irrationella investerare agerar mer på negativa nyheter. Från Wall Street Journals dagligt uppdaterade artikelserie "Abreast of the Market" analyserar Tetlock (2007) interaktionen mellan media och aktiemarknaden. Tetlock uppmärksammar då hur en pessimistisk medial kommunikation medför en negativ aktiekursutveckling innan aktiepriserna därefter återhämtar sig mot dess fundamentala värde. Med utgångspunkt i att investerare är riskaversiva kan han därför påvisa att graden av pessimism i "Abreast of the Market" är representativ för negativt investerarsentiment. Vad som undersökningen därtill presenterar är att i de fall som det förekommer ovanligt stor eller ovanligt liten mängd negativa ord, kan aktiemarknaden vänta sig en högre handelsvolym. Likväl uppmärksammas hur storleken på aktiebolaget har avgörande effekt, där negativt sentiment tenderar att påverka mindre aktier i högre grad än större aktier.

2.2.2. Google Trender och aktieavkastning

Ett sätt att mäta investerarsentiment är genom sökaktivitet på nätet med hjälp av Googles sökvolymindex, *GSVI*, som presenteras på hemsidan Google Trender. I artikeln “The Sum of All FEARS Investor Sentiment and Asset Prices” (Da, Engelberg & Gao, 2015) mäter författarna investerarsentiment genom att undersöka förändringar i sökintresset på Google. Orden som de använder i sin studie är värdeladdade ord relaterade till ekonomi, till exempel “*gold prices*”, “*crisis*”, “*bankruptcy*” och “*frugal*”. Författarna skapar ett så kallat FEARS-index med negativt laddade sökord från Google. Höga indexvärden indikerar då att användare söker på de negativa orden i större utsträckning än vanligt. De finner bland annat ett prisvändningsmönster, då höga indexvärden förutspår nedgångar på aktiemarknaden de nästkommande dagarna och en uppgång därefter. Detta är ett viktigt resultat eftersom författarna inte bara hittar ett samtida samband mellan FEARS-indexet och aktieavkastningen, utan FEARS-indexet har även en prediktionsförmåga. Ett prisvändningsmönster är även centralt eftersom det bekräftar de beteendekonomiska teorierna om att priser tillfälligt kan avvika från fundamentala värden som en konsekvens av förändringar i investerarsentiment. Det finns sedan tidigare inget FEARS-index för den svenska aktiemarknaden.

Fördelarna med att använda sökaktivitet på nätet för att mäta investerarsentiment är många jämfört med andra metoder. För det första är datamängden stor, snabbt tillgänglig och har en hög frekvens (Beer, Herve & Zouaoui, 2013; Da, Engelberg & Gao, 2015). Sökaktiviteten på Google går nämligen att följa ända ner på minutbasis (Google, 2021). Data från Google Trender blir därför mer aktuell än kvalitativa undersökningar som bygger på enkätsvar från enstaka tusen utfrågade (Beer, Herve & Zouaoui, 2013; Da, Engelberg & Gao, 2015). I Sverige finns exempelvis Konjunkturinstitutet som försöker mäta svenskarnas inställning till ekonomin, men deras rapporter är bara tillgängliga på månatlig basis (Konjunkturinstitutet, 2021). Da, Engelberg och Gao (2015) visar till och med att data från Google Trender är så aktuell att deras FEARS-index kan förutse resultaten av dessa månatliga enkätundersökningar. För det andra *avslöjar* sökintresset investerarnas sinnesstämning, till skillnad från undersökningar som uttryckligen *frågar* efter den. Da, Engelberg och Gao (2015) argumenterar att det kan vara känsligt att svara sanningsenligt på frågor som är negativt laddade, men på nätet söker personer utan hämningar. Slutligen fångar sökdata intresset från icke-professionella investerare, vilket blir särskilt passande för ämnesområdet

beteendeekonomi eftersom man vill undersöka investerare vars beslut påverkas av investerarens sinnesstämning (Beer, Herve & Zouaoui, 2013; Da, Engelberg & Gao, 2011).

2.2.3. Google Trender, volatilitet och handelsvolym

Förutom att mäta investerarnas sinnesstämning har Google Trender även använts inom beteendeekonomi för att mäta investerarnas uppmärksamhet (Aouadi, Arouri & Teulon, 2013; Joseph, Wintoki & Zhang, 2011; Kim, Lučivjanská, Molnár & Villa, 2019; Takeda & Wakao, 2014). GSVI fungerar då som ett mått på vad investerare uppmärksammar för tillfället. Höga GSVI-värden innebär att investeraren har lagt märke till det som söks efter. Tidigare studier inom beteendeekonomi har då undersökt hur investerare uppmärksammar olika tillgångar på aktiemarknaden. Överlag visar det sig att höga GSVI-värden på tickersymboler och företagsnamn kan förutse ovanligt stor avkastning för aktier som söks på Google. Dessa resultaten håller på fler marknader än den amerikanska, och volatila aktier är mer känsliga för förändringar i sökintresse (Aouadi, Arouri & Teulon, 2013; Joseph, Wintoki & Zhang, 2011; Kim, Lučivjanská, Molnár & Villa, 2019; Takeda & Wakao, 2014). Även Da, Engelberg och Gao (2011) använder Google Trender som ett verktyg för investerares uppmärksamhet. Förutom en högre avkastning de kommande två veckorna finner de att en hög uppmärksamhet för en tickersymbol förutspår en högre förstadaysavkastning för börsnoteringar. Börsnoteringar som föregås av en hög uppmärksamhet på Google Trender predikterar även en lägre långsiktig aktieutveckling.

Ett vanligt förekommande resultat i forskningen som använder Google Trender för att mäta investerarsentiment och investerares uppmärksamhet, är att sökintresset även kan förutspå annan aktivitet på aktiemarknaden. Höga GSVI-värden på tickersymboler och företagsnamn predikterar både handelsvolym och volatilitet de kommande dagarna. En hög sökaktivitet leder alltså till att handelsaktivitet kring företagen som det söks efter ökar (Aouadi, Arouri & Teulon, 2013; Joseph, Wintoki & Zhang, 2011; Kim et al., 2019; Takeda & Wakao, 2014). Detta ger belägg för teorin om att uppmärksamhet är en begränsad resurs som avgör vilka aktier som kommer handlas. En investerare, och framförallt icke-professionella investerare, kan inte uppmärksamma alla aktier på marknaden och därför handlas endast de aktier som investeraren har lagt märke till (Barber & Odean, 2008). Återigen håller resultaten om ökad volatilitet och handelsvolym på fler marknader än den amerikanska, och aktier med hög

volatilitet är särskilt känsliga för dessa förändringar i sökintresse (Aouadi, Arouri & Teulon, 2013; Joseph, Wintoki & Zhang, 2011; Kim et al., 2019; Takeda & Wakao, 2014).

2.3. Sammanställning av teorier och tidigare forskning

I tabell 1 återfinns en sammanställning av de teorier och studier som presenterats i detta avsnittet. Där lyfter vi även de viktigaste forskningsresultaten som kommer behandlas vidare i denna uppsatsen.

Tabell 1. Sammanställning av teorier, tidigare forskning och resultat som tillämpas i denna uppsatsen.

Artiklar	Ämnesområde	Resultat
Fama, 1970; Malkiel, 2003; Shaker, 2013	Effektiva marknadshypotesen.	På en effektiv marknad återspeglar priserna all tillgänglig information. Det finns anledning att misstänka att reglerade och oreglerade svenska marknader inte är lika effektiva. Detta ger utrymme för beteendekonomi.
Baker & Wurgler, 2006; Baker & Wurgler, 2007; Barberis & Thaler, 2003; Brown & Cliff, 2005; De Long et al., 1990	Beteendekonomi och investerarsentiment.	Investerarsentiment kan orsaka irrationellt handlande. Priser kan därmed avvika från rationellt grundade värden på grund av investerarens sinnesstämning.
Barberis & Thaler, 2003; Baumeister et al., 2001	Positivt och negativt sentiment.	Negativt laddade ord har en starkare effekt än positiva ord. I en ekonomisk kontext är detta förknippat med dispositionseffekten.
Barberis & Thaler, 2003; Barber, Odean & Zhu, 2009; Foucault, Sraer & Thesmar, 2011	Vilka investerare påverkas av sin sinnesstämning?	Spekulativa investerare påverkas av sin sinnesstämning. Dessa är ofta icke-professionella investerare.
Baker & Wurgler, 2007; Tetlock, 2007	Mått på investerarsentiment.	Nyhetskolumnen "Abreast of the market" och andra klassiska mått kan mäta investerarsentiment.
Beer, Herve & Zouaoui, 2013; Da, Engelberg & Gao, 2011; Da, Engelberg och Gao, 2015	Google Trender och aktieavkastning.	GSVI kan användas som mått på investerarsentiment och kan prediktera aktiekursförändringar med ett FEARS-index. Sökaktivitet på Google är ett tillgängligt och aktuellt mått på investerares sentiment och uppmärksamhet. Sökaktiviteten fångar främst beteende från icke-professionella investerare.
Aouadi, Aroui & Teulon, 2013; Beer, Herve & Zouaoui, 2013; Da, Engelberg & Gao, 2011; Joseph, Wintoki & Zhang, 2011; Kim et al., 2019; Takeda & Wakao, 2014	Google Trender, volatilitet och handelsvolym.	GSVI kan prediktera förändringar i volatilitet och handelsvolym. Små företag med hög volatilitet är särskilt känsliga för förändringar i investerarsentiment. Resultaten håller på fler marknader än den amerikanska.

2.4. Litteraturkritik

Vald litteratur och tidigare studier har bedömts med utgångspunkt i källans transparens i beskrivning av metoder, aktualitet och objektivitet. Kvalitetskontroll av tidskrifter och dess auktoritetshierarki har gjorts med hjälp av SCImago Journal Ranking (SJR) som rankar efter auktoritet inom olika forskningsfält (SCImago, n.d.). Auktoritet bestäms med hänsyn till tidsskriftens påverkan och prestige genom att tjänsten beräknar ett genomsnitt av antalet citationer, där referenser i framstående tidsskrifter viktas tyngre. Litteratur av Fama (1965; 1970) inom ämnet effektiva marknadshypotesen samt Brown och Cliffs (2005) bidrag inom beteendekonomi anses, trots sin ålder, som auktoritära inom ämnet tack vare dess påverkan på fältet och tidskrifternas position. "The Sum of all FEARS Investor Sentiment and Asset Prices" av Da, Engelberg och Gao (2015) utgör en central del arbetets metod. Artikeln är relativt modern, vilket är viktigt då metoden som används kan påverkas av faktorer såsom Googles uppdateringar av sina tjänster, vilket vi bedömer inte har genomgått betydelsefulla förändringar sen publiceringen av deras forskning.

Artikeln av Shaker (2013) är publicerad i en tidskrift som inte är rankad av SJR, men har å andra sidan genomgått en kollegial utvärdering (European Journal of Business and Social Sciences, n.d.). Artikeln föranleder enbart att ämnet är intressant att studera vidare med tanke på att författaren indikerar att OMXS30 inte är en effektiv marknad.

I litteraturen om investerarsentiment finns det ingen allmänt vedertagen metod för att mäta investerarsentiment. Olika mått har använts i olika studier. De studier som har skapat FEARS-indexet är inte heller ense om FEARS-indexet ska baseras på enbart negativa ord, enbart positiva eller en blandning av de två. Detta är både en fördel och en nackdel inom forskningsområdet. Å ena sidan verkar det inte vara fastställt vilken metod som är bäst. Å andra sidan är det en styrka att resultaten i stora drag är överensstämmande, även fast olika studier har använt sig av olika tillvägagångssätt. Detta är något som ökar både forskningsområdets validitet och reliabilitet.

En vanlig frågeställning i forskning som har använt sig av sökintresse på Google är hur förändringar i investerarsentiment påverkar handelsvolym, volatilitet och avkastning. Detta är en av de frågor som dessutom har undersökts på flera marknader internationellt och med

blandade resultat. Vissa studier kan visa ett signifikant samband (Aouadi, Arouri & Teulon, 2013; Joseph, Wintoki & Zhang, 2011) medan andra inte kan det (Kim et al., 2019; Takeda & Wakao, 2014). Några ser en skillnad mellan de olika typerna av företag som undersöks, och andra gör inte det. Här har forskningen en svårighet med att förklara vilka faktorer som gör att det ibland finns ett signifikant samband. Detta är en klar nackdel med forskningen om investerarsentiment. Ibland finner forskningen ett signifikant resultat, men kan då bara teoretisera kring vad det är som föranlett det. Mycket av det beror på att forskningen om investerarsentiment försöker kvantifiera något som är kvalitativt.

2.5. Utvecklande av hypoteser

Efter att ha redogjort för den teoretiska referensramen kan vi formulera välgrundade hypoteser som ämnar att besvara studiens syfte. Tidigare forskning, och kritiken mot denna, fungerar som stöd för att identifiera vilka aspekter av investerarsentiment som är intressanta att undersöka på den svenska aktiemarknaden.

2.5.1. FEARS-indexets kraft på olika marknader

Som nämnt är spekulativa aktier, det vill säga små företag med hög volatilitet, särskilt känsliga för förändringar i investerarnas uppmärksamhet (Baker & Wurgler, 2006; Baker & Wurgler, 2007; Beer, Herve & Zouaoui, 2013). Ett viktigt resultat i studien av Baker och Wurgler (2006) är just att företagets karaktär är centralt för att avgöra hur stor påverkan investerarsentimentet har för aktiekursutvecklingen. Av denna anledningen blir det intressant att skilja på en reglerad och en oreglerad aktiemarknad och jämföra om investerarsentiment påverkar företagen på respektive marknad olika mycket. Som bekant har företagen på First North Sweden annorlunda egenskaper jämfört med de större och mer etablerade företagen på Stockholmsbörsen. Det är därför rimligt att anta att investerarnas sinnesstämning påverkar aktiekursutvecklingen olika mycket på de två marknaderna. Vidare visar statistiken från SCB (2021a) i figur 1 att de svenska hushållen har en större andel av sitt aktieinnehav på oreglerade marknader jämfört med andra samhällssektorer. Vi vet med oss att denna typen av investerare tenderar att vara mer spekulativa, påverkas mer av sin sinnesstämning och ta mindre informerade beslut. Dessutom vet vi sedan tidigare att Googles sökvolymindex fångar sökintresset hos just dessa oinformerade investerare. Mot bakgrund av detta kommer alla kommande hypoteser prövas på både First North Sweden och på Stockholmsbörsen.

2.5.2. Negativt och positivt sentiment

Av de studier som har skapat FEARS-index på olika marknader råder det ingen konsensus om FEARS-indexet borde innehålla enbart negativt laddade ord, enbart positivt laddade ord, eller både och. Da, Engelberg och Gao (2015) använde bara negativt laddade ord eftersom de fann att bland de ord som hade starkast samband med marknadsavkastningen var sambandet oftast negativt. En förklaring till detta kan vara att tidsperioden som författarna undersökte, 2004 till 2011, bland annat innefattade finanskrisen. Denna period kan mycket väl tänkas skapa oroligheter bland investerare, och skulle således motivera varför negativt sentiment får ett större inflytande på marknadsavkastningen. Författarnas tidsperiod innehöll både uppgångar, nedgångar och stagnationer på aktiemarknaden, till skillnad från vår tidsperiod som till övervägande del karaktäriseras av uppgångar. Dessutom finns det en asymmetri mellan olika marknadsfaser, där fallande aktiekurser har en större påverkan på framtida volatilitet än stigande (Francq & Zakoian, 2010).

Brochado (2020) skapar ett FEARS-index på den portugisiska aktiemarknaden och undersöker perioden 2009 till 2014. Den distinkta skillnaden mellan denna tidsperioden jämfört med den från Da, Engelberg och Gao (2015) är att Brochado (2020) undersöker en tidsperiod med tydliga uppgångar på aktiemarknaden. Brochado (2020) skapar ett FEARS-index med positivt laddade ord och ett FEARS-index med negativt laddade ord. Författaren finner därefter att det positiva FEARS-indexet har en starkare inverkan på aktiemarknadsutvecklingen än det negativa indexet, både när det gäller prediktion av avkastning och handelsvolym. Dock är både det negativa och det positiva indexet signifikanta i sina prediktioner.

Studierna antyder att negativt sentiment och positivt sentiment på aktiemarknader kan variera beroende på vilken tidsperiod som undersöks, och detta får konsekvenser för senare resultat. Därför är det viktigt att fastställa vilken sorts sentiment som är mest framträdande i den perioden och den geografiska regionen som undersöks. I vårt fall motiverar detta varför det är intressant att undersöka om det är någon skillnad mellan positivt laddade ord och negativt laddade ord i konstruktionen av FEARS-indexet. Dessa argument leder oss till följande hypoteser:

- H1a: Det finns ingen skillnad på positiva och negativa ord i investerarsentiment på Stockholmsbörsen.
- H1b: Det finns ingen skillnad på positiva och negativa ord i investerarsentiment på First North Sweden.
- H1c: Det finns ingen skillnad på positiva ord i investerarsentiment på Stockholmsbörsen jämfört med på First North Sweden.
- H1d: Det finns ingen skillnad på negativa ord i investerarsentiment på Stockholmsbörsen jämfört med på First North Sweden.

2.5.3. Investor sentiment med Google Trender

Av tidigare avsnitt att döma finns det flera fördelar med *GSVI* som mått på investerarsentiment, jämfört med exempelvis enkäter eller ett VIX-index. I kontexten av den här uppsatsen finns det ytterligare en nackdel med dessa mått. De mäter sentiment på hela marknaden i stort, de skiljer inte på reglerade och oreglerade marknader. Ett FEARS-index går däremot att anpassa utefter exakt vad man vill undersöka. Detta motiverar varför det är intressant att ta reda på om det går att konstruera ett FEARS-index för den svenska aktiemarknaden. Det hade varit användbart med ett FEARS-index som förutspår aktiekursförändringar även på den svenska marknaden och inte bara på den amerikanska eller den portugisiska. FEARS-indexet är enkelt att framställa, billigt, aktuellt, anpassningsbart och försvagas inte av oärliga enkätsvar. FEARS-indexen har dessutom en bevisad prediktionsförmåga. Detta gör att FEARS-indexen faktiskt blir användbara, till skillnad från om det bara skulle finnas ett samtida samband med aktieavkastningen. Således konstruerar vi följande hypoteser:

- H2a: FEARS-index påverkar inte aktiekursförändringar på Stockholmsbörsen.
- H2b: FEARS-index påverkar inte aktiekursförändringar på First North Sweden.
- H2c: FEARS-index påverkan på aktiekursförändringar skiljer sig inte mellan Stockholmsbörsen och First North Sweden.

En skillnad mellan *GSVI* som ett mått på investerares uppmärksamhet och investerares sinnesstämning är att sökorden i det senare fallet är värdeladdade. I övrigt kan *GSVI* för investerarsentiment ses som en förlängning av *GSVI* för investerares uppmärksamhet eftersom investeraren först måste uppmärksamma något innan denne kan skapa sig en attityd

gentemot det. Detta gör att resultaten inom forskningen om investerares uppmärksamhet kan tänkas vara användbara även i studier om investerarsentiment. FEARS-indexet skulle då inte bara påverka aktiekursförändringar, utan även förändringar i volatilitet och handelsvolym. Att pröva FEARS-indexets relation till andra variabler än aktieavkastningen blir dessutom ett test för att se om FEARS-indexet fungerar och faktiskt mäter investerarsentiment. Vi kommer senare förklara att FEARS-indexet är skapat genom att utgå från de ord som har starkast samband med aktieavkastning. Om FEARS-indexet fortfarande kan påverka andra variabler än aktieavkastning blir det därmed ett tecken på att indexet är robust och mäter det som det avser att mäta. Av den orsaken bygger vi följande hypoteser om volatilitet:

- H3a: FEARS-index påverkar inte förändringar i volatilitet på Stockholmsbörsen.
- H3b: FEARS-index påverkar inte förändringar i volatilitet på First North Sweden.
- H3c: FEARS-index påverkan på volatilitet skiljer sig inte mellan Stockholmsbörsen och First North Sweden.

Vi konstruerar även följande hypoteser om handelsvolym:

- H4a: FEARS-index påverkar inte förändringar i handelsvolym på Stockholmsbörsen.
- H4b: FEARS-index påverkar inte förändringar i handelsvolym på First North Sweden.
- H4c: FEARS-index påverkan på handelsvolym skiljer sig inte mellan Stockholmsbörsen och First North Sweden.

3. METOD OCH DATAINSAMLING

I detta avsnittet redogörs och motiveras studiens tillvägagångssätt. Insamlingen och behandlingen av data från Google Trender, samt konstruktionen av FEARS-indexet, följer till stora delar Da, Engelberg och Gao (2015) som skapade det ursprungliga FEARS-indexet. Flera delar av metodiken kompletteras även med kunskap från tidigare liknande studier som antingen använt sig av Google Trender eller som har undersökt investerarsentiment.

3.1. Val av metod

Studien tar en deduktiv ansats genom att utvald teori härleder till fyra hypoteser och dess begrepp översätts till operationella termer (Bryman & Bell, 2017). Tidsserieanalys används med syftet att undersöka positiva och negativa ord i det konstruerade FEARS-indexet och dess samband med avkastning, volatilitet och handelsvolym på reglerade och oreglerade svenska aktiemarknader över tid.

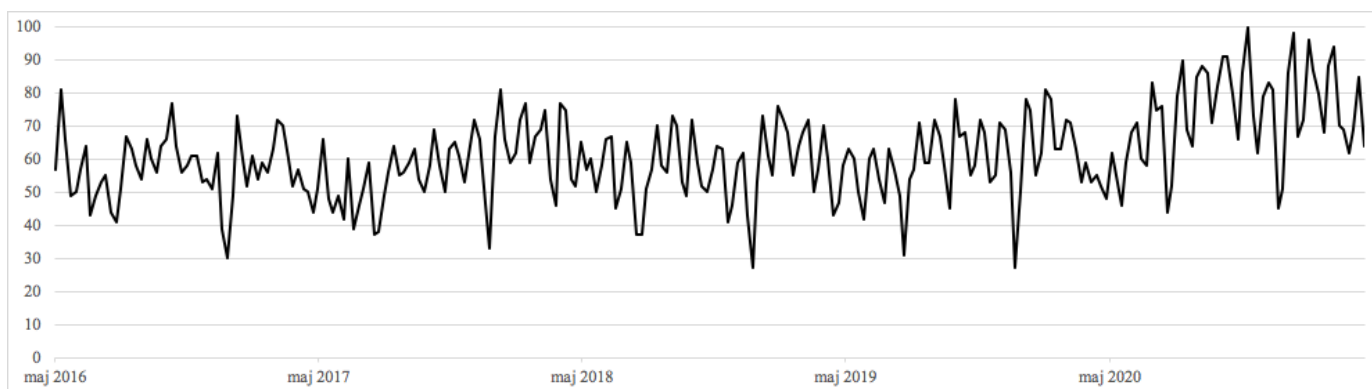
Tidsserieanalys ämnar att undersöka sambandet mellan variabler som mäts vid flera tidpunkter över en given tidsperiod (Dougherty, 2016). En viktig egenskap med tidsserieanalys är att ordningen på observationerna spelar roll. Tidsserieanalys utgår därför från andra antaganden och kräver andra metoder än exempelvis analyser av tvärsnittsdata. Med tidsseriedata är det exempelvis inte längre lämpligt att anta att observationerna är en sekvens av oberoende och slumpmässiga variabler dragna från samma fördelning. Två tidpunkter i en tidsserie tenderar istället att vara beroende av varandra (Dougherty, 2016).

3.2. Datainsamling och databehandling

Listan av ord som utgör FEARS-indexet baseras på översatta ord från Harvard IV-4 Dictionary och Lasswell Value Dictionary (General Inquirer, n.d.) och dess relaterade ord i Google Trender (Google, n.d.). FEARS-indexets konstruktion och datainsamling behandlas i mer detalj senare i samma avsnitt. Prisindex och handelsvolym på Nasdaq OMX Stockholm och First North Sweden har hämtats genom Bloombergs databas. Kontrollvariabler är hämtade från Konjunkturinstitutet, Statistiska centralbyrån och Riksbanken.

3.2.1. Google sökvolymsindex

“Google Trender” är ett verktyg som låter användaren se och ladda ner söktrenden för ett visst sökord på Google. Användaren skriver in ett sökord, en tidsperiod och en geografisk region och Google Trender visar därefter ordets sökintresse. Sökintresset uttrycks som ett indexvärde mellan 0 och 100, i vad som kallas GSVI. Sökordet kan vara både ett ord och en längre fras (Google, 2021). Figur 2 visar sökintresset för ordet “Arbetsförmedlingen” i Sverige mellan den 1a maj 2016 och den 18e april 2021.



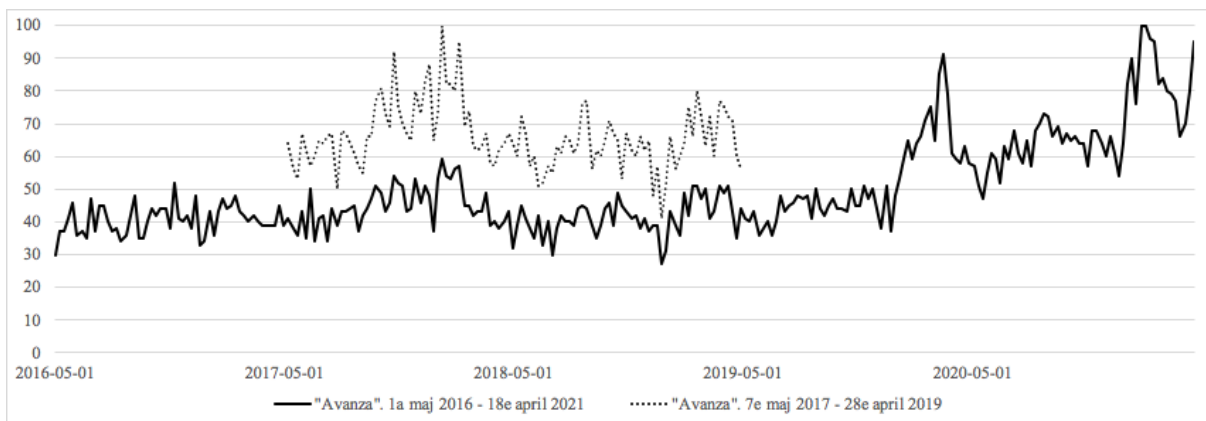
Figur 2. Sökintresset för “Arbetsförmedlingen” i Sverige mellan den 1a maj 2016 och den 18e april 2021. De lägsta punkterna visar ett tydligt mönster att intresset för Arbetsförmedlingen är lågt över julhelger och i juli månad.

För en given tidpunkt, t , och ett givet sökord, k , beräknas $GSVI_{k,t}$ genom att dividera antalet sökningar på sökordet med det totala antalet sökningar inom den givna tidsperioden och geografiska området. Därefter divideras de relativa sökfrekvenserna inom tidsperioden med den högsta relativa sökfrekvensen för tidsperioden. Slutligen multipliceras kvoten med 100 för att beräkna $GSVI_{k,t}$. Exempelvis, om det har gjorts 100 sökningar på ett givet sökord i tidpunkt t , och det har gjorts totalt 1000 sökningar över hela tidsperioden, blir den relativa sökfrekvensen vid tidpunkt t 0,1. Om 0,4 är den högsta relativa sökfrekvensen för någon tidpunkt i hela tidsperioden, kommer indexvärdet $GSVI_{k,t}$ i tidpunkt t bli 0,25 multiplicerat med 100, det vill säga 25.

För att förbättra svarstiden använder Google Trender bara en delmängd av de faktiska sökningarna i dessa beräkningar (Google, 2021). Vilken delmängd som används ändras vid

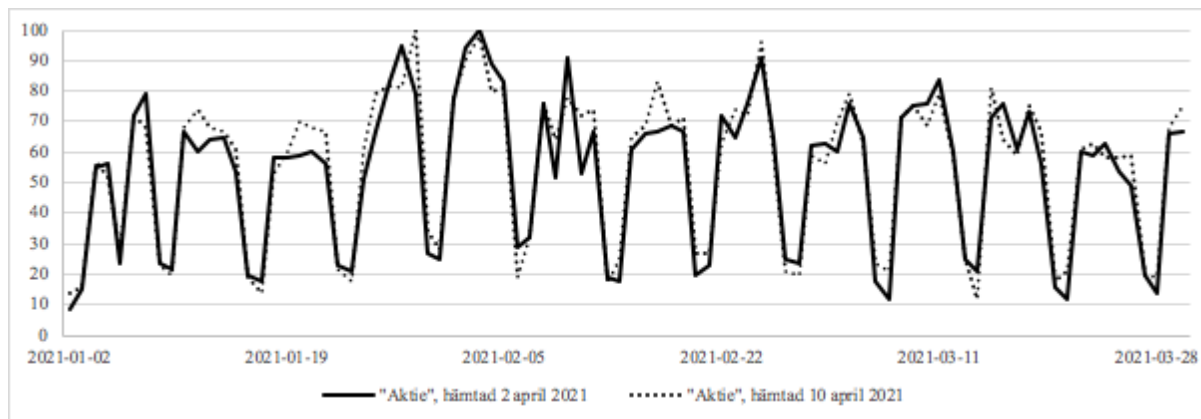
jämna mellanrum. Google skriver på sin hjälpsida att delmängden är ett representativt urval av de faktiska sökningarna, men berättar inte hur urvalet görs.

Beräkningsmetoden och dataurvalet får två konsekvenser som måste beaktas i vår studie. För det första beror $GSVI_{k,t}$ på den valda tidsperioden. Detta gör att $GSVI_{k,t}$ för en given tidpunkt, t , kan anta olika värden beroende på vilken tidsperiod som studeras och därför måste tidsperioden som studeras väljas med omtanke. I vårt fall exkluderas tidsperioden från år 2020 och framåt då både börsdata och sökintresset på Google kan präglas av covid-19-pandemin. Eftersom $GSVI_{k,t}$ -värdet för en given tidpunkt beror på hela tidsperioden påverkar sökintressen under 2020 även indexvärden tidigare i perioden. Om vi hade inkluderat 2020 i tidsperioden hade vi antagligen fått sökintressen som inte är generaliserbara utanför en krissituation. Dessa hade i sin tur påverkat indexvärdena för resten av tidsperioden eftersom sökintresset för en given tidpunkt sätts i relation till sökintresset över hela tidsperioden. Man får alltså inte bara exceptionella värden under 2020, utan detta riskerar även att färga av sig på indexvärdena för resten av perioden. Denna förutsättningen visualiseras i figur 3.



Figur 3. $GSVI_{k,t}$ för sökordet “Avanza” i Sverige, hämtat över två olika tidsperioder. I den överlappande tidsperioden skiljer sig indexvärdena åt mellan de två serierna. Detta beror på att den längre tidsperioden inkluderar år 2020 och framåt, vilket är en period då sökintresset för sökordet “Avanza” var ovanligt stort som en konsekvens av covid-19-pandemin. Detta gör perioderna mindre jämförbara och skapar stora problem med heteroskedasticitet, brytpunkter i tidsserien och icke-stationäritet.

För det andra kan $GSVI_{k,t}$ anta olika värden beroende på när data hämtas, vilket måste beaktas vid datainsamlingen. Denna förutsättningen visualiseras i figur 4 och behandlas i nästkommande avsnitt.



Figur 4. $GSVI_{k,t}$ för sökordet “aktie” från den 2a januari 2021 till den 30e mars 2021. De två tidsserierna är hämtade på två olika dagar. Korrelationen mellan de två serierna är cirka 96 %. Serierna avviker från varandra framförallt kring extremvärden, vilket motiverar varför datan bör justeras för extremvärden.

3.2.2. Databehandling av Google sökvolymindex

Eftersom $GSVI_{k,t}$ kan anta olika värden vid olika nedladdningar, d , följer vi Preis, Moat och Stanley (2013) och laddar ner data flera dagar i rad för alla våra sökord. Vi gör totalt 8 nedladdningar, D , mellan den 6 april och den 13 april. Därefter räknar vi ut det genomsnittliga indexvärdet över alla nedladdningar för ett givet sökord och för en given tidpunkt. Resultatet betecknas $SVI_{k,t}$ och är härnäst den huvudsakliga variabeln som kommande beräkningar utgår ifrån:

$$SVI_{k,t} = \frac{1}{D} \cdot \sum_{d=1}^D GSVI_{k,t,d} \quad (1)$$

Härnäst beräknar vi förändringen i den naturliga logaritmen av $SVI_{k,t}$:

$$\Delta SVI_{k,t} = \ln\left(\frac{SVI_{k,t}}{SVI_{k,t-1}}\right) . \quad (2)$$

Med denna omvandling räknar vi ut förändringar i sökintresset för ett givet sökord. En förutsättning för att tidsseriedata ska vara jämförbar är att den är stationär (Cryer & Chan, 2008). En monoton transformation med hjälp av den naturliga logaritmen ändrar skalan för sökintresset och gör att sambandet blir linjärt och tidsserien stationär då variansen av $SVI_{k,t}$ inte längre beror på t . Detta hade annars varit ett problem om sökaktiviteten på Google ökar över tiden. Vi behandlas stationäritet i mer detalj senare i avsnittet.

För att justera för extremvärden winsoriserar vi varje tidsserie med 2,5 % i varje ände. Winsoriseringen begränsar extremvärdenas inflytande genom att reducera värden över den 97,5:e percentilen till värdet av den 97,5:e percentilen och genom att öka värden under den 2,5:e percentilen till värdet av den 2,5:e percentilen. Anledningen till varför vi använder oss av en winsorisering istället för att exkludera extremvärden är för att vi vill bevara höga värden i datasetet men samtidigt förhålla oss kritiska till att extremvärden kan vara anomalier. Extremvärdena kan vara anomalier eftersom Googles sökvolymsindex inte är baserat på alla faktiska sökningar utan endast ett ospecificerat urval av populationen.

Sökintresset på Google kan följa en naturlig säsongsvariation. Figur 4 visar en tydlig säsongsvariation för sökordet "aktie". Tolkningen är intuitiv: sökintresset är lågt på helgerna när aktiehandeln är stängd. Om detta inte beaktas kommer datan förvrängas av de stora förändringarna i sökintresset mellan vardagar och helger. Förändringar i sökintresset för ett visst ord blir i sådana fall inte ett mått på investerarsentiment utan snarare ett mått på när det är vardag. Även om vi använder oss av veckodata istället för daglig data som i figur 4 kan det fortfarande finnas säsongsvariationer som beror på något annat än förändringar i investerarnas sentiment. Därför behöver tidsserierna säsongjusteras, vilket vi förklarar härnäst.

För att få en bättre förståelse för de faktorer som förklarar tidsseriens utveckling, har denna brutits ned. Tidsserien kan ses som en funktion av tre delkomponenter: trend, säsongsvariation och residualer. Tidsseriedatas *trend* beräknas utifrån glidande medelvärden

och visar den utjämnade utvecklingen, där fluktuationer på medel- och hög frekvens har filtrerats bort (Athanasopoulos & Hyndman, 2018). På så vis tar denna inte hänsyn till kortsiktiga upp- och nedgångar utan ger en överblickande förklaring till utvecklingen över längre tid (OECD, 2005a). *Säsongskomponenten*, i sin tur, representerar den variation som är mer eller mindre återkommande med avseende på timing, riktning och magnitud på årlig basis (OECD, 2005b). Denna beräknas genom att ta genomsnittet för varje tidsenhet över samtliga perioder innan den slutligen centreras. Det skapas således dummyvariabler för olika dagar, veckor och månader i tidsserien (Athanasopoulos & Hyndman, 2018). *Residualen* är det som återstår efter att trend- och säsongskomponenter har exkluderats från den observerade tidsserien. Residualkomponenten består alltså endast av oregelbunden variation som uppkommer från utomstående, icke-systematiska faktorer (OECD, 2005c). En observation i en tidsserie kan alltså ses som en funktion av tre komponenter: ett glidande medelvärde, dummyvariabler beroende på vilken månad, vecka och dag som observationen tillhör, samt residualen. Vid en undersökning huruvida ett FEARS-index påverkar förändringar i tidsseriedata från OMXS och First North Sverige, kommer endast residualerna undersökas vidare. Residualerna representerar sporadiska impulser och humörsvängningar, och inte systematiska sökningar på Google. Det är värt att notera att när trendkomponenten exkluderas från den observerade tidsserien försvinner observationer i början och slutet av tidsserien över residualer. Detta eftersom tidsseriens trend är ett glidande medelvärde som är baserad på de sex månaderna innan och efter tidpunkt t .

Avslutningsvis normaliserar vi alla tidsserier till medelvärdet 0 och standardavvikelsen 1 genom att dividera varje tidsserie med dess standardavvikelse, detta enligt Da, Engelberg och Gao (2015) samt Baker och Wurgler (2006). Normaliseringen är nödvändig eftersom tidsserierna för varje sökord senare kommer användas i konstruktionen av FEARS-indexet och behöver då följa samma skala.

3.2.3. Tidsperiod och geografisk region

Vi har valt att undersöka tidsperioden från den 1a januari 2015 till den 31 december 2019. Med denna perioden får vi så aktuell data som möjligt. Vi har medvetet undvikit tidsperioden från år 2020 och framåt då aktiemarknaden under denna perioden präglas av en större osäkerhet och ovanligt stora fluktuationer som en konsekvens av covid-19-pandemins utbrott. Det går att resonera för att investerare kan tänkas agera och påverkas annorlunda av

sentiment under en krissituation och därför inkluderar vi inte senare data. Resultaten av vår studie har därför förhoppningen att vara mer applicerbara under mer “normala” omständigheter. Avslutningsvis är vår valda tidsperiod inte längre än fem år eftersom dataupplösningen från Google Trender övergår från veckodata till månatlig data om tidsperioden är längre än fem år. Vi anser att månatlig data har för låg datafrekvens för att mäta förändringar i investerarsentiment. Med den valda tidsperioden får vi 261 observationer per tidsserie. Detta antal minskar först till 260 eftersom vi förlorar den första observationen i tidsserien när vi beräknar förändringen i indexvärde enligt ekvation 2. Senare minskar antalet observationer till 193 eftersom säsongjusteringen bland annat är baserad på glidande medelvärden. Ett glidande medelvärde baserat på ett års observationer innebär således att observationerna de första och sista sex månaderna kommer att försvinna.

Vi har därtill valt att avgränsa den geografiska regionen på Google Trender till Sverige eftersom vi vill undersöka den svenska aktiemarknaden.

3.2.4. Ordlistan

För att kunna använda oss av Google Trender behöver vi veta vilka ord vi ska hämta sökdata från. Vårt tillvägagångssätt för att bygga en ordlista följer till stor del Da, Engelberg och Gao (2015) och Brochado (2020). Vi utgår från Harvard IV-4 Dictionary och Lasswell Value Dictionary enligt Tetlock (2007) samt Tetlock, Saar-Tschansky och Macskassy (2008). Detta lexikon delar in engelska ord och begrepp i olika kategorier beroende på vilket ämne de tillhör och hur värdeladdade de är. Vi filtrerar lexikonet på ord som tillhör en ekonomi-kategori och som antingen är positiva, negativa, uttrycker styrka eller svaghet (“Econ@”, “ECON”, “Exch”, “positiv”, “negativ”, “pstv”, “ngtv”, “strong” respektive “weak”). Detta ger oss 278 ord att utgå ifrån. Vi väljer dessa laddade kategorier då vi vill att orden ska mäta en icke-neutral sinnesstämning hos investerarna. Dessutom visar tidigare studier att ord kan få ett annat värde när de används i kontexten att bygga ett FEARS-index. Da, Engelberg och Gao (2015) finner exempelvis ett negativt samband mellan sökintresset för ordet “gold” och marknadsavkastningen, samtidigt som lexikonet kategoriserar ordet som positivt laddat.

I nästa steg översätter vi de engelska orden till svenska med hjälp av Nationalencyklopedins engelska ordbok (NE, n.d.). Vi inkluderar alla synonyma översättningar som kan tänkas vara relaterade till ekonomi. Efter detta steget har vi 435 ord.

Därefter för vi in varje ord i Google Trender och hämtar alla ord som är relaterade till ekonomi bland de tio mest populära relaterade söktermerna. Detta gör vi för att utöka vår ordlista och för att täcka fler av de orden som sentimentala investerare använder. Ord som används i icke-ekonomiska sammanhang exkluderas från ordlistan. Om ett ord både kan användas i ett ekonomisk och ett icke-ekonomiskt sammanhang inkluderas ordet i ordlistan. Till exempel är ordet "fattig" med i ordlistan, även om det framgår av de relaterade sökningarna att ordet i vissa sammanhang används för att söka efter Astrid Lindgrens låt "Fattig bonddräng". Ursprungsordet "fattig" är alltså i en gråzon eftersom ordet både kan ha en ekonomisk och en icke-ekonomisk innebörd. Det finns anledning att tro att ord som har en övervägande icke-ekonomisk innebörd kommer att filtreras bort. I avsnitt 3.7. kommer denna felkälla utvecklas vidare.

Efter att ha rensat bort dubletter innehåller vår ordlista 777 sökord och fraser. Vi laddar ner $GSVI_{k,t}$ för alla ord i denna ordlistan en gång om dagen mellan den 6e april och den 13e april. Om det inte går att ladda ner $GSVI_{k,t}$ för ett sökord på grund av att Google inte har tillräckligt med data exkluderas ordet från ordlistan. Om fler än 10 observationer i en tidsserie antar värdet 0 exkluderas ordet från ordlistan eftersom låga sökvolymen visas som 0 på Google Trender (Google, 2021). Vår slutgiltiga ordlista, hädanefter "baslistan", innehåller 180 ord och fraser. Baslistan återfinns i tabell A.1 i appendix 1.

3.2.5. Konstruktion av FEARS-indexet

Av alla ord och fraser i baslistan vill vi bara inkludera de som bäst mäter investerarsentiment. För att identifiera dessa räknar vi ut korrelationen mellan varje sökord och marknadsavkastningen på OMXS respektive First North Sweden. För varje sökord görs alltså två skattningar av korrelationen baserad på historisk data, en med OMXS och en med First North Sweden. Observationerna för alla sökord och börsdata är därmed tidsserier. Från korrelationsanalysen får varje sökord ett t-värde som berättar hur signifikant den skattade korrelationskoefficienten är, baserat på standardavvikelsen och den skattade korrelationskoefficienten för varje tidsserie. Om t-värdet är negativt innebär det att sökordet

har en negativ korrelation med marknadsavkastningen och är således ett negativt laddat ord. Det omvända gäller för positiva t-värden. Fortsättningsvis är detta definitionen av vad som är ett negativt laddat sökord och vad som är ett positivt laddat sökord (Da, Engelberg & Gao, 2015). Sökorden rangordnas sedan utifrån största t-värde i absoluta tal, och därefter väljer vi de 30 sökord med störst t-värde för respektive marknadsplats. Vi rangordnar utifrån t-värden eftersom FEARS-indexet då baseras på de sökord som har det mest signifikanta sambandet med marknadsavkastningen. Vi väljer de 30 mest signifikanta orden eftersom vi vid detta antal kan anta att vi har diversifierat bort idiosynkratiskt brus (Da, Engelberg & Gao, 2015).

Slutligen formas ett FEARS-index genom att beräkna det genomsnittliga $\Delta SVI_{k,t}$ över alla 30 sökord, k , vid tidpunkt t . FEARS-indexets värde vid tidpunkt t blir alltså det genomsnittliga $\Delta SVI_{k,t}$ för de 30 mest signifikanta sökorden och indexet uttrycks som en tidsserie över vår valda tidsperiod:

$$FEARS_t = \frac{\sum_{k=1}^{30} \Delta SVI_{k,t}}{30}. \quad (3)$$

I slutändan har vi alltså två FEARS-index, ett för Stockholmsbörsen och ett för First North Sweden. Dessa betecknas $FEARS_{OMXS}$ och $FEARS_{FNSAS}$. I tabell A.2 i appendix 2 framgår det vilka sökord som ingår i respektive FEARS-index.

3.2.6. Nasdaq OMX Stockholm och First North Sweden

Genom Bloombergs databas hämtas stängningskurser och volym på aktiekursindexen OMXS och First North Sweden All-Share, vilket är aktiekursindexen för samtliga svenska aktier på OMX Stockholm respektive First North Sweden. Vi väljer att undersöka breda aktieindex eftersom vi vill jämföra reglerade och oreglerade marknader, och då vill vi inkludera alla företag som tillhör respektive marknad för att göra urvalet mer representativt. First North Sweden är den överlägset största oreglerade marknaden i Sverige (SCB, 2021a), och det finns inget aktieindex som innehåller alla företag från alla oreglerade marknader i Sverige.

Vi använder stängningskursen på den sista börsdagen för en given vecka som den veckans observation för indexkursen, P . Detta för att matcha datafrekvensen med den för $\Delta SVI_{k,t}$. Vi beräknar därefter den logaritmerade avkastningen på respektive marknad:

$$OMXS_t = \ln\left(\frac{P_{OMXS,t}}{P_{OMXS,t-1}}\right) \quad (4)$$

$$FNSAS_t = \ln\left(\frac{P_{FNSAS,t}}{P_{FNSAS,t-1}}\right) \quad (5)$$

Den logaritmerade avkastningen från ekvation 4 och 5 betecknas härnäst *OMXS* respektive *FNSAS*. Det är dessa som senare kommer användas som beroende variabler i vår regressionsmodell när vi prövar hypotes H2a, H2b och H2c.

När vi undersöker volatiliteten på respektive aktiemarknad använder vi den realiserade volatiliteten. Den realiserade volatiliteten, $Vol_{OMXS,t}$ och $Vol_{FNSAS,t}$ är baserad på daglig data över aktiekursindexens prisnivå P från den senaste veckan, det vill säga de fem senaste handelsdagarna (Brochado, 2020):

$$Vol_t = \sqrt{\sum_{t=t-4}^t \ln\left(\frac{P_t}{P_{t-1}}\right)^2} \quad (6)$$

$Vol_{OMXS,t}$ och $Vol_{FNSAS,t}$ från ekvation 6 kommer användas som beroende variabel när vi prövar hypotes H3a, H3b och H3c.

För att undersöka sambandet mellan FEARS-indexen och handelsvolym, *HVOL*, använder vi den senaste veckans handelsvolym för respektive aktiekursindex och beräknar återigen den logaritmerade förändringen:

$$HVOL_t = \ln\left(\frac{Handelsvolym_t}{Handelsvolym_{t-1}}\right) \quad (7)$$

$HVOL_{OMXS,t}$ och $HVOL_{FNSAS,t}$ från ekvation 7 kommer användas som beroende variabel när vi prövar hypotes H4a, H4b och H4c om förändringar i handelsvolym på respektive marknad.

3.2.7. Kontrollvariabler

Kontrollvariabler används med syftet att kontrollera för faktorer som påverkar våra beroende variabler och investerarsentiment, men som inte ska ingå i FEARS-indexet.

Näringslivets konfidensindikator (*NKI*), är ett mått på det svenska näringslivets uppfattning om nuläget och framtidsutsikterna (Konjunkturinstitutet, 2021). För att mäta detta ställer Konjunkturinstitutet frågor om nuläget och förväntningar på bland annat företagets orderstock, försäljningsvolym, antal anställda och efterfrågan på företagets produkter och tjänster. Konjunkturinstitutet mäter även hushållens uppfattningar och förväntningar, och skapar ett mikroindex (*Mikro*) och ett makroindex (*Makro*) utifrån detta. Frågorna som skapar mikroindexet rör hushållens ekonomi och om det är förmånligt att handla kapitalvaror. Makroindexet berör hushållens syn på den svenska ekonomin i stort. Både *Mikro* och *Makro* används i modellen. *NKI*, *Mikro* och *Makro* är både framåtblickande och en nulägesbedömning eftersom de tillfrågade både svarar på förväntningar inför framtiden och jämför nuläget med vilka förväntningar de haft tidigare. Konjunkturinstitutets tillvägagångssätt följer EU:s riktlinjer för denna typen av undersökningar (Konjunkturinstitutet, 2021).

Förändringen i indexvärdena från föregående månad används som oberoende kontrollvariabel i regressionen. Anledningen till varför vi har valt just dessa kontrollvariabler är för att alla kan tänkas påverka investerarsentiment. Förväntningar om förändringar i inflation, reporäntor, investerarens privatekonomi och de svenska företagens prestationer är faktorer som lär påverka vilka beslut en investerare tar på aktiemarknaden.

Vi använder även ett index som mäter osäkerhet kring ekonomisk politik (*EPO*). Indexet baseras på nyckelord såsom “nationalekonomi” och “osäkerhet” i Expressen, Dagens Industri, Aftonbladet och Svenska dagbladet (Armelius, Hull & Stenbacka Köhler, 2016). Indexet används för att kontrollera för ändrat investerarbeteende till följd av ekonomipolitisk osäkerhet.

Slutligen används även tre makroekonomiska kontrollvariabler som kan påverka aktieavkastningen, köpaktivitet och investerarsentiment. Den första är den kvartalsvisa förändringen i Sveriges bruttonationalprodukt jämfört med föregående kvartal, uttryckt i procent (*BNP*) (SCB, 2021b). Den andra är den månatliga förändringen i konsumentprisindex (*KPI*) (SCB, 2021c). Den tredje kontrollvariabeln är riksbankens reporänta (*Repo*) (Riksbanken, 2021). Eftersom reporäntan är tillgänglig i daglig frekvens använder vi den bestämda reporäntan vid veckans slut som den innevarande veckans observation.

NKI, Mikro, Makro och *EPO* ges månadsvis och har därför omvandlas till veckodata genom linjär interpolering mellan månadsobservationerna. Den linjära interpoleringen drar en rät linje mellan två månadsobservationer. Veckorna som infaller mellan de två månadsobservationerna tilldelas därefter det värde som ligger på den räta linjen vid samma tidpunkt som veckan infaller. Data som *BNP* är baserad på uppdateras kvartalsvis och har också omvandlats till veckodata genom linjär interpolering. Vi logariterar förändringen i *NKI, Mikro, Makro, EPO och BNP* genom att beräkna logaritmen av kvoten mellan värdet i tidpunkt t och värdet i tidpunkt $t-1$. Kontrollvariabeln *Repo* anger den faktiska räntan och logariteras inte eftersom den ibland är oförändrad.

Det är viktigt att notera skillnaden mellan kontrollvariabeln *Makro* och de makroekonomiska kontrollvariablerna *NKI, BNP, KPI* och *Repo*. Skillnaden är att *Makro* är baserad på hushållens förväntningar om den makroekonomiska utvecklingen, och är således mer förknippat med framtidstro. *NKI, BNP, KPI* och *Repo* är den faktiska makroekonomiska utvecklingen och är därmed mer förknippat med förändringar i sentiment i nuläget.

Da, Engelberg och Gao (2015) använde sig av VIX-indexet som kontrollvariabel för investerarsentiment, men det saknas en direkt motsvarighet på den svenska aktiemarknaden. Eftersom det amerikanska VIX-indexet inte tillför något i våra modeller kommer det inte inkluderas som kontrollvariabel.

3.3. Metod för hypotesprövning

I det följande avsnittet förklarar vi de statistiska tester som utförs för att besvara våra hypoteser. Hypotes 1 testas med hjälp av ett *Wilcoxon rank sum*-test, medan hypotes 2-4 prövas med hjälp av olika regressionsmodeller. Vi kommer även presentera de antaganden som modellerna bygger på och olika sorters tester som prövar dessa antaganden.

3.3.1. Wilcoxon rank sum test

Med hjälp av hypotes 1 belyser och bemöter vi den ovisshet om hur positiva och negativa ord kan och bör användas i ett FEARS-index. Tidigare studier har ibland använt enbart positiva, enbart negativa eller både positiva och negativa ord. Resultaten har varit olika signifikanta i

olika studier, och därför är det viktigt att fastställa vilken typ av ord som bör användas i konstruktionen av våra FEARS-index. Från dessa hypoteser konstruerar vi därefter FEARS-index för Stockholmsbörsen och First North Sweden.

Wilcoxon rank sum-test, ibland kallat *Mann-Whitney-Wilcoxon-test* eller bara *Mann-Whitney U-test*, testar nollhypotesen att två oberoende populationer följer samma fördelning. Observationer från de två populationerna rangordnas utifrån dess absoluta värde i en lista i fallande skala. Varje observation tilldelas därefter en binär kategori beroende på vilken population som observationen är tagen ifrån, exempelvis A eller B. Därefter summeras positionstalen för alla observationer som ursprungligen tillhör population A respektive population B. Till exempel, om den första, andra och fjärde observationen i listan kommer från population A blir summan av positionstalen 7 för population A. Denna summa kallas W-värdet. W-värdet för respektive population utvärderas därefter utifrån nollhypotesens antagande att population A och B följer identiska fördelningar. Testet gör inte något antagande om vilken fördelning som de två populationerna följer, utan prövar bara om fördelningen är densamma genom att jämföra medianer. Om nollhypotesen förkastas tyder det på att medianen för en av fördelningarna är förskjuten i någon riktning bort från den andra fördelningens median. Det går att göra både ett tvåsidigt test för att testa om det finns någon förskjutning över huvud taget, och ett ensidigt test för att undersöka förskjutningens riktning. Testet kräver att populationerna är oberoende av varandra och att observationerna åtminstone följer en ordinal skala. Wilcoxon rank sum-test är därför lämpligare att använda när antagandena för *Student's t-test* inte är uppfyllda (Nachar, 2008).

För att pröva våra första hypoteset om att positiva och negativa ord inte väger lika tungt utgår vi från t-värdena som beräknades i avsnitt 3.2.5. Sökorden i baslistan delas upp i två listor, en med de sökord som har positiva t-värden och en med de sökord som har negativa t-värden. Listorna rangordnas därefter utifrån de största t-värdena i absoluta tal. Därefter jämförs de 30 starkast negativa t-värdena med de 30 starkast positiva t-värdena med ett Wilcoxon rank sum-test. Med detta kan vi ta reda på vilken typ av sentiment som har starkast samband med marknadsavkastningen. Vi använder 30 ord eftersom detta är samma antal som ingår i FEARS-indexet. Det kan även uppstå problem om alla ord i baslistan ingår i testet eftersom testet inte tar hänsyn till det absoluta värdet på t-värdet, utan endast följer en ordinal skala. Det kan till exempel vara liten skillnad i absoluta termer mellan t-värden som befinner sig precis över 0 och precis under 0. Wilcoxon rank sum-test tar inte hänsyn till detta utan antar

att värdena kommer från två olika populationer (Nachar, 2008), trots att det är liten skillnad mellan de två observationerna. Detta kan bli missvisande. Vi använder istället de 30 mest representativa orden från varje fördelning.

T-värdena uppfyller kraven för Wilcoxon rank sum-test eftersom de går att rangordna, populationerna är oberoende av varandra och vi behöver inte göra något antagande om vilken fördelning populationerna följer. Om nollhypotesen förkastas betyder det att fördelningarna av t-värdena för positivt och negativt laddade ord skiljer sig från varandra. Detta skulle i sådana fall betyda att den fördelningen som är förskjuten i en positiv riktning har större t-värden i absoluta tal, och således har denna typen av ord ett starkare samband med marknadsavkastningen. Med denna metoden undersöker vi alltså om positivt och negativt sentiment skiljer sig åt enligt hypotes H1a-H1d. Testet utförs först på First North Sweden och Stockholmsbörsen för att pröva hypotes H1a och hypotes H1b. Därefter jämförs de positiva orden på First North Sweden med de positiva orden på Stockholmsbörsen för att pröva hypotes H1c. Slutligen testas de negativa orden på respektive handelsplattform för att pröva hypotes H1d. De 30 mest positiva och negativa t-värdena på respektive marknad presenteras i tabell A.3 i appendix 2.

3.3.2. Regressionsanalyser med minstakvadratmetoden

Metoderna som ligger bakom regressionsmodellerna utgörs av en vanlig minsta kvadratmetod samt en generaliserbar minsta kvadratmetod. Båda metoderna bygger på flertalet antaganden, vilka kommer presenteras i 3.3.3. och 3.3.4.. Vad som däremot skiljer de åt är hur de behandlar korrelerade feltermar. Till skillnad från den vanliga minsta kvadratmetoden som förutsätter att feltermerna är okorrelerade vid jämförelser mellan ekvationer kommer vi utgå från en generaliserbar minsta kvadratmetod när vi undersöker ifall det föreligger en skillnad mellan marknaderna. Mer specifikt kommer den typ av generaliserbara minsta kvadratmetod vara vad som kallas *seemingly unrelated regression* (SUR). Denna tillåter att vi undersöker skillnaden mellan två modeller som väntas ha korrelerade feltermar mellan ekvationer för en given observation men som är okorrelerad mellan observationer (Baltagi, 2011). Korrelation mellan observationer inom en ekvation kommer vidare förklaras i 3.3.4.3..

Beräkningen av skillnaden mellan de två marknaderna kommer genomföras på två sätt. Det första steget är att genomföra beräkningar för de två marknaderna var för sig, vilket utförs

enligt minsta kvadratmetoden. Därefter testar vi ifall riktningskoefficienten mellan de två ekvationerna är skild från noll. Eftersom det är sannolikt att dessa ekvationer kommer ha korrelerade feltermar, trots att vi undersöker för olika marknader med olika FEARS-index, kommer beräkningen genomföras med SUR. Det kan noteras att ifall det trots allt inte skulle förekomma några korrelerade feltermar kommer SUR ge samma resultat som om beräkningen hade genomförts enligt minsta kvadratmetoden, och resultaten kan därför tolkas likadant (Baltagi, 2011).

3.3.3. Variabelanalys

Vi utför diagnostiska tester med syftet att kontrollera att regressionsmodellens antaganden uppfylls. Mer specifikt kommer det för undersökningen vara av intresse att undersöka och kontrollera för multikollinearitet och förekomst av linjärt samband mellan modellens variabler.

3.3.3.1. Multikollinearitet

En regressionsmodell med flera förklarande variabler kan uppvisa multikollinearitet (Dougherty, 2016). Multikollinearitet uppstår när flera av de förklarande variablerna är linjärt korrelerade och gör att modellens skattningar blir mer osäkra. En modell med multikollinearitet kan därför ha en hög förklaringsgrad och samtidigt sakna signifikanta skattningar av de förklarande variablerna. Med hög multikollinearitet blir det därför problematiskt att dra slutsatser från regressionsresultaten (Dougherty, 2016).

Ett sätt att hantera multikollinearitet är genom att exkludera de förklarande variabler som orsakar det (James, Witten, Hastie & Tibshirani, 2013). Detta kan göras genom att beräkna *VIF* (*Variance Inflation Factor*) för de förklarande variablerna och exkludera de variabler med ett högt VIF-värde. VIF-värdet för en enskild förklarande variabel beräknas genom att dividera variansen i flervariabelmodellen med variansen i en modell som endast innehåller den enskilda förklarande variabeln. Ett högt VIF-värde är ett tecken på att den förklarande variabeln har en hög grad av linjär korrelation med andra förklarande variabler. VIF-värdet kan som lägst vara 1 och det betyder att modellen helt saknar multikollinearitet. Som tumregel brukar VIF-värden över 5 användas som tröskelvärde för när modellen lider av för mycket multikollinearitet (James, Witten, Hastie & Tibshirani, 2013). Regressionsmodellerna

i ekvationerna 9-26 utvärderas för multikollinearitet var för sig. Variabler som uppvisar en hög grad av multikollinearitet utesluts och variablerna som kvarstår utgör den slutgiltiga regressionsmodellen. Variablernas VIF-värden presenteras i tabell 8.

3.3.3.2. Linjäritet

En linjär regression förutsätter att sambandet mellan variablerna är linjärt (Dougherty, 2016). Detta antagandet kan testas med hjälp av ett *Ramsey RESET-test (Regression Equation Specification Error Test)*. Testet använder skattningar av den beroende variabeln i en regressionsmodell och adderar olika polynom av dessa skattningar som förklarande variabler i modellen. Testet prövar därefter nollhypotesen att riktningskoefficienterna för polynomen är lika med 0 med hjälp av ett F-test. Om nollhypotesen förkastas antyder det att den ursprungliga linjära regressionen är felspecificerad och att icke-linjära variabler kan förklara delar av modellen, vilket ger belägg för att antagandet om ett linjärt samband inte håller.

Ett sätt att hantera en felspecificerad modell är att transformera variablerna eller att addera icke-linjära variabler (Dougherty, 2016). Eftersom vi redan har logaritmerat våra variabler kan man förvänta sig att modellens samband är linjärt. Logaritmeringen är en monoton transformation som kan skapa ett linjärt samband av ett icke-linjärt samband, eftersom logaritmeringen gör att sambandet går från multiplikativt till additivt.

3.3.4. Residualanalys

Residualer, eller felterm, är den del av datamaterialets variation som modellen inte kan förklara. Utöver att studera de antaganden som hör till variabelselektering, bygger en linjär regression på att följande fem antaganden uppfylls av modellens residualer för att kunna dra slutsatser från regressionen (Brooks, 2014):

1. Väntevärdet av feltermerna är noll.
2. Homoskedasticitet: variansen av feltermerna är konstant.
3. Ingen autokorrelation: feltermerna är oberoende av varandra.
4. Exogena variabler: feltermerna är inte korrelerad med någon av variablerna.
5. Normalfördelade felterm.

Antagande 2, 3 och 5 kommer prövas i denna uppsatsen och således förklaras de i mer detalj i de kommande avsnitten. Antagande 1 uppfylls eftersom våra regressionsmodeller innehåller en konstant term (Dougherty, 2016). Antagande 4 om exogenitet uppnås genom att vi inkluderar kontrollvariabler i modellerna. Utöver dessa antagandena kommer vi även undersöka om vårt datamaterial uppfyller kravet på stationäritet, vilket är ett viktigt antagande vid tidsserieanalys.

3.3.4.1. Homoskedacitet

Antagandet om homoskedacitet innebär att variansen av modellens residualer är konstant oavsett värde hos oberoende variabler (Brooks, 2014). Dess motsats, heteroskedasticitet, betyder att residualerna skiljer sig beroende på vilket värde som den oberoende variabeln antar. Detta kan leda till att modellens estimerade koefficienter lider av ett systematiskt fel. Innan vi testar antagandet om homoskedacitet har modellens beroende variabel logaritmerats. En monoton transformation av den beroende variabeln kommer endast ändra skalan den undersöks utifrån, utan att påverka den på något annat sätt. Vi testar därefter antagandet om homoskedacitet med hjälp av ett *Breusch Pagan*-test i tabell 8 med diagnostiska tester. Breusch Pagan-testet undersöker ifall residualernas variation beror på värdena för de oberoende variablerna, i vilket fall det då skulle innebära att vi erhåller heteroskedasticitet.

3.3.4.2. Stationäritet

Stationäritet är en förutsättning för att tidsseriesdata ska vara jämförbar och för att regressionsresultat med tidsseriesdata ska vara pålitliga (Cryer & Chan, 2008). Stationäritet är något unikt för tidsserieanalys jämfört med tvärsnittsdata. En stationär tidsserie bygger på tre antaganden. De första två antagandena är att variabelns väntevärde och varians är konstanta över hela tidsperioden, och därmed oberoende av t . Detta är en av anledningarna till varför tidsseriesdatan normaliserades och den naturliga logaritmen av sökintresset differentierades i ekvation 2. Det tredje antagandet är att kovariansen mellan två tidpunkter i tidsserien, y_t och y_{t-s} , beror på s men inte av t . Antagandena om en stationär tidsserie kan alltså kort sammanfattas som att datamaterialet uppvisar samma egenskaper oavsett var i tidsseriesdatan observeras (Dougherty, 2016).

Eftersom vårt slutgiltiga FEARS-index är en tidsserie behöver vi testa om tidsserien är stationär. Detta görs med ett *Dickey-Fuller*-test (Dougherty, 2016). Testet prövar nollhypotesen att värdet på en observation i en tidsserie endast utgörs av en oberoende felterm samt det föregående värdet multiplicerat med en faktor större än eller lika med ett. Om det stämmer skulle det betyda att tidsserien är icke-stationär eftersom medelvärdet och variansen inte längre är konstant över tiden. Den bästa prediktionen av nästa värde i tidsserien är helt enkelt det föregående värdet, eftersom resten av variationen är den oberoende feltermen. Mothypotesen är att tidsserien tenderar att återgå till samma värde. Nollhypotesen i Dickey-Fuller-testet ska således förkastas för att tidsserien ska vara stationär (Dougherty, 2016). Notera att i övriga diagnostiska tester brukar man inte vilja förkasta nollhypotesen för det innebär att antagandet bryts. I Dickey-Fuller-testet vill vi förkasta nollhypotesen.

3.3.4.3. Ingen autokorrelation

Det tredje antagandet är att regressionsmodellens feltermen är oberoende av varandra, vilket därigenom också innebär att dess autokorrelation kommer vara noll (Brooks, 2014). Autokorrelation mäter korrelationen mellan processens olika tidpunkter och givet att modellens residualer inte endast utgörs av vitt brus innebär detta att det går att modellera framtida värden genom att känna till dess historiska. Vitt brus kallas även för en *random walk*. Modellen kommer i sådana fall behöva kompletteras med en eller flera tidsfördröjda variabler till dess att residualerna endast utgörs av vitt brus.

Eventuell autokorrelation i modellen kommer testas utifrån en *Box Jenkins*-ansats. Detta tillvägagångssätt, till skillnad från ett *Durbin Watson*-test, gör att vi inte bara förstår ifall autokorrelation förekommer, utan ger även information om hur modellen bör utvecklas för residualerna ska utgöras av vitt brus. Ansatsen förutsätter stationaritet och studerar sekvenser genererade av slumpmässiga tal över tid, vilka därefter modelleras med hjälp av en så kallad *ARMA*-modell (Cryer & Chan, 2008). Denna modell består av en autoregressiv (AR) komponent och ett glidande medelvärde (MA). AR-komponenten undersöker ifall en observation beror på dess tidigare värden och MA-komponenten testar beroendet mellan observationen och standardfelen från ett glidande medelvärde av tidigare observationer.

De två komponenterna, MA och AR, som utgör denna ARMA-modell, testas var för sig med hjälp av så kallad autokorrelationsfunktion (ACF) respektive partiell autokorrelationsfunktion (PACF) (Cryer & Chan, 2008). Beroendeförhållanden med tidigare värden i processen utvärderas utifrån att en tidsfördröjd variabel har en påverkan på det nutida värdet i en omfattning som överstiger två standardavvikelser. Ifall detta, å andra sidan, inte påträffas innebär det att autokorrelation inte heller kan påvisas och vi erhåller vad som kallas en ARMA (0,0). Residualerna utgörs då endast av vitt brus. Fram till dess att residualerna inte längre återger ett samband med tidigare värden kommer tidsfördröjda observationer inkluderas i modellen. Detta är anledningen till varför ekvationerna för att testa hypotes 2 skiljer sig från ekvationerna som testar hypotes 3 och 4.

3.3.4.4. Normalfördelade felterm

Antagandet om normalfördelade residualer uppkommer ur den så kallade centrala gränsvärdessatsen. Denna menar att vid ett större antal observationer kommer dess fördelning gå mot normalfördelningen (Brooks, 2014). I de fall där modellen inte har normalfördelade residualer kan det ge en indikation på att modellens residualer inte uppkommer slumpmässigt, utan istället följer en underliggande fördelning. Vi använder oss av ett *Jarque Bera*-test för att testa ifall residualerna är normalfördelade.

Även fast den centrala gränsvärdessatsen säger att residualerna kommer följa en normalfördelning när antalet observationer är stort, finns det anledning att tro att detta inte nödvändigtvis kommer hålla i vårt fall. I finansiell forskning följer nämligen residualer ofta fördelningar med svansar som är större än hos normalfördelningen (Rachev, 2003). Detta innebär att sannolikheten för outliers är större än vad som tillåts förekomma enligt normalfördelningen. Avslutningsvis innebär det att ett *Jarque Bera*-test som förkastar normalitet inte nödvändigtvis behöver innebära att utformningen av modellen är olämplig.

3.4. Akaike Information Criterion

För att pröva hur robusta våra resultat är använder vi även en annan metod för att konstruera FEARS-indexet. Istället för korrelationsanalysen som beskrivs i avsnitt 3.2.5. använder vi även en stegvis variabelselektering utifrån *Akaike Information Criterion* (AIC) för att välja ord som ska inkluderas i FEARS-indexet. Vi skapar alltså ytterligare ett FEARS-index med

hjälp av AIC för att undersöka om våra resultat håller även om vi använder en annan metod för att konstruera FEARS-indexet.

AIC utförs i modellselekteringssyfte genom att välja de oberoende variabler i en modell som bäst förklarar data (Sheather, 2009). AIC utgår från den relativa nyttan varje ytterligare variabel tillför, genom att inkludera en penalisering för antalet variabler som modellen använder. På så vis hoppas modellen endast inkludera de variabler som tillför tillräckligt stor nytta för att det ska överstiga en eventuell överanpassad modell.

3.5. FEARS-index och avkastning

Hypotes H2a, H2b och H2c prövas med hjälp av en multipel linjär regression, varav H2c genomförs enligt en generaliserbar minsta kvadratmetod. I syfte att underlätta genomgången av ekvation 9-26, kommer de kontrollvariabler som förekommer för samtliga ekvationer ersättas av en summa av kontrollvariabler, KV_t , från ekvation 8. Regressionen på Stockholmsbörsen utförs enligt ekvation 9 och 10. Regressionen på First North Sweden utförs enligt ekvation 11 och 12. Vi utvärderar även skillnaden mellan marknaderna med en generaliserad linjär regression enligt ekvation 13 och 14:

$$KV_t = \beta_2 \cdot NKI_t + \beta_3 \cdot Mikro_t + \beta_4 \cdot Makro_t + \beta_5 \cdot BNP_t + \beta_6 \cdot EPO_t + \beta_7 \cdot KPI_t + \beta_8 \cdot Repo_t. \quad (8)$$

$$OMXS_t = \beta_0 + \beta_{1,OMXS} \cdot FEARS_{OMXS,t} + KV_t + \varepsilon_t. \quad (9)$$

$$OMXS_{t+1} = \beta_0 + \beta_{1,OMXS} \cdot FEARS_{OMXS,t} + KV_t + \varepsilon_t. \quad (10)$$

$$FNSAS_t = \beta_0 + \beta_{1,FNSAS} \cdot FEARS_{FNSAS,t} + KV_t + \varepsilon_t. \quad (11)$$

$$FNSAS_{t+1} = \beta_0 + \beta_{1,FNSAS} \cdot FEARS_{FNSAS,t} + KV_t + \varepsilon_t. \quad (12)$$

$$\begin{aligned}
OMXS_t - FNSAS_t &= (\beta_{0,OMXS} - \beta_{0,FNSAS}) + \\
&\beta_1 \cdot (FEARS_{OMXS,t} - FEARS_{FNSAS,t}) + KV_t + \varepsilon_t.
\end{aligned} \tag{13}$$

$$\begin{aligned}
OMXS_{t+1} - FNSAS_{t+1} &= (\beta_{0,OMXS} - \beta_{0,FNSAS}) + \\
&\beta_1 \cdot (FEARS_{OMXS,t} - FEARS_{FNSAS,t}) + KV_t + \varepsilon_t.
\end{aligned} \tag{14}$$

Den oberoende variabeln är FEARS-indexet för respektive marknad vid tidpunkt t . Den beroende variabeln används både i tidpunkt t och $t+1$. I modellen ingår även kontrollvariablerna *NKI*, *Mikro*, *Makro*, *EPO*, *Repo*, *BNP* och *KPI* för tidpunkt t .

Märk väl att den slutgiltiga modellen är baserad på aktuell data, t . De beroende variablerna kommer även undersökas i tidpunkt $t+1$. På detta sättet undersöks huruvida FEARS-indexen vid tidpunkt t påverkar framtida avkastning, och inte bara att det finns ett samband i nutid. När tidsfördröjda värden används på detta sättet kan det dock uppstå problem med bland annat autokorrelation (Dougherty, 2016). Därför måste modellens antaganden testas varje gång nya tidsfördröjningar används.

3.6. FEARS-index, volatilitet och handelsvolym

De tredje hypoteserna prövas på samma sätt med hjälp av multipel linjär regression. Den beroende variabeln är volatiliteten på Stockholmsbörsen (ekvation 15 och 16), First North Sweden (ekvation 17 och 18) samt skillnaden mellan de två marknaderna (ekvation 19 och 20). Volatiliteten är beräknad enligt ekvation 6. Till skillnad från regressionerna för hypotes 2 har vi adderat en och två veckors tidsfördröjningar av den beroende variabeln volatilitet som kontrollvariabel. Detta tillägg till modellerna har gjorts för att undvika autokorrelerade residualer, vilket inte behövdes i de tidigare regressionerna. Annorlunda uttryckt har vi adderat tidsfördröjda variabler till dess att vi kan påvisa residualer som återger vitt brus.

$$\begin{aligned}
Vol_{OMXS,t} &= \beta_0 + \beta_1 \cdot FEARS_{OMXS,t} + KV_t + \\
&\beta_9 \cdot Vol_{OMXS,t-1} + \beta_{10} \cdot Vol_{OMXS,t-2} + \varepsilon_t.
\end{aligned} \tag{15}$$

$$\begin{aligned}
Vol_{OMXS,t+1} &= \beta_0 + \beta_1 \cdot FEARS_{OMXS,t} + KV_t + \\
&\beta_9 \cdot Vol_{OMXS,t-1} + \beta_{10} \cdot Vol_{OMXS,t-2} + \varepsilon_t.
\end{aligned} \tag{16}$$

$$\begin{aligned}
Vol_{FNSAS,t} &= \beta_0 + \beta_1 \cdot FEARS_{FNSAS,t} + KV_t + \\
&\beta_9 \cdot Vol_{FNSAS,t-1} + \beta_{10} \cdot Vol_{FNSAS,t-2} + \varepsilon_t.
\end{aligned} \tag{17}$$

$$\begin{aligned}
Vol_{FNSAS,t+1} &= \beta_0 + \beta_1 \cdot FEARS_{FNSAS,t} + KV_t + \\
&\beta_9 \cdot Vol_{FNSAS,t-1} + \beta_{10} \cdot Vol_{FNSAS,t-2} + \varepsilon_t.
\end{aligned} \tag{18}$$

$$\begin{aligned}
Vol_{OMXS,t} - Vol_{FNSAS,t} &= (\beta_{0,OMXS} - \beta_{0,FNSAS}) + \\
&\beta_1 \cdot (FEARS_{OMXS,t} - FEARS_{FNSAS,t}) + KV_t + \\
&\beta_9 \cdot (VOL_{OMXS,t-1} - VOL_{FNSAS,t-1}) + \\
&\beta_{10} \cdot (Vol_{OMXS,t-2} - Vol_{FNSAS,t-2}) + \varepsilon_t.
\end{aligned} \tag{19}$$

$$\begin{aligned}
Vol_{OMXS,t+1} - Vol_{FNSAS,t+1} &= (\beta_{0,OMXS} - \beta_{0,FNSAS}) + \\
&\beta_1 \cdot (FEARS_{OMXS,t} - FEARS_{FNSAS,t}) + KV_t + \\
&\beta_9 \cdot (VOL_{OMXS,t-1} - VOL_{FNSAS,t-1}) + \\
&\beta_{10} \cdot (Vol_{OMXS,t-2} - Vol_{FNSAS,t-2}) + \varepsilon_t.
\end{aligned} \tag{20}$$

För att testa våra fjärde hypoteser utför vi även regressioner på liknande vis som i hypotes 3, förutom att vi nu använder handelsvolymen från ekvation 7 som beroende variabel. Återigen har tidsfördröjningar av den beroende variabeln inkluderats som kontrollvariabel till dess att vi inte längre har autokorrelerade feltermen.

$$\begin{aligned}
HVOL_{OMXS,t} &= \beta_0 + \beta_1 \cdot FEARS_{OMXS,t} + KV_t + \\
&\beta_9 \cdot HVOL_{OMXS,t-1} + \beta_{10} \cdot HVOL_{OMXS,t-2} + \varepsilon_t.
\end{aligned} \tag{21}$$

$$\begin{aligned}
HVOL_{OMXS,t+1} &= \beta_0 + \beta_1 \cdot FEARS_{OMXS,t} + KV_t + \\
&\beta_9 \cdot HVOL_{OMXS,t-1} + \beta_{10} \cdot HVOL_{OMXS,t-2} + \varepsilon_t.
\end{aligned} \tag{22}$$

$$\begin{aligned}
HVOL_{FNSAS,t} &= \beta_0 + \beta_1 \cdot FEARS_{FNSAS,t} + KV_t + \\
&\beta_9 \cdot HVOL_{FNSAS,t-1} + \beta_{10} \cdot HVOL_{FNSAS,t-2} + \varepsilon_t.
\end{aligned} \tag{23}$$

$$\begin{aligned}
HVOL_{FNSAS,t+1} &= \beta_0 + \beta_1 \cdot FEARS_{FNSAS,t} + KV_t + \\
&\beta_9 \cdot HVOL_{FNSAS,t-1} + \beta_{10} \cdot HVOL_{FNSAS,t-2} + \varepsilon_t.
\end{aligned} \tag{24}$$

$$\begin{aligned}
HVOL_{OMXS,t} - HVOL_{FNSAS,t} &= (\beta_{0,OMXS} - \beta_{0,FNSAS}) + \\
&\beta_1 \cdot (FEARS_{OMXS,t} - FEARS_{FNSAS,t}) + KV_t + \\
&\beta_9 \cdot (HVOL_{OMXS,t-1} - HVOL_{FNSAS,t-1}) + \\
&\beta_{10} \cdot (HVOL_{OMXS,t-2} - HVOL_{FNSAS,t-2}) + \varepsilon_t.
\end{aligned} \tag{25}$$

$$\begin{aligned}
HVOL_{OMXS,t+1} - HVOL_{FNSAS,t+1} &= (\beta_{0,OMXS} - \beta_{0,FNSAS}) + \\
&\beta_1 \cdot (FEARS_{OMXS,t} - FEARS_{FNSAS,t}) + KV_t + \\
&\beta_9 \cdot (HVOL_{OMXS,t-1} - HVOL_{FNSAS,t-1}) + \\
&\beta_{10} \cdot (HVOL_{OMXS,t-2} - HVOL_{FNSAS,t-2}) + \varepsilon_t.
\end{aligned} \tag{26}$$

I regressionsavsnitten är det alltså sammanfattningsvis avkastning, volatilitet, handelsvolym och FEARS-indexen på respektive marknad som är av intresse i denna uppsatsen. Dessa variabler finns även sammanställda i tabellen över den deskriptiva statistiken i tabell A.5 appendix 2.

3.7. Metod- och datakritik

En svaghet med att använda data från Google Trender är att index-värdena är baserade på en delmängd av de faktiska sökningarna på Google. Dessutom saknas information hur Google gör detta urval när de räknar ut GSVI. Som bekant gör Google detta för att förbättra svarstiden på sin hemsida. Detta skapar reliabilitetsproblem eftersom $GSVI_{k,t}$ antar olika värden om data hämtas vid olika tidpunkter. Av denna anledningen följer vi Preis, Moat och Stanley (2013) och använder det genomsnittliga indexvärdet över flera datahämtningar för att få indexvärden som är mer representativa. Antalet hämtningar utförs till dess att en ytterligare hämtning inte kommer medföra ett signifikant annorlunda resultat, där differensen testas med hjälp av ett t-test (se tabell A.4 i appendix 2). Trots detta kvarstår kritiken mot studiens reliabilitet, och detta problem går inte att undvika utan att få tillgång till Googles obehandlade data.

I vår studie använder vi oss enbart av sökordens svenska översättningar och inte engelska. Det finns därför risk att vi når en annan målgrupp jämfört med om vi enbart använt engelska översättningar. Det är till exempel inte otänkbart att unga svenskar antagligen är mer benägna att söka på engelska jämfört med äldre. Dessutom gör detta att vi inte får med alla relaterade sökningar som en investerare kan tänkas göras. Om en investerare först söker på svenska och sedan på engelska kommer vår undersökning endast täcka den svenska översättningen. Vi utförde stickprov på engelska ord med avgränsning till sökningar gjorda i Sverige, och insåg att sökintresset ofta var för lågt. Dessa sökord hade sorterats bort på grund av begränsad datamängd. Att inte inkludera engelska översättningar påverkar alltså studiens validitet eftersom modellen inte fullt ut mäter allt det sentiment som investerare har. Samtidigt förenklas tillvägagångssättet genom att inte inkludera engelska översättningar. En viktig fördel med FEARS-indexet gentemot andra mått på investerarsentiment är just dess enkelhet och tillgänglighet. Denna fördelen kan gå förlorad om modellen blir alltför komplicerad.

Vissa sökord går även att använda i en annan kontext än den tilltänkta. Som nämnt kan sökningar på exempelvis sökordet "fattig" ha andra avsikter än de vi anat. Detta är ett tydligt exempel på problem med att använda GSVI som mått på investerarsentiment: vi vet inte avsikten bakom sökningen på Google. Detta är anledningen till varför vi använder ord som har de mest signifikanta t-värdena eftersom det är ett sätt att "let the data speak for itself." (Da, Engelberg & Gao, 2015, sid. 9). Trots att vi endast använder oss av de 30 mest

signifikant t-värdena kvarstår risken att ord vars användning kan förekomma i flera olika sammanhang fördunklar resultatet.

Eftersom vi har avgränsat den geografiska regionen till Sverige, både sett till vilken aktiemarknad vi undersöker och vilka sökresultat från Google som vi använder, kommer investerare som handlar över landsgränsen att uteslutas från undersökningen. Sökiintresset från svenskar som placerar sitt kapital utomlands inkluderas i GSVI, men deras investeringar omfattas inte av studien. Det omvända gäller för utländska investerare som placerar sitt kapital i Sverige. Deras investeringar omfattas av studien men deras sökiintresse på Google fångas inte av GSVI. Eftersom GSVI framförallt mäter icke-professionella investerares sentiment är vi intresserade av hushållens sentiment, och det finns anledning att tro att just de utländska hushållen utgör en liten del av ägandeskapet i den svenska aktiemarknaden. I SCBs (2021a) statistik skiljer Statistiska centralbyrån inte på olika utländska sektorer, och därför går det inte att säga exakt hur stor andel de utländska hushållen äger. Vi uppskattar att detta inte påverkar våra resultat avsevärt.

Det går även att diskutera ifall man verkligen behöver följa Da, Engelberg och Gao (2015) och winsoriserat datamaterialet från Google Trender. Data som vi winsoriserar är som bekant ett genomsnitt av data från flera hämtningar. Om ett extremvärde finns kvar även efter genomsnittsberäkningen finns det anledning att tro att extremvärdet faktiskt inte är en anomali. För att ett extremvärde ska finnas kvar efter genomsnittsberäkningen krävs det nämligen att samma observation uppvisar ett extremvärde vid flera nedladdningar. Genomsnittsberäkningen i sig kan därför ses som ett sätt att justera för extremvärden och därför går det att argumentera för att winsoriseringen inte är nödvändig och att den istället leder till informationsbortfall.

Det finns flera brister med våra kontrollvariabler. Flera av kontrollvariablerna har en annan datafrekvens än FEARS-indexen och de beroende variablerna. Dessa variabler är inte tillgängliga i en högre frekvens. Det är något som annars hade kunnat öka precisionen. För att hantera detta använde vi oss av linjär interpolering mellan månadsobservationerna för att skapa observationer för varje vecka i datamaterialet. Detta är en approximation, och i detta steget fanns det flera andra sätt att gå till väga som hade påverkat vårt resultat annorlunda. Exempelvis går det att använda icke-linjär interpolering eller anta att alla veckor som tillhör samma månad antar samma värde. Det kan, trots allt, anses som ett rimligt antagande att

använda sig av linjär interpolering eftersom det är osannolikt att de makroekonomiska variablerna skulle variera stort mellan månadsobservationerna. Därför menar vi att en linjär förändring är en bra uppskattning av den verkliga veckoförändringen. Det är även naturligt att interpolera eftersom kontrollvariablerna antingen uttrycker ett genomsnittlig eller en summa för den innevarande månaden. De är alltså inte mätningar vid en enskild tidpunkt. Exempelvis är inflationen uttryckt som den genomsnittliga prisförändringen för den innevarande månaden och BNP mäts som förändringen av BNP för *hela* det innevarande kvartalet jämfört med föregående kvartal. Slutligen är det även en nackdel att våra kontrollvariabler är kontrollvariabler för hela marknaden, medan våra beroende variabler endast utgör en liten del av marknaden. Detta innebär att våra kontrollvariabler inte är helt anpassade för undersökningen.

Det hade varit gynnsamt om datamaterialet hade daglig frekvens. Med daglig frekvens hade vi kunnat undersöka mer granulära förändringar i sentiment som en konsekvens av plötsliga impulser hos investeraren. Med veckodata kan vi inte se förändringar i sentiment som sker inom samma vecka. Vi kan inte heller styra vilka sju dagar som ska ingå i en observation, utan observationerna motsvarar kalenderveckor. Samtidigt, om vi hade använt dagliga värden från Google Trender hade det inte gått att undersöka mer än en period på nio månader, eftersom datafrekvens övergår till veckodata om tidsperioden överstiger nio månader. Med daglig data hade vi dessutom haft färre observationer i varje tidsserie. Det hade inte heller gått att sammanfoga flera på varandra följande sekvenser av nio månaders daglig data, eftersom GSVI för en given dag är beroende av vilken tidsperiod som observationen tillhör. Detta är en konsekvens av hur Google Trender fungerar och en klar begränsning med verktyget. Vi hade även kunnat använda månatlig frekvens på FEARS-indexen för att matcha kontrollvariablernas frekvens. Vi menar dock att detta inte hade lett till lika meningsfulla resultat eftersom vi vet med oss att förändringar i investerarsentiment framförallt får konsekvenser de nästkommande dagarna. Däremot är detta något som hade varit intressant att undersöka i vidare forskning.

3.8. Sammanställning av hypoteser och metodik

Tabell 2 sammanfattar metodavsnittet genom att redogöra för vilka hypoteser som ska prövas i denna uppsatsen, samt förklara hur hypotesprövningen utförs.

Tabell 2. Sammanställning av uppsatsens hypoteser och tillvägagångssätt för hur dessa ska utvärderas.

Hypotes	Variabel	Test	Utförande
H1a: Det finns ingen skillnad på positiva och negativa ord i investerarsentiment på Stockholmsbörsen.	T-värdet för korrelationen mellan sökord och marknadsavkastning på <i>OMXS</i> .	Wilcoxon rank sum test.	De 30 mest positiva t-värdena jämförs med de 30 mest negativa.
H1b: Det finns ingen skillnad på positiva och negativa ord i investerarsentiment på First North Sweden.	T-värdet för korrelationen mellan sökord och marknadsavkastning på <i>FNSAS</i> .	Wilcoxon rank sum test.	De 30 mest positiva t-värdena jämförs med de 30 mest negativa.
H1c: Det finns ingen skillnad på positiva ord i investerarsentiment på Stockholmsbörsen jämfört med på First North Sweden.	T-värdet för korrelationen mellan sökord och marknadsavkastning på respektive marknad.	Wilcoxon rank sum test.	De 30 mest positiva t-värdena på Stockholmsbörsen jämförs med de 30 mest positiva t-värdena på First North Sweden.
H1d: Det finns ingen skillnad på negativa ord i investerarsentiment på Stockholmsbörsen jämfört med på First North Sweden.	T-värdet för korrelationen mellan sökord och marknadsavkastning på respektive marknad.	Wilcoxon rank sum test.	De 30 mest negativa t-värdena på Stockholmsbörsen jämförs med de 30 mest negativa t-värdena på First North Sweden.
H2a: FEARS-index påverkar inte aktiekursförändringar på Stockholmsbörsen.	$FEARS_{OMXS}$ och <i>OMXS</i> .	Multipel linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 9 och 10.
H2b: FEARS-index påverkar inte aktiekursförändringar på First North Sweden.	$FEARS_{FNSAS}$ och <i>FNSAS</i> .	Multipel linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 11 och 12
H2c: FEARS-index påverkan på aktiekursförändringar skiljer sig inte mellan Stockholmsbörsen och First North Sweden.	$FEARS_{OMXS}$, $FEARS_{FNSAS}$, <i>OMXS</i> och <i>FNSAS</i> .	Generaliserad linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 13 och 14.

H3a: FEARS-index påverkar inte förändringar i volatilitet på Stockholmsbörsen.	$FEARS_{OMXS}$ och Vol_{OMXS} .	Multipel linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 15 och 16.
H3b: FEARS-index påverkar inte förändringar i volatilitet på First North Sweden.	$FEARS_{FNSAS}$ och Vol_{FNSAS} .	Multipel linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 17 och 18.
H3c: FEARS-index påverkan på volatilitet skiljer sig inte mellan Stockholmsbörsen och First North Sweden.	$FEARS_{OMXS}$, $FEARS_{FNSAS}$, Vol_{OMXS} och Vol_{FNSAS} .	Generaliserad linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 19 och 20.
H4a: FEARS-index påverkar inte förändringar i handelsvolym på Stockholmsbörsen.	$FEARS_{OMXS}$ och $HVOL_{OMXS}$.	Multipel linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 21 och 22.
H4b: FEARS-index påverkar inte förändringar i handelsvolym på First North Sweden.	$FEARS_{FNSAS}$ och $HVOL_{FNSAS}$.	Multipel linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 23 och 24.
H4c: FEARS-index påverkan på handelsvolym skiljer sig inte mellan Stockholmsbörsen och First North Sweden.	$FEARS_{OMXS}$, $FEARS_{FNSAS}$, $HVOL_{OMXS}$ och $HVOL_{FNSAS}$.	Generaliserad linjär regression.	Analys av deskriptiv statistik från regressionsmodellen i ekvation 25 och 26.

4. RESULTAT OCH ANALYS

I detta avsnittet presenterar och analyserar vi de resultat som studien kommit fram till. Avsnittet börjar med att vi redogör för resultaten från Wilcoxon rank sum-testerna och regressionsmodellerna som prövar våra hypoteser. Därefter presenterar vi resultaten av AIC-modellen och de diagnostiska testerna som prövar modellernas antaganden.

4.1. Hypotes 1 - Jämförelse av positivt och negativt sentiment

I hypotes 1 undersöker vi om det är skillnad på positivt och negativt sentiment, både inom och mellan reglerade och oreglerade aktiemarknader i Sverige. Detta görs utifrån testresultat och en visualisering av datamaterialet. I tabell A.3 i appendix 2 framgår det vilka sökord som används till Wilcoxon rank sum-testerna. Utifrån dessa får vi sedan de sökord som används i FEARS-indexen i hypotes 2, 3 och 4, vilket framgår i tabell A.2 i appendix 2.

4.1.1. Resultat för hypotes H1a-H1d

För att testa de första hypoteserna utför vi Wilcoxon rank sum-tester. Resultatet av dessa testerna presenteras i tabell 3. Tabellen visar nollhypotes, mothypotes, W-värde och motsvarande p-värde för hypotes H1a-H1d. Tabellen är även indelad i två segment, ett för tvåsidiga tester och ett för ensidiga tester.

Resultatet av Wilcoxon rank sum-test indikerar att det är skillnad i fördelningen av negativt och positivt laddade ord på både Stockholmsbörsen och First North Sweden. Testets nollhypotes, att fördelningarna av positivt och negativt laddade ord inte skiljer sig åt, förkastas på båda aktiemarknaderna. Detta föranleder att vi förkastar hypotes H1a och H1b. Testen är tvåsidiga och jämför fördelningarnas medianer. Därför fastslås det inte i vilken riktning de två fördelningarna är förskjutna i förhållande till varandra. Det säger därför ingenting om huruvida det är de negativa eller positiva orden som har starkast samband med aktieavkastningen. Om vi däremot undersöker listorna över t-värden i tabell A.3 i appendix 2 indikerar den höga andelen positiva t-värden att det är positiva ord som har störst påverkan på avkastning på Stockholmsbörsen. Figur 5 visualiserar detta resultat med hjälp av ett lådagram, och det blir därmed tydligt att de positiva orden borde ha en fördelning av t-värden som tenderar att vara högre i absoluta tal än de negativa orden.

Vi prövar detta genom att genomföra tester med samma nollhypotes men med en ensidig mothypotes. Nollhypoteserna förkastas återigen med hög signifikansnivå på båda aktiemarknaderna. Resultaten säger att förskjutningen mellan fördelningarnas medianer är negativ. I vårt fall är datamaterialet ordnat på ett sådant sätt att en negativ förskjutning innebär att fördelningen av positiva ord har en större median i absoluta termer jämfört med negativa ord. Detta är intuitivt utifrån figur 5, där de positiva orden tydligt är fördelade högre upp i lådagrammet. Vi kan således fastställa att H1a och H1b förkastas. Det är skillnad på positiva och negativa ord, både på Stockholmsbörsen och First North Sweden. Skillnaden är att positivt laddade ord har ett starkare samband med aktieavkastningen jämfört med negativa ord.

Tabell 3. Resultat av ensidiga och tvåsidiga Wilcoxon rank sum-tester för att pröva hypotes H1a, H1b, H1c och H1d.

<i>Wilcoxon rank sum test</i>				
<i>Tvåsidigt test:</i>				
<i>H₀: Förskjutning = 0</i>				
<i>H_A: Förskjutning ≠ 0</i>	<i>H1a</i>	<i>H1b</i>	<i>H1c</i>	<i>H1d</i>
W-värde	31	15	199	494
P-värde	5,998E-13	1,157E-14	1,34E-4	0,5229
<i>Ensidigt test:</i>				
<i>H₀: Förskjutning = 0</i>				
<i>H_A: Förskjutning < 0</i>	<i>H1a</i>	<i>H1b</i>	<i>H1c</i>	<i>H1d</i>
W-värde	31	15	199	-
P-värde	2,999E-13	5,784E-15	6,698E-5	-

H1a: Det är skillnad på positiva och negativa ord i FEARS på Stockholmsbörsen.

H1b: Det är skillnad på positiva och negativa ord i FEARS på First North Sweden.

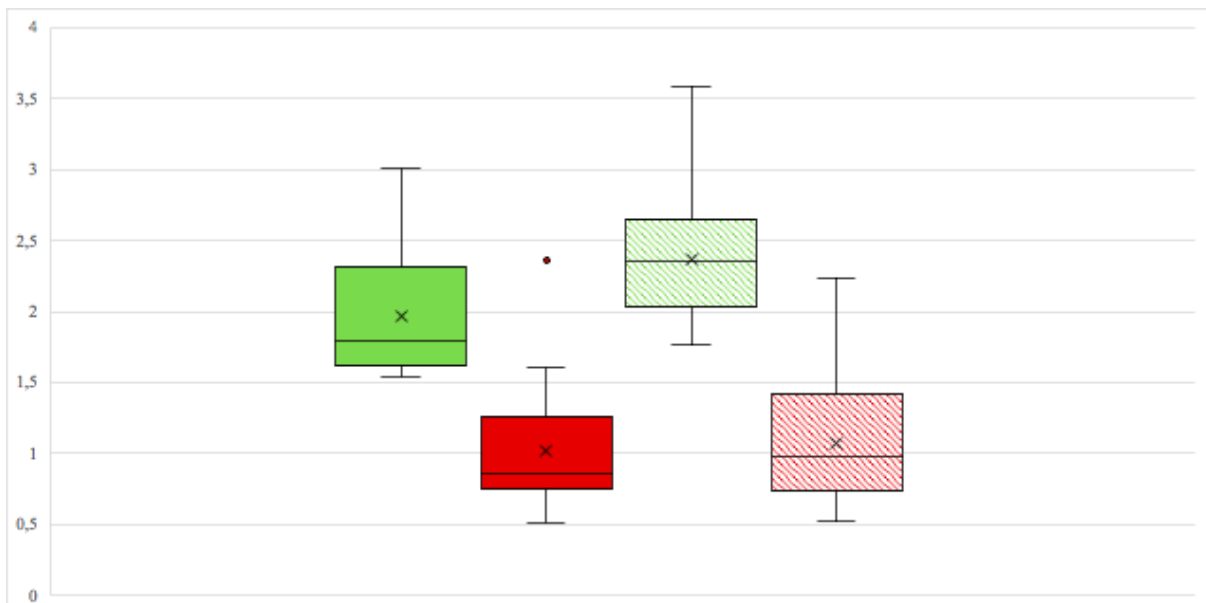
H1c: Det är skillnad på positiva ord i FEARS på Stockholmsbörsen jämfört med på First North Sweden.

H1d: Det är skillnad på negativa ord i FEARS på Stockholmsbörsen jämfört med på First North Sweden.

I tabell 3 jämför vi även om det är skillnad mellan Stockholmsbörsen och First North Sweden. I det tvåsidiga testet som jämför skillnaden mellan positiva ord på Stockholmsbörsen och First North Sweden kan nollhypotesen förkastas. I det motsvarande

ensidiga testet kan vi återigen förkasta nollhypotesen med hög signifikansnivå. Datamaterialet är ordnat på ett sådant sätt att den alternativa hypotesen säger att de positiva orden på First North Sweden har en högre fördelning än de positiva orden på Stockholmsbörsen. Denna skillnaden syns tydligt i figur 5 då det randiga gröna lådagrammet är högre upp än det helfärgade gröna lådagrammet. Positivt laddade ord har alltså ett starkare samband med aktieavkastningen på First North Sweden än på Stockholmsbörsen.

Den enda tvåsidiga nollhypotesen som vi inte kan förkasta är den som prövar om det finns någon skillnad mellan negativt laddade ord på Stockholmsbörsen och First North Sweden. Enligt resultaten går det inte att påvisa någon skillnad. Eftersom en tvåsidig nollhypotes inte kunde förkastas kommer inte ett ensidigt test genomföras.



Figur 5. Lådagram över fördelningarna av de 30 mest positiva (grönt) och negativa (rött) t-värdena på Stockholmsbörsen (helfärgat) och på First North Sweden (randigt). T-värden uttryckta i absoluta tal. Medelvärdet är markerat med ett kryss och medianen med en horisontell linje.

4.1.2. Analys av hypotes H1a-H1d

Resultaten att positivt sentiment är starkare på just First North Sweden stämmer väl överens med teorierna om vilken typ av investerare som påverkas av sin sinnesstämning. Statistiken

från SCB (2021a) visar att de svenska hushållen utgör en större del av de oreglerade svenska aktiemarknaderna jämfört med de reglerade marknaderna. Hushållen är alltså en större aktör på oreglerade marknader än på reglerade marknader. Dessutom säger teorierna att icke-professionella investerare tenderar att vara mer spekulativa och påverkas mer av sin sinnesstämning. Därför blir det naturligt att sambandet mellan värdeladdade ord och aktieavkastning är starkare på de marknader där spekulativa investerare är en större aktör. Resultaten från Da, Engelberg och Gao (2011), att sökaktivitet på Google framförallt fångar icke-professionella investerares sökaktivitet, stödjer också våra resultat. Google Trender har lyckats skilja på marknader som har olika stor andel icke-professionella investerare.

Våra resultat talar emot resultatet av Baumeister et al. (2001), att negativt laddade ord väger tyngre än positivt laddade ord. En möjlig förklaring är vårt val av tidsperiod. Denna är präglad av stigande aktiekurser och saknar börskrasch, till skillnad från Da, Engelberg och Gao (2015). Da, Engelberg och Gao finner att bland de ord som har starkast samband med aktieavkastningen är sambandet oftast negativt. Deras tidsperiod innehåller bland annat finanskrisen 2008 som gav kraftigt fallande aktiekurser. Börskraschen kan även tänkas påverka investerarnas sinnesstämning och göra dem mer försiktiga och pessimistiska inför framtiden. Brochado (2020) som undersöker perioden 2009 till 2014 har uteslutit börskraschen från sin tidsperiod, och finner att de positivt laddade orden har ett starkare samband med aktiekursutvecklingen. Vilka ord som har starkast samband med aktiekursutvecklingen tycks därför bero på vilken tidsperiod som undersöks, och vår tidsperiod innehåller mest uppgångar.

4.2. Hypotes 2 - FEARS-index och avkastning

I tabell 4 presenteras resultatet för regressionerna mellan avkastningen och FEARS-indexen som är skapade med korrelationsmetoden. Vi börjar med att redogöra och analysera resultatet för Stockholmsbörsen och First North Sweden i tidpunkt t . Detta motsvarar kolumn 1 och 3 i tabell 4. Därefter presenterar och analyserar vi dess påverkan för nästkommande vecka, $t+1$, i kolumn 2 och 4. Slutligen presenterar och analyserar vi om den generaliserade linjära regressionen kan visa på någon skillnad mellan de två marknaderna, både i tidpunkt t och $t+1$. Detta motsvarar kolumn 5 och 6 i tabell 4. Analyserna görs löpande för att lättare knyta an till resultatet.

4.2.1. FEARS-index och avkastning i tidpunkt t

I detta avsnittet undersöker vi resultaten i kolumn 1 och 3 i tabell 4. Resultaten visar att FEARS-indexet kan förutspå samma veckas stängningskurs på både Stockholmsbörsen och First North Sweden. Vi kan således förkasta hypotes H2a och H2b med trestjärnig signifikans. Likväl kan vi utläsa positiva samband mellan de två marknaderna och dess respektive FEARS-index. En 1%-ökning i $FEARS_{OMXS}$ ger en ökning med ungefär 0,029% i den beroende variabeln $OMXS$ samma vecka. På liknande vis ökar $FNSAS$ med ungefär 0,025% vid en 1%-ökning i $FEARS_{FNSAS}$ samma vecka. Resultatet för de två marknaderna är därigenom i linje med det resultat Da, Engelberg och Gao (2015) observerade på den amerikanska aktiemarknaden.

Tabell 4: Regressionsstabell för aktieutveckling på Stockholmsbörsen och First North Sweden för att pröva hypotes H2a, H2b och H2c

Avkastning	Aktiemarknad					
	Stockholmsbörsen		First North Sweden		Skillnad	
	<i>t</i>	<i>t+1</i>	<i>t</i>	<i>t+1</i>	<i>t</i>	<i>t+1</i>
(Konstant)	-0,006059 (0,007209)	0,002661 (0,007635)	-0,004989 (0,007916)	-0,014177. (0,008351)	-0,002993 (0,00190)	-0,00544 (0,00202)
FEARS	0,028655*** (0,004666)	-0,005505 (0,004928)	0,024830*** (0,004315)	-0,008376. (0,004824)		
Skillnad i FEARS					0,007223 (0,006059)	0,00320 (0,00708)
Kontrollvariabler	Ja	Ja	Ja	Ja	Ja	Ja
Förklaringsgrad	0,2119	0,1088	0,2383	0,1220		
Justerad förklaringsgrad	0,1777	0,0593	0,1965	0,0788		
F-värde	6,185***	2,222*	5,695***	2,825**		
Observationer	193	193	193	193	386	386

Signifikansnivåer: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1

Utöver att undersöka hur väl FEARS-indexen, modellernas huvudvariabel, kan påverka aktiekursutveckling, är det även viktigt att studera hur väl modellen som helhet är utformad. Vi gör detta genom att studera den justerade förklaringsgraden, vilken mäter hur stor del av variationen i den beroende variabeln som förklaras av modellens oberoende variabler. Eftersom den icke-justerade förklaringsgraden inte tar hänsyn till den relativa nyttan av att inkludera ytterligare variabler, kommer vi endast utgå från den justerade förklaringsgraden. Till skillnad från dess icke-justerade motsvarighet, använder denna en penalisering för antalet variabler som inkluderas i modellen. För regressionerna på Stockholmsbörsen och First North Sweden erhåller vi justerade förklaringsgrader som överstiger 17 respektive 19 %.

Det går även att undersöka hur väl modellen som helhet förklarar den beroende variabeln genom att undersöka modellens F-värde. För både regressionen för Stockholmsbörsen och First North Sweden kan vi med säkerhet säga att de angivna modellerna är signifikanta. Med F-värden motsvarande 6,185 respektive 5,964 erhåller vi trestjärniga signifikanser. Vid jämförelser mot tidigare studier av hur väl GSVI kan påverka aktiekursutveckling håller våra modeller, relativt sett, hög nivå (se Brochado, 2020; Da, Engelberg & Gao, 2015). Med tanke på att FEARS-indexen har skapats med hjälp av de 30 mest signifikant korrelerade orden för aktiekursutvecklingen blir det logiskt att vi erhåller ett signifikant samband i nutid. För att FEARS-indexet ska vara mer användbart i predikterande syfte behöver man dock undersöka dess samband med framtida värden.

4.2.2. FEARS-index och avkastning i tidpunkt $t+1$

Omvänd kausalitet bör diskuteras eftersom det kan vara osäkert i vilken riktning som sambandet mellan FEARS och den beroende variabeln går. Det är inte tydligt vad som är orsak och verkan. Av denna anledningen använder vi tidsfördröjningar av de beroende variablerna. I kolumn 2 och 4 i tabell 4 kan vi se att FEARS-indexen för den innevarande veckan, t , påverkar aktiekursförändringar nästkommande vecka, $t+1$. Signifikansnivån är dock svagare jämfört med när vi undersöker det samtida sambandet. Notera dock att vi ser samma prisvändningsmönster som tidigare studier (Aouadi, Arouri & Teulon, 2013; Beer, Herve & Zouaoui, 2013; Joseph, Wintoki & Zhang, 2011; Kim et al., 2019). Sambandet mellan FEARS-indexen och aktieavkastningen för den innevarande veckan är nämligen positivt, medan sambandet med nästkommande vecka är negativt. Eftersom våra FEARS-index framförallt består av positivt laddade ord är tolkningen naturlig. När

investerare söker efter positivt laddade ord i större utsträckning än vanligt har de en positiv sinnesstämning och är optimistiska inför framtiden. Investerares val av ord och priserna drivs upp från dess fundamentala värden. Nästkommande vecka faller priserna och återgår till fundamentala värden. Det går även att tolka detta som en konsekvens av dispositionseffekten eftersom investerarna väljer att inte behålla aktier när de har gått upp i värde (Barberis & Thaler, 2003). Resultatet är följaktligen i linje med teorier och studier som ifrågasätter effektiva marknadshypotesen (Brown & Cliff, 2005; De Long et al., 1990; Malkiel, 2003; Ritter, 2003; Shaker, 2013) och motstrider Fama (1965; 1970) eftersom tidigare värden har ett samband med framtida.

4.2.3. Skillnaden mellan FEARS-index och avkastning på olika marknader

I tabellens femte kolumn presenteras resultatet från hypotes H2c vid tidpunkt t . Denna utförs med hjälp av en generaliserbar linjär modell, där vi mäter ifall det är skillnad på riktningskoefficienten hos de två tidigare modellerna. Eftersom FEARS-indexen inte är desamma på de två marknaderna finns det anledning att tro att feltermerna hos de två ekvationerna kommer vara korrelerade om vi undersöker skillnaden mellan dem. Modellernas konstanta term eller FEARS-koefficienterna kan därför inte jämföras rakt av utifrån kolumn 1-4. Givet beräkningen från den generaliserbara linjära modellen kan vi fastslå att det inte går att se någon skillnad mellan hur väl FEARS kan påverka aktiekursutveckling på Stockholmsbörsen eller First North Sweden.

Att vi inte kan skilja på hur mycket FEARS-indexen påverkar förändringar på sina respektive marknader talar emot vad tidigare litteratur presenterat om att spekulativa aktier och investerare skulle vara känsligare för investerarsentiment (Baker & Wurgler, 2007; Barber, Odean & Zhu, 2009; Black, 1986; Foucault, Sraer & Thesmar, 2011). Enligt detta resonemang bör vi se en skillnad mellan de två marknaderna, där FEARS-indexet på First North Sweden borde påverka mer än FEARS-indexet på Stockholmsbörsen. Det går att resonera för att detta beror på att arbitrageinvesterare agerar som motvikt mot spekulativa investerare. Trots att de har en mindre andel av sitt kapital på oreglerade marknader jämfört med de spekulativa investerarna, kommer ett högre absolut värde av deras totala kapital väga upp för den inverkan de har på aktiekurserna.

Även om den generaliserbara linjära modellen inte kunde påvisa någon skillnad mellan prediktionsförmågan på Stockholmsbörsen och First North Sweden kan vi jämföra regressionsresultaten för respektive marknad. Riktningskoefficienten på Stockholmsbörsen, 0,028655, är ungefär 15% större än riktningskoefficienten på First North Sweden, 0,024830. Enligt tidigare resonemang kan man förvänta sig att investerarsentiment skulle ha en större inverkan på den oreglerade marknaden, eftersom de svenska hushållen utgör en större andel av denna marknaden. Detta ser vi däremot inget tecken på. Tidigare studier, exempelvis Baker och Wurgler (2006), har även visat att volatila aktier är känsligare för förändringar i investerarsentiment. Vid första anblick verkar det som att detta framförallt berör First North Sweden, eftersom mindre företag är listade där och de tenderar att vara mer volatila. Förvånande nog har aktieindexet OMXS varit mer volatilt än FNSAS under tidsperioden som vi undersökt. För att säkerställa att detta stämmer har vi jämfört olika mått på volatilitet från Bloombergs databas och Nasdaqs hemsida. Vi har själva beräknat realiserad volatilitet, standardavvikelse och använt olika långa tidsperioder. Alla metoder tyder på att OMXS har varit mer volatilt. Detta skulle i sådana fall vara en förklaring till varför FEARS-indexet påverkar större aktiekursförändringar på Stockholmsbörsen än på First North Sweden.

I kolumn 6 kan vi inte heller urskilja en skillnad mellan marknaderna för nästkommande tidsperiod. Detta går att härleda till att vi saknar ett signifikant resultat för Stockholmsbörsen, samt endast har en svag signifikans för First North Sweden. Även om vi inte kan påvisa en signifikant skillnad mellan marknaderna visar det positiva tecknet framför koefficienten att prediktionsförmågan på First North Sweden skulle kunna vara annorlunda från vad vi ser på Stockholmsbörsen. Som sagt är signifikansnivån för svag för att denna skillnad ska kunna säkerställas.

4.3. Hypotes 3 - FEARS-index och volatilitet

I tabell 5 presenteras resultatet för regressionerna mellan volatiliteten och FEARS-indexen. Eftersom resultaten i detta avsnitt inte är signifikanta följer avsnittet en annan disposition. Vi lägger istället större vikt vid att analysera vad som gör att resultaten inte är signifikanta. Vi börjar med att redogöra resultatet för Stockholmsbörsen och First North Sweden i både tidpunkt t och $t+1$. Detta motsvarar kolumn 1-4 i tabell 5. Därefter analyserar vi dessa resultat och argumenterar varför resultaten inte är signifikanta. Slutligen presenterar och

analyserar vi om den generaliserade linjära regressionen kan visa på någon skillnad mellan de två marknaderna, både i tidpunkt t och $t+1$. Dessa motsvarar kolumn 5 och 6 i tabell 5.

4.3.1. FEARS-index och volatilitet i tidpunkt t och $t+1$

I kolumn 1 och 3 i tabell 5 kan vi klargöra att FEARS-indexet inte påverkar volatilitet i den nuvarande tidpunkten, t , på någon av marknaderna. Vi kan således inte förkasta hypotes H3a och H3b. Vi kan dock utvärdera modellen som helhet. Ett F-värde på 3,309 och 5,491 på Stockholmsbörsen respektive First North Sweden, vilket motsvarar en trestjärnig signifikansnivå, innebär att modellen som helhet har ett samband med aktievolatilitet på de två marknaderna. Den justerade förklaringsgraden hos de två modellerna visar därtill hur modellerna förklarar ungefär 11 respektive 19 % av variationen i volatilitet i de två marknaderna.

I kolumn 2 och 4 för tidsperioden $t+1$ undersöks hur väl FEARS-indexen påverkar nästkommande veckas förändringar i volatilitet. Föga förvånande är det varken ett samband med volatilitet eller möjligt att urskilja ett värdningsmönster för någon av marknaderna. Eftersom FEARS-indexen inte kunde förklara en nutida förändring i volatilitet blir det naturligt att indexen inte heller kan prediktera förändringar i kommande tidsperioder.

Tabell 5: Regressionstabell för volatilitet på Stockholmsbörsen och First North Sweden.

Volatilitet	Aktiemarknad					
	Stockholmsbörsen		First North Sweden		Skillnad	
	<i>t</i>	<i>t+1</i>	<i>t</i>	<i>t+1</i>	<i>t</i>	<i>t+1</i>
(Konstant)	-0,01456 (0,07689)	0,032097 (0,07114)	-0,01363** (0,00421)	0,01101** (0,00416)	-0,001703*** (0,00098)	-0,01603*** (0,00093)
FEARS	0,002915 (0,04976)	-0,032802 (0,04612)	0,000221 (0,01180)	0,000418 (0,00178)		
Tidsfördröjd volatilitet (1 vecka)	-0,3684*** (0,07177)	-0,42966*** (0,06719)	-0,19476** (0,07153)	0,29982*** (0,07331)		
Tidsfördröjd volatilitet (2 veckor)	-0,2456*** (0,07164)	-0,26485*** (0,06657)	-0,22962** (0,07135)	-0,21007** (0,07049)		
Skillnad i FEARS					0,021754 (0,05235)	0,02217 (0,05353)
Kontrollvariabler	Ja	Ja	Ja	Ja	Ja	Ja
Förklaringsgrad	0,1539	0,284	0,2318	0,2566		
Justerad förklaringsgrad	0,1074	0,2363	0,1896	0,2115		
F-värde	3,309***	5,950***	5,491***	5,680***		
Observationer	193	193	193	193	386	386

Signifikansnivåer: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1

4.3.2. Analys av FEARS-index och volatilitet

Eftersom vi inte kan urskilja ett samband med volatilitet strider vårt resultat mot studier på amerikanska reglerade marknader (Da, Engelberg & Gao, 2015) samt andra europeiska reglerade marknader (Aouadi, Arouri & Teulon, 2013; Kim et al., 2019). Vi har identifierat fyra anledningar till att vi inte kan påvisa ett signifikant resultat. För det första kan FEARS svaga samband med volatilitet bero på att dess påverkan på volatilitet är kortvarig, och därför kommer det inte fångas av en veckas genomsnittlig volatilitet. Da, Engelberg och Gao (2015) fann exempelvis endast ett statistiskt signifikant samband med samtida volatilitet, baserat på intradagsdata, och inga signifikanta samband med kommande dagar. Därför kan det anses vara väntat att vi inte finner något signifikant samband när vi använder veckodata.

För det andra är FEARS-indexens tidsperiod präglad av en relativt jämn börsuppgång med få negativa ord. Detta kan medföra att dess förmåga att förutse volatilitet försvagas i och med att perioder med hög volatilitet oftast sker i samband med börsnedgångar, som observerat av Francq och Zakoian (2010).

För det tredje kan konstruktionen av FEARS-indexen ha inverkan på dess förmåga att påverka volatilitet. Till skillnad från Da, Engelberg och Gao (2015) består vårt index av merparten positiva ord i stället för enbart negativa. Våra FEARS-index bör därför uppvisa andra egenskaper. Detta förklarar varför vi får en positiv korrelation med aktieavkastningen medan Da, Engelberg och Gao (2015) har ett negativt samband. De har även en positiv korrelation med sitt FEARS-index och volatilitet. Vi skulle således kunna förvänta oss ett negativt samband mellan vårt FEARS-index och volatilitet. Eftersom tidigare studier har undersökt sambandet mellan negativt sentiment och volatilitet, och inte positivt sentiment som vi har gjort, är det därför möjligt att framförallt negativt sentiment har ett samband med volatilitet.

Slutligen är avsaknaden av ett svenskt VIX-index som kontrollvariabel en möjlig faktor. VIX-index har en tydligare koppling till investerares förväntningar genom att ta hänsyn till optioner på S&P 100-indexet. Konjunkturinstitutets indikatorer fångar förväntningar från icke-investerare och dess effekt som kontrollvariabel kan därför vara svagare.

4.3.3. Skillnaden mellan FEARS-index och volatilitet på olika marknader

Vi kan inte konstatera att det förekommer någon skillnad på modellernas riktningskoefficienter. Det går alltså inte att påvisa någon skillnad i FEARS-indexens påverkan på de två marknaderna. Detta gäller både för nuvarande tidsperiod t i kolumn 5 och nästkommande tidsperiod $t+1$ i kolumn 6.

Vi har tidigare diskuterat anledningar till varför FEARS-indexen inte kan förutspå förändringar i volatilitet, varken för nuvarande eller kommande tidsperiod. Det återstår att identifiera bakomliggande orsaker till varför vi inte heller kan se någon skillnad mellan de två typerna av marknader. För det första är det möjligt att FEARS-index inte kan prediktera någon förändring alls, vilket innebär att de därför inte heller kan påvisa någon skillnad mellan marknaderna. För det andra går det att resonera att skillnaden i informationsasymmetri på reglerade och oreglerade marknader i Sverige inte är tillräckligt stor för att ge uttryck som en signifikant skillnad. Det är även möjligt att det förekommer en kombination av båda alternativen.

Som tidigare nämnt i analysen av sambandet mellan FEARS-indexen och avkastning är det oväntat att volatiliteten var högre på Stockholmsbörsen, och vi har använt olika källor och metoder för att validera detta. Även om allt indikerar att Stockholmsbörsen haft högre volatilitet återstår fortfarande en osäkerhet i frågan eftersom det inte känns intuitivt. Detta skulle också kunna vara en förklaring till varför vi inte kan prediktera volatilitet, även om mycket pekar på att det inte är fel i datamaterialet.

4.4. Hypotes 4 - FEARS-index och handelsvolym

I följande avsnitt presenteras resultatet för regressionerna mellan handelsvolym och FEARS-indexen. Avsnittet följer samma disposition som för hypotes 2. Vi börjar med att redogöra och analysera resultatet för Stockholmsbörsen och First North Sweden i tidpunkt t . Detta motsvaras av kolumn 1 och 3 i tabell 6. Därefter presenterar och analyserar vi prediktionsförmågan för nästkommande vecka, $t+1$, i kolumn 2 och 4. Slutligen presenterar och analyserar vi om den generaliserade linjära regressionen kan visa på någon skillnad mellan de två marknaderna, både i tidpunkt t och $t+1$. Dessa motsvarar kolumn 5 och 6 i tabell 6. Analyserna görs löpande för att lättare knyta an till resultatet.

4.4.1. FEARS-index och handelsvolym i tidpunkt t

Riktningskoefficienterna för FEARS i kolumn 1 och 3 i tabell 6 har ett positivt tecken. Detta visar på ett positivt samband mellan FEARS och den beroende variabeln *HVOL*. En 1%-ökning i FEARS motsvarar en 0,56843 % ökning på Stockholmsbörsen och 0,37123 % på First North Sweden. Riktningskoefficienterna har en trestjärnig signifikans på båda marknaderna. Detta betyder att vi kan förkasta hypotes H4a och H4b. FEARS kan således användas för att förutspå förändringar i handelsvolymen på både en reglerad och oreglerad marknad i Sverige. Modellens justerade förklaringsgrad visar att så mycket som 35,65 % av variationen i handelsvolym förklaras av de oberoende variablerna på Stockholmsbörsen och 19,65 % på First North Sweden. Resultatet ligger i linje med studier som visar hur en ökad sökaktivitet resulterar i ökad handelsvolym (Aouadi, Arouri & Teulon, 2013; Joseph, Wintoki & Zhang, 2011; Kim et al., 2019; Takeda & Wakao, 2014). Det betyder således att GSVI kan förklara nutida ökningarna i handelsvolym.

Tabell 6: Regressionstabell för handelsvolym på Stockholmsbörsen och First North Sweden

Handelsvolym	Aktiemarknad					
	Stockholmsbörsen		First North Sweden		Skillnad	
	<i>t</i>	<i>t+1</i>	<i>t</i>	<i>t+1</i>	<i>t</i>	<i>t+1</i>
(Konstant)	-0,21421* (0,10078)	0,12636 (0,09171)	-0,23079*** (0,04562)	-0,17601*** (0,04841)	-0,18241*** (0,03449)	-0,07617* (0,03697)
FEARS	0,56843*** (0,06648)	-0,23649*** (0,05942)	0,37123*** (0,06783)	0,12247. (0,07371)		
Tidsfördröjd handelsvolym (1 vecka)	-0,21436*** (0,06219)	-0,40516*** (0,06976)	-0,29864*** (0,06831)	-0,34371*** (0,07129)		
Tidsfördröjd handelsvolym (2 veckor)	-0,22186*** (0,06101)	-0,21681** (0,07176)	-0,14334* (0,06858)	-0,19792** (0,07135)		
Skillnad i FEARS					0,23034* (0,09684)	-0,28258** (0,09800)
Kontrollvariabler	Ja	Ja	Ja	Ja	Ja	Ja
Förklaringsgrad	0,3902	0,2356	0,2383	0,2255		
Justerad förklaringsgrad	0,3565	0,1892	0,1965	0,1739		
F-värde	7,452***	5,073***	5,695***	4,368***		
Observationer	193	193	193	193	386	386

Signifikansnivåer: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1

4.4.2. FEARS-index och handelsvolym i tidpunkt $t+1$

I kolumn 2 kan vi se ett vändningsmönster på Stockholmsbörsen eftersom riktningskoefficienten för FEARS har övergått till ett negativt tecken i tidpunkt $t+1$. Detta tolkas som att när FEARS-indexet antar höga värden kommer handelsvolymen vara högre den innevarande veckan och lägre nästkommande vecka. På First North Sweden är mönstret avtagande istället för en distinkt vändning eftersom riktningskoefficienten inte byter tecken i kolumn 4. Det är fortfarande tydligt att förändringen i handelsvolym är betydligt större i tidpunkt t jämfört med tidpunkt $t+1$. Signifikansnivån i tidpunkt $t+1$ är högre på Stockholmsbörsen än på First North Sweden som endast är svagt signifikant.

Likt tidigare studier (Da, Engelberg & Gao, 2015; Tetlock, 2007) finns det ett samband mellan sentiment och handelsvolym på Stockholmsbörsen. För att förstå på vilket sätt FEARS-indexen påverkar handelsvolym i nästföljande tidsperiod behöver vi återigen bryta ned beroendeförhållandet mellan de två variablerna. Eftersom båda variablerna mäts utifrån förändringen gentemot föregående värde kan det negativa tecknet framför koefficienten för FEARS på Stockholmsbörsen uppstå ur två scenarion. Det första händelseförloppet, givet att vi även studerar riktningen vid tidpunkt t , kan innebära att en ökning i GSVI leder till att volymen går upp vid nuvarande tidsperiod och sjunker under nästföljande tidsperiod. Detta skulle betyda att handelsvolymen återgår till vad som kan ses som ett normalt värde. Det andra scenariot är istället att det förekommer en viss fördröjning i ökningen av handelsvolym samtidigt som GSVI återgår till ett normalt värde. Oavsett scenario skulle vi erhålla ett negativt samband mellan FEARS och handelsvolym för tidpunkt $t+1$.

Båda scenarierna kan förekomma samtidigt, men en kan vara mer frekvent än den andra. För att förstå vilket alternativ som är mest korrekt behöver vi studera de tidsfördröjda kontrollvariablerna för handelsvolymen. Variablerna som ursprungligen introducerades till modellen för att motverka autokorrelation kommer nu förklara orsaken till vändningsmönstret. Det negativa signifikanta sambandet mellan tidigare veckas handelsvolym innebär att en ökning i handelsvolym tenderar att följas av en minskning, allt annat lika. Detta betyder därför att vi inte har en fördröjning i handelsvolym, det vill säga inte scenario två. Istället kan vi slå fast att handelsvolymen kommer minska för att återgå till vad som kan anses som normal nivå.

4.4.3. Skillnaden mellan FEARS-index och volatilitet på olika marknader

Vår sista hypotes, H4c, att det inte är skillnad mellan hur väl FEARS-indexen påverkar handelsvolym på Stockholmsbörsen respektive First North Sweden. I kolumn 5 och 6 kan vi se att skillnaden mellan de två marknaderna är signifikant, både för den nuvarande tidpunkten t och nästkommande tidpunkt, $t+1$. Vi kan därför förkasta hypotes H4c.

I kolumn 5 undersöks ifall det förekommer en skillnad i de två regressionernas riktningskoefficienter vid nuvarande tidpunkt. Med en enstjärnig signifikans kan nollhypotesen för H4c, att FEARS-indexen skulle förutspå lika stora förändringar i handelsvolym på Stockholmsbörsen och First North Sweden, förkastas. Vad som däremot gör resultatet intressant är att den positiva koefficienten innebär att FEARS-indexet påvisar större förändringar på Stockholmsbörsen än på First North Sweden. Det talar alltså emot tidigare resonemang om att investerare på oreglerade marknader skulle vara mer påverkbara. Vändningsmönstret på Stockholmsbörsen vid tidpunkt $t+1$ förklarar i sin tur varför vi ser prov på ett negativt tecken över skillnaden mellan de två marknaderna vid denna tidpunkt. Den förstärkta signifikansnivån mellan de två marknaderna beror därför på att vi ser olika beteenden. Stockholmsbörsen har ett vändningsmönster vid tidpunkt $t+1$, medan First North Sweden endast har ett avtagande mönster.

4.5. FEARS-indexets robusthet - AIC-modellen

Reliabiliteten för resultaten av hypotes 2, 3 och 4 prövas med hjälp av ett robusthetstest. Vi utför samma regressioner som tidigare i detta avsnittet, men FEARS-indexet är ersatt med ett FEARS-index som är konstruerat utifrån AIC istället för korrelationsmetoden. I tabell 7 återfinns en kortare sammanställning av regressionsresultat från regressioner där FEARS-indexen är konstruerade med hjälp av AIC. Vi ser att våra resultat och analyser håller även i en AIC-modell. Givetvis skiljer sig värdena åt jämfört med tidigare modeller, men överlag är värdena och signifikansnivåerna i paritet med koefficienterna i tidigare modeller. Vi kan alltså konstruera huvudvariabeln FEARS med en annan teknik och fortfarande få samma resultat. Detta blir belägg för att resultaten inte bara är en tillfällighet, vilket ökar studiens validitet och reliabilitet.

Tabell 7. Regressionsresultat där FEARS-indexen är konstruerade med hjälp av AIC.

	Aktiemarknad					
	Stockholmsbörsen		First North Sweden		Skillnad	
	H2a		H2b		H2c	
	t	t+1	t	t+1	t	t+1
FEARS	0,02048**	-0,00468	0,03037***	-0,00168	0,02839***	-0,00722
	(0,00714)	(0,00706)	(0,00752)	(0,00781)	(0,00761)	(0,00498)
Justerad förklaringsgrad	0,05159	0,05566	0,102	0,02934		
F-värde	2,305.	2,132.	3,725***	1,725		
	H3a		H3b		H3c	
	t	t+1	t	t+1	t	t+1
FEARS	0,01188	-0,02853	0,00448	-0,00193	0,00360	-0,00205
	(0,07083)	(0,06588)	(0,00343)	(0,00349)	(0,00379)	(0,00379)
Justerad förklaringsgrad	0,1075	0,23	0,1971	0,166		
F-värde	3,312***	6,214***	5,712***	4,821***		
	H4a		H4b		H4c	
	t	t+1	t	t+1	t	t+1
FEARS	0,60638***	-0,23425**	0,61239***	0,28707*	0,15978	0,15978
	(0,07892)	(0,08620)	(0,13191)	(0,13619)	(0,13217)	(0,13217)
Justerad förklaringsgrad	0,32	0,1563	0,1633	0,1846		
F-värde	9,987***	4,556***	4,748***	4,952***		

Signifikansnivåer: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1

4.6. Diagnostiska tester

Alla regressionsmodeller har genomgått tester för att säkerställa att modellernas antaganden uppfylls. En sammanställning av dessa resultaten återfinns i tabell 8. Alla antaganden, förutom antagandet om normalfördelade residualer, uppfylls i alla regressioner. Detta var väntat eftersom det inom empiriska studier är ovanligt att residualerna följer en normalfördelning. Fördelningens svansar brukar istället ofta vara tjockare på grund av att extremvärden är mer vanligt förekommande än vad normalfördelningen antyder (Rachev, 2003).

Tabell 8. Resultat av de diagnostiska testerna. Rubrikerna visar vilka test som utförts, vilket antagande som är förknippat med testet, samt testets nollhypotes. Gråmarkerade celler indikerar antaganden som inte uppfylls.

<i>Test</i>		<i>VIF-värde</i>	<i>Ramsey RESET</i>	<i>Breusch-Pagan</i>	<i>Box-Jenkins</i>	<i>Jarque-Bera</i>	<i>Dickey-Fuller</i>
<i>Antagande</i>		<i>Multikollinearitet</i>	<i>Linjärt samband</i>	<i>Homoskedasticitet</i>	<i>Ingen autokorrelation</i>	<i>Normalfördelade residualer</i>	<i>Stationäritet</i>
<i>H₀</i>		-	<i>Linjärt samband</i>	<i>Homoskedasticitet</i>	-	<i>Normalfördelade residualer</i>	<i>Icke-stationär</i>
H2a	t	<1,4	0,5877	0,2382	ARMA (0,0)	0,0120	<0,01
	t+1	<1,4	0,4800	0,1309	ARMA (0,0)	0,0032	<0,01
H2b	t	<1,4	0,8883	0,2073	ARMA (0,0)	2,47E-04	<0,01
	t+1	<1,4	0,2737	0,1068	ARMA (0,0)	0,0027	<0,01
H3a	t	<1,5	0,0934	0,9098	ARMA (0,0)	1,51E-10	<0,01
	t+1	<1,5	0,2327	0,6768	ARMA (0,0)	4,70E-11	<0,01
H3b	t	<1,5	0,3201	0,0727	ARMA (0,0)	2,20E-16	<0,01
	t+1	<1,5	0,1688	0,4428	ARMA (0,0)	3,99E-15	<0,01
H4a	t	<1,5	0,9590	0,7391	ARMA (0,0)	3,88E-06	<0,01
	t+1	<1,5	0,4970	0,2762	ARMA (0,0)	6,63E-04	<0,01
H4b	t	<1,5	0,5317	0,5162	ARMA (0,0)	4,88E-06	<0,01
	t+1	<1,5	0,4015	0,1838	ARMA (0,0)	6,39E-07	<0,01

5. SLUTSATSER OCH DISKUSSION

Vi avslutar med att koppla samman våra resultat med studiens syfte och redogör för uppsatsens kunskapsbidrag. Vi resonerar även kring potentiella frågeställningar för vidare forskning baserat på våra resultat.

5.1. Slutsatser

Till att börja med har vår studie bidragit med en bredare förståelse för investerarsentiment på den svenska aktiemarknaden, vilket är ett forskningsområde med få bidrag sedan tidigare. Exempelvis har vi kommit fram till att positivt sentiment väger tyngst i den tidsperioden vi har undersökt och att det kan vara viktigt att skilja på reglerade och oreglerade marknader i vidare forskning. Detta är relevanta kunskapsbidrag när aktiemarknaden blir mer och mer tillgänglig för icke-professionella investerare. Vi har även visat att Google Trender kan vara ett användbart verktyg inom forskningen, vilket är relevant eftersom Google Trender är ett billigt och anpassningsbart verktyg med en stor databas som är tillgänglig för alla.

Eftersom FEARS-index kan förutspå prisförändringar och förändringar i handelsvolym betyder det att FEARS-indexet är ett meningsfullt mått på investerarsentiment. Vi har således bidragit med ett verktyg för den som vill känna av sinnesstämningen på marknaden. Verktöget är bland annat billigare, enklare och mer tillgängligt än andra mått på investerarsentiment.

5.2. Diskussion

Våra resultat kan vara förvridna om det finns utelämnade variabler som hade haft en förklarande effekt om de hade inkluderats. I vårt fall kan utelämnade variabler med en förklarande kraft vara ett VIX-index som är anpassat för just den svenska marknaden. Som tidigare nämnt är kontrollvariablerna från Konjunkturinstitutet inte uttryckligen riktade mot aktiemarknaden, utan samhället i stort. Om det hade funnits andra mått på svenska investerares sinnesstämning finns det anledning att tro att dessa skulle ha haft en större förklarande kraft än Konjunktursinstitutets mätningar. Eftersom dessa variabler inte finns kommer deras effekter istället hänföras till FEARS-indexen och kontrollvariablerna som vi har använt. På detta sättet kan våra resultat vara förvridna.

Det finns ett antagande i regressionsmodellerna som vi inte har testat i denna uppsatsen, och det är antagandet om parameterstabilitet. Enkelt uttryckt handlar parameterstabilitet om att olika urval av datamaterialet uppvisar samma egenskaper. Variabler och resultat är således stabila även om undersökningen bara utgår från en del av det tillgängliga datamaterialet. Inom tidsserieanalys brukar detta testas med hjälp av ett *Chow*-test. Testet förutsätter att det finns en förutbestämd brytpunkt som delar in datamaterialet, vilket vanligtvis är större historiska händelser som kan påverka variablerna. Anledningen till varför vi inte har utfört något *Chow*-test är för att det inte finns någon uppenbar historisk händelse som kan vara en brytpunkt i vår tidsserie, eftersom vi har använt en relativt kort tidsperiod. En kortare tidsperiod brukar innebära färre potentiella brytpunkter. Vi har även testat att göra olika indelningar av datamaterialet utan att hitta en uppenbar brytpunkt. Detta talar för att våra parametrar är konstanta över tidsperioden, även om vi inte har utfört ett *Chow*-test.

I diskussionen kring studiens validitet kan det vara värt att studera tabellerna i appendix 2 med ord som ingår i respektive FEARS-index och positivt och negativt laddade ord. Kan man tänka sig att det är dessa orden som investerare söker efter, och finns det en uppenbar skillnad mellan de ord som vi identifierat som positivt laddade jämfört med de som ska vara negativt laddade? Om vi börjar med att studera orden som ingår i FEARS-indexen verkar de flesta orden vara rimliga sökord som en sentimental investerare kan tänkas använda. Vi har ord och fraser såsom "Aktie", "Avanza", "Aktier", "Nedgång", "Avkastning", "Arbetsförmedlingen jobb" och "Tjäna pengar" som kan tänkas fånga det vi är ute efter och det tycks inte vara en ren slump att dessa orden har identifierats. Samtidigt finns det ord såsom "Nätverk", "Leda", "Stena" och "Producent" som har en otydligare koppling till investerarsentiment. Detta kan vara en konsekvens av att baslistan är väldigt bred, och om man undersöker tillräckligt många ord kommer man hitta några irrelevanta ord som råkar bli signifikanta av en slump. Vi diskuterar just sökordens signifikansnivå senare.

Om vi istället undersöker de positivt och negativt laddade orden på respektive marknad ser vi att flera ord har tilldelats en rimlig kategori. Exempelvis är "Avanza" med bland de positiva orden och "Nedgång" är ett starkt negativt ord. Överlag verkar dock skillnaden inte vara alltför tydlig. Hade man kunnat gissa sig till vilken lista som var positiv och vilken som var negativ om rubriken inte redan var satt? Möjligtvis om man endast hade studerat de mest signifikanta orden i respektive lista, men ju fler man jämför desto svårare blir det att se en

tydlig skillnad mellan positiva och negativa ord. Det beror på att orden är sorterade efter signifikansnivå i fallande skala, och orden längst ner i listan har p-värden som ofta motsvarar en svag signifikans. Deras t-värden är inte nödvändigtvis, med statistisk säkerhet, skilda från noll. Därför är det rimligt att det är svårare att se skillnad på vad som ska vara positivt laddade ord och negativt laddade ord ju längre ner i listan orden förekommer. Vi diskuterar konsekvenserna av detta i det följande.

Det bör noteras att även om Wilcoxon rank sum-testerna i avsnitt 4 kan påvisa en signifikant skillnad mellan fördelningarna av de 30 mest signifikanta orden, är orden *i sig* inte alltid signifikanta. De valda orden är som nämnt de 30 *mest* signifikanta orden, men detta betyder inte att alla de 30 orden har en hög signifikansnivå. T-värdena som finns tabell A.3 i appendix 2 och i figur 5 motsvarar p-värden mellan $<0,1\%$ för de mest signifikanta orden och 61% för de minst signifikanta orden. Märk väl att det fortfarande är meningsfullt att inkludera ord som har höga p-värden i Wilcoxon rank sum-testerna eftersom p-värdet endast mäter signifikansnivån för skattningen av själva korrelationskoefficienten. I syftet att jämföra fördelningarna av t-värden är det fortfarande meningsfullt att inkludera ord som har höga p-värden eftersom t-värdena fortfarande är ett mått på hur starkt samband respektive ord har med marknadsavkastningen. Det är alltså viktigt att skilja på signifikansnivån för den skattade korrelationskoefficienten och fördelningen av t-värden. Höga p-värden är istället något som borde fått konsekvenser för hur väl FEARS-indexen lyckas förklara aktiviteten på aktiemarknaden.

5.3. Vidare forskning

Det kan vara intressant att skapa FEARS-index med en lägre datafrekvens. Månadsdata kan exempelvis användas för att undersöka om ackumulerat sentiment för en hel månad får mer långtgående effekter på aktiemarknaden. Potentiella användningsområden är att undersöka den långsiktiga prestationen för börsnoteringar, uppköp, sammanslagningar eller avknoppningar. Här överlappar även användningsområdet med mikro- och makroekonomi. Vidare forskning kan exempelvis undersöka om det går att skapa FEARS-index som predikterar förändringar i mikro- och makroekonomiska variabler såsom husköp, sparande och arbetstillfällen.

Ett argument från vår analys är att konjunkturcykeln kan vara avgörande för om positivt eller negativt sentiment bäst förutspår utvecklingen på marknaden. Vidare forskning kan bygga vidare på detta resonemanget och dela in tidsperioder i olika faser, exempelvis uppgång, nedgång, stagnation och krissituation. Därefter kan man utvärdera FEARS-indexets egenskaper i de olika faserna för att få en bättre förståelse för när vilken sorts sentiment väger tyngst.

Det första steget i forskningen var att visa att FEARS-indexet går att använda på en svensk aktiemarknad. Detta har vi gjort genom att tillverka relativt breda FEARS-index som mäter de bredaste aktiekursindexen på Stockholmsbörsen och First North Sweden. När detta är gjort är det bara fantasin som sätter gränserna för verktygets användningsområden eftersom användaren själv kan bestämma vilka sökord som ska studeras. Nästa steg kan då vara att tillverka mer nischade index, till exempel för en given sektor, ett givet företag eller för en annan sorts värdepapper än aktier. Det är i synnerhet intressant att undersöka tillgångar som är särskilt känsliga för förändringar i investerarsentiment, såsom kryptovalutor. Användningsområdet kan kanske vara att vidareutveckla en investeringsstrategi baserat på förändringar i sentiment. Kanske är det så att ju mer specialiserat FEARS-indexet är på ett visst instrument, desto mer användbart är det för att skapa investeringsstrategier. Det är något som den som söker skall finna.

KÄLLFÖRTECKNING

Aouadi, A., Arouri, M., & Teulon, F. (2013). Investor Attention and Stock Market Activity: Evidence from France, *Economic Modelling*, vol. 35, sid. 674-681, tillgänglig via

LUBsearch:

<https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S0264999313003507?via%3Dihub> [hämtad 12 april 2021]

Armelius, H., Hull, I. & Stenbacka Köhler, H. (2016). The timing of uncertainty shocks in a small open economy. *Sveriges Riksbank Working Paper Series*, no. 334, tillgänglig online:

<https://www.riksbank.se/globalassets/media/rapporter/working-papers/2016/no.-334-the-timing-of-uncertainty-shocks-in-a-small-open-economy.pdf> [hämtad 23 maj 2021]

Athanasopoulos, G. & Hyndman, R. J. (2018). *Forecasting: Principles and Practice*, 2:a upplagan, Otexts

Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-section of Stock Returns, *The Journal of Finance*, vol. 61, no. 4, sid. 1645-1680, tillgänglig online:

<https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/j.1540-6261.2006.00885.x> [hämtad 15 april 2021]

Baker, M., & Wurgler, J. (2007). Investor Sentiment in The Stock Market, *Journal of Economic Perspectives*, vol. 21, no. 2, sid. 129-152, tillgänglig online:

<https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.21.2.129> [hämtad 15 april 2021]

Baltagi, B. H. (2011). *Econometrics*, 5:e upplagan, New York, NY: Springer New York

Barberis, N., & Thaler, R. (2003). A Survey of Behavioral Finance, i George M.

Constantinides, Milton Harris, & Rene Stulz (eds), *Handbook of the Economics of Finance*, Amsterdam: Elsevier, sid. 1053-1128, tillgänglig via LUBsearch:

https://www-nber-org.ludwig.lub.lu.se/system/files/working_papers/w9222/w9222.pdf

[hämtad 16 april 2021]

Barber, B. M., & Odean, T. (2008). All That Glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *The Review of Financial Studies*, vol. 21, no. 2, sid. 785-818, tillgänglig via LUBsearch:

<http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bt&AN=31859884&site=eds-live&scope=site> [hämtad 16 april 2021]

Barber, B. M., Odean, T., & Zhu, N. (2009). Systematic Noise, *Journal of Financial Markets*, vol. 12, no. 4, sid. 547-569, tillgänglig via LUBsearch:

<https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S1386418109000196?via%3Dihub> [hämtad 16 april 2021]

Baumeister, R., Bratslavsky, E., Finkenauer, C., & Vohs, K. (2001). Bad is Stronger than Good, *Review of General Psychology*. vol. 5, no. 4. sid. 323-370, tillgänglig online

https://www.researchgate.net/publication/46608952_Bad_Is_Stronger_than_Good [hämtad 13 april 2021]

Beer, F., Hervé, F., & Zouaoui, M. (2013). Is Big Brother Watching Us? Google, investor Sentiment and the stock market, *Economics Bulletin*, vol. 33, no. 1, sid. 454-466, tillgänglig vid LUBsearch:

<https://www-scopus-com.ludwig.lub.lu.se/record/display.uri?eid=2-s2.0-84882962545&origin=inward&txGid=0905b3eb47a2258a50d1dc2a87a570f0> [hämtad 13 april 2021]

Black, F. (1986). Noise, *The Journal of Finance*, vol. 41, no. 3, sid. 528-543, tillgänglig via LUBsearch: <http://ludwig.lub.lu.se/login?url=https://www.jstor.org/stable/2328481> [hämtad 16 april 2021]

Brochado, A. (2020). Google Search Based Sentiment Indexes, *IIMB Management Review*, vol. 32, no. 3, sid. 325-335, tillgänglig via LUBsearch:

<https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S0970389617302719?via%3Dihub> [hämtad 15 april 2021]

Brooks, C. (2014). *Introductory Econometrics for Finance*, 3:e upplagan Cambridge: Cambridge University Press.

Brown, G. W., & Cliff, M. T. (2005). Investor Sentiment and Asset Valuation, *The Journal of Business*, vol. 78, no. 2, sid. 405-440, tillgänglig online:

<http://ludwig.lub.lu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsjsr&AN=edsjsr.10.1086.427633&site=eds-live&scope=site> [hämtad 15 april 2021]

Bryman, A. & Bell, E. (2017). Företagsekonomiska forskningsmetoder. 3:e uppl. Stockholm, Liber AB

Cryer, J. & Chan, K-S. (2008). Time Series Analysis, 2:a upplagan, New York, NY: Springer New York

Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention, *The Journal of Finance*, vol. 66, no. 5, sid. 1461-1499, tillgänglig online:

<https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/j.1540-6261.2011.01679.x> [hämtad 12 april 2021]

Da, Z., Engelberg, J., & Gao, P. (2015). The Sum of all FEARS Investor Sentiment and Asset Prices, *The Review of Financial Studies*, vol. 28, no. 1, sid. 1-32, tillgänglig online:

<https://academic.oup.com/rfs/article/28/1/1/1682440> [hämtad 12 april 2021]

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets, *Journal of Political Economy*, vol. 98, no. 4, sid. 703-738, tillgänglig via LUBsearch:

http://www.nccr-finrisk.uzh.ch/media/pdf/DeLongShleiferSummersWaldmann_JPE1990.pdf [hämtad 16 april 2021]

Dougherty, C. (2016). Introduction to Econometrics, 5e upplagan, Oxford: Oxford University Press

European Journal of Business and Social Sciences. (n.d.) Journal Description, tillgänglig online: <https://ejbss.org/index.php/ejbss> [hämtad 2021-05-27]

Fama, E. F. (1965). The Behavior of Stock-Market Prices, *The Journal of Business*, vol. 38, no. 1, sid. 34-105, tillgänglig via LUBsearch:

<https://www-jstor-org.ludwig.lub.lu.se/stable/2350752> [hämtad 16 april 2021]

Fama, E. F. (1970). Efficient Capital Markets: A review of theory and empirical work, *The Journal of Finance*, vol. 25, no. 2, sid. 383-417, tillgänglig via LUBsearch:

https://www-jstor-org.ludwig.lub.lu.se/stable/2325486?seq=1#metadata_info_tab_contents
[hämtad 12 april 2021]

Foucault, T., Sraer, D., & Thesmar, D. J. (2011). Individual Investors and Volatility, *The Journal of Finance*, vol. 66, no. 4, sid. 1369-1406, tillgänglig via LUBsearch:

<http://ludwig.lub.lu.se/login?url=https://www.jstor.org/stable/29789818> [hämtad 16 april 2021]

Franco, C., & Zakoian, J., (2010). Inconsistency of the MLE and inference based on weighted LS for LARCH models, *Journal of Econometrics*, vol. 159, no. 1, sid 151-165, tillgänglig online: <https://www.sciencedirect.com/science/article/pii/S0304407610001284> [hämtad 19 maj 2021]

General Inquirer. (n.d.) Harvard IV-4 Dictionary och Lasswell Value Dictionary, tillgänglig online: <http://www.wjh.harvard.edu/~inquirer/> [hämtad 27 maj 2021]

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*, 19 februari, tillgänglig online: <https://www.nature.com/articles/nature07634> [hämtad 14 maj 2021]

Goldman Sachs. (2009). Alternative Trading Systems, tillgänglig online:

<https://www.goldmansachs.com/insights/archive/alt-trading-sys.html> [hämtad 30 april 2021]

Google. (2021). Vanliga frågor om data i Google Trender, tillgänglig online:

<https://support.google.com/trends/answer/4365533?hl=sv> [hämtad 11 april 2021]

Google. (n.d.) Google trends, tillgänglig online: <https://trends.google.se/trends/> [hämtad 27 maj 2021]

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning, [e-bok] New York, NY: Springer New York, tillgänglig via LUBsearch: <https://link-springer-com.ludwig.lub.lu.se/content/pdf/10.1007%2F978-1-4614-7138-7.pdf> [hämtad 4 maj 2021]

Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting Abnormal Stock Returns And Trading Volume using Investor Sentiment: Evidence from online search, *International Journal of Forecasting*, vol. 27, no. 4, sid. 1116-1127, tillgänglig online: https://wintoki.ku.edu/myssi/_pdf/Online_Search_Returns_IJF_preprint.pdf [hämtad 12 april 2021]

Keynes, J.M. (1936). The General Theory of Employment, Interest, and Money, [e-bok] Cham: Palgrave Macmillan, tillgänglig via LUBsearch: <https://link-springer-com.ludwig.lub.lu.se/content/pdf/10.1007%2F978-3-319-70344-2.pdf> [hämtad 11 maj 2021]

Kim, N., Lučivjanská, K., Molnár, P., & Villa, R. (2019). Google Searches and Stock Market Activity: Evidence from Norway, *Finance Research Letters*, vol. 28, sid. 208-220, tillgänglig via LUBsearch: <https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S1544612317307377> [hämtad 12 april 2021]

Konjunkturinstitutet. (2021). Metodbok för Konjunkturbarometern, tillgänglig online: <https://www.konj.se/metodbok> [hämtad 3 maj 2021]

Lo, A., Mackinlay, C., (1988). Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test, *The Review of Financial Studies*, vol.1, no. 1, sid. 41–66, tillgänglig via: <https://academic.oup.com/rfs/article/1/1/41/1601244?login=true> [hämtad 25 maj 2021]

Malkiel, B. G. (2003). The Efficient Market Hypothesis and its Critics, *Journal of Economic Perspectives*, vol 17, no. 1, sid. 59-82, tillgänglig online: <https://pubs.aeaweb.org/doi/pdfplus/10.1257/089533003321164958> [hämtad 12 april 2021]

Nachar, N. (2008). The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution, *Tutorials in quantitative Methods for Psychology*, vol. 4, no. 1, sid. 13-20, tillgänglig online:

<https://pdfs.semanticscholar.org/007b/c0936646c34abd369ceda930000c3d142228.pdf>

[hämtad 24 maj 2021]

Nasdaq. (2021a). Nordic Main Market Rulebook for Issuers of Shares, tillgänglig online:

<https://www.nasdaq.com/solutions/rules-regulations-stockholm> [hämtad 29 april 2021]

Nasdaq. (2021b). Nasdaq First North Growth Market – Rulebook, tillgänglig online:

<https://www.nasdaq.com/solutions/rules-regulations-first-north-mtf-rules> [hämtad 29 april 2021]

Nasdaq OMX Nordic. (2021). Var handlar man aktier?, tillgänglig online:

<http://www.nasdaqomxnordic.com/utbildning/aktier/varhandlarmanaktier/?languageId=3>

[hämtad 29 april 2021]

NE. (n.d.). Ordbokstjänster från NE, tillgänglig online: <https://ne.ord.se/> [hämtad 27 maj 2021]

OECD (2005a). Trend Component of a Time Series, tillgänglig online:

<https://stats.oecd.org/glossary/detail.asp?ID=6692> [hämtad 26 april 2021]

OECD (2005b). Seasonal Component of a Time Series, tillgänglig online:

<https://stats.oecd.org/glossary/detail.asp?ID=6695> [hämtad 26 april 2021]

OECD (2005c). Irregular Component of a Time Series, tillgänglig online:

<https://stats.oecd.org/glossary/detail.asp?ID=6697> [hämtad 26 april 2021]

Oldstrand, R. (2020). Ett år som går till historien, Swedbank Aktiellt, 29 december, tillgänglig online:

https://www.swedbank-aktiellt.se/2020/december/ett_ar_som_gar_till_historien.csp [hämtad

23 maj 2021]

Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying Trading Behavior in Financial Markets using Google Trends, *Scientific reports*, vol. 3, no. 1, sid. 1684, tillgänglig online: https://www.nature.com/articles/srep01684?__hstc=113740504.2a1e835c34ab7bf88e972fdd7a7debc8.1424476800061.1424476800062.1424476800063.1&__hssc=113740504.1.1424476800064&__hsfp=3972014050 [hämtad 12 april 2021]

Rachev, S. T. (2003). Handbook of Heavy Tailed Distributions in Finance, [e-bok] Nordholland: Elsevier, tillgänglig via LUBsearch: <https://www-sciencedirect-com.ludwig.lub.lu.se/book/9780444508966/handbook-of-heavy-tailed-distributions-in-finance> [hämtad 27 maj 2021]

Riksbanken. (2021). Reporänta, in- och utlåningsränta, tillgänglig online: <https://www.riksbank.se/sv/statistik/sok-rantor--valutakurser/reporanta-in--och-utlaningsranta/> [hämtad 25 maj 2021]

Ritter, J. R. (2003). Behavioral Finance, *Pacific-Basin Finance Journal*, vol. 11, no. 4, sid. 429-437, tillgänglig via LUBsearch: <https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S0927538X03000489?via%3Dihub> [hämtad 16 april 2021]

SCB (2021a). Aktieäggande i bolag noterade på svensk marknadsplats, ställningsvärden, marknadsplatsandelar, procent efter sektor, marknadsplats och 2 ggr/år, tillgänglig online: <https://www.statistikdatabasen.scb.se/sq/107397> [hämtad 14 april 2021]

SCB (2021b). Nationalräkenskaper, BNP (ENS2010), volymförändring. Kvartal 1993K1 - 2020K4, tillgänglig online: https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START_NR_NR0103_NR0103S/NR0103ENS10SnabbStat/ [hämtad 23 maj 2021]

SCB (2021c). Konsumentprisindex (KPI), totalt, 1980=100. Månad 1980M01 - 2021M04, tillgänglig online: https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START_PR_PR0101_PR0101A/KPItotM/ [hämtad 25 maj 2021]

SCImago, (n.d.). About us, tillgänglig online: <https://www.scimagojr.com/help.php> [hämtad 25 maj 2021]

Shaker, A. (2013). Testing the Weak-Form Market Efficiency of the Finnish and Swedish Stock Markets, *European Journal of Business and Social Sciences*, vol. 2, no. 9, sid. 176-185, tillgänglig online:
https://www.researchgate.net/publication/259493181_Testing_the_Weak-Form_Efficiency_of_the_Finnish_and_Swedish_Stock_Markets [hämtad 24 april 2021]

Sheather, S. J. (2009). *A Modern Approach to Regression with R*, 6:e upplagan, Texas: Springer.

Takeda, F., & Wakao, T. (2014). Google Search Intensity and its Relationship With Returns and Trading Volume of Japanese Stocks. *Pacific-Basin Finance Journal*, vol. 27, sid. 1-18, tillgänglig online:
<https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S0927538X14000043>
[hämtad 12 april 2021]

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The role of media in the stock market, *The Journal of Finance*, vol. 62, no. 3, sid. 1139-1168, tillgänglig via LUBsearch:
<http://ludwig.lub.lu.se/login?url=https://www.jstor.org/stable/4622297> [hämtad 15 april 2021]

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than Words: Quantifying language to measure firms' fundamentals, *The Journal of Finance*, vol. 63, no. 3, sid. 1437-1467, tillgänglig via LUBsearch:
<http://ludwig.lub.lu.se/login?url=https://www.jstor.org/stable/25094477> [hämtad 15 april 2021]

Weber Shandwick. (2021). The Gamestop Phenomenon, tillgänglig online:
<https://www.webershandwick.com/wp-content/uploads/2021/02/Reputation-Alert-The-Gamestop-Investor-Phenomenon.pdf> [hämtad 8 maj 2021]

Appendix 1. Baslista med sökord

Tabell A.1: Lista över ord som ingår baslistan.

ADEL	CSN	IMPORTERA	LEVERANTÖR	SEB BANK
AFA	CSN BIDRAG	INDUSTRIELL	LYCKA	SEK TILL DOLLAR
AFA FÖRSÄKRING	CSN LÅN	INFLATION	LYX	SEMESTERERSÄTTNING
AFFÄRER	DAGENS INDUSTRI	INKOMST	LÅN	SILVER
AFFÄRSVÄRLDEN	DEPRESSION	INNEBOENDE	LÅNA	SKULD
AKASSA	DI	INVESTERA	LÅNA PENGAR	SKULDSANERING
AKTIE	DOLLAR EURO	JOBB	LÄGENHETER	SKÄNKA
AKTIEBOLAG	DOLLAR TILL SEK	JOBB GÖTEBORG	LÄNSFÖRSÄKRINGAR	SMS LÅN
AKTIER	DOLLAR SEK	JOBB MALMÖ	LÄNSFÖRSÄKRINGAR BANK	STARTA EGET
ALLA BOLAG	DONATION	JOBB STOCKHOLM	LÖN	STARTA FÖRETAG
ALLABOLAG	DONERA	JURIDISK	LÖN EFTER SKATT	STATLIG SKATT
ANSTÄLLD	EFTER SKATT	KAPITAL	LÖN SKATT	STENA
ARBETSFÖRMEDLINGEN	EKONOMISK	KLARNA	MARKNAD	STÅL
ARBETSFÖRMEDLINGEN JOBB	ENSKILD FIRMA	KOLLEGA	METALL	SWEDBANK LÅN
ARBETSLÖS	EURO	KONKURS	MONOPOL	SÄGA UPP
ARBETSLÖSHET	FINGERPRINT	KONTANTER	NEDGÅNG	SÖKA JOBB
AVANZA	FOREX	KONTRAKT	NORDEA BANK	TILLGÅNG

AVKASTNING	FYND	KOPPAR	NORDEA FÖRETAG	TILLVERKARE
AVSKRIVNING	FÖRDEL	KOSTNAD	NYTTA	TILLÄGG
AVTAL	FÖRETAGS	KR DOLLAR	NÄRINGSLIV	TIMLÖN
BANK LÅN	FÖRSÄKRA	KRIS	NÄTVERK	TJÄNA
BERÄKNA SKATT	FÖRSÄKRING	KRONOFOGDEN	OLJA	TJÄNA PENGAR
BIDRAG	FÖRSÄKRING LÄNSFÖRSÄKRINGAR	KÖP	PANTBREV	UNDERHÅLL
BILFÖRSÄKRING	FÖRSÄKRINGSKASSAN	KÖP OCH SÄLJ	PARTNER	UNDERHÅLLSSTÖD
BILLIG	FÖRSÄLJNING	KÖPA AKTIER	PENGAR	UPPSÄGNINGSTID
BITCOIN	GULD	KÖPA BIL	POLICY	US DOLLAR
BLOCKET JOBB	GÅVA	KÖPA BOSTAD	POLITIK	VAGABOND
BONDE	HANDEL	KÖPA HUND	PRIS GULD	VECKANS AFFÄRER
BONUS	HANDELS	KÖPA HUS	PRISJAKT	VINNA
BOOM	HANDELSBOLAG	KÖPA IPHONE	PRODUCENT	VINST
BRIST	HJÄLPA	KÖPA LÄGENHET	REA	VÄRDE
BUDGET	HUR MYCKET SKATT	LAGFART	RESURS BANK	ZINK
BYGGLOV	HYRA	LEDA	RIK	ÄGA
BÖTER	HYRA BIL	LEDIGA JOBB	RÄKNA SKATT	ÄGARE
CASINO	HYRA STUGA	LEDIGA JOBB ARBETSFÖRMEDLINGEN	RÄKNA UT SKATT	ÖKA
CHECK	IMPORT	LEDNING	RÄNTA LÅN	ÖVERSKOTT

Appendix 2. Testresultat och statistik

Tabell A.2: Lista över ord som ingår i respektive FEARS-index.

<i>FEARS_{OMXS}</i>			<i>FEARS_{FNSAS}</i>		
<i>AFFÄRER</i>	<i>DAGENS INDUSTRI</i>	<i>LÄNSFÖRSÄKRINGAR</i>	<i>AKTIER</i>	<i>EKONOMISK</i>	<i>KRONOFOGDEN</i>
<i>AKTIE</i>	<i>DONATION</i>	<i>LÄNSFÖRSÄKRINGAR BANK</i>	<i>ARBETSFÖRMEDLINGEN JOBB</i>	<i>FÖRSÄKRINGSKASSAN</i>	<i>LEDA</i>
<i>AKTIER</i>	<i>FYND</i>	<i>NEDGÅNG</i>	<i>AVANZA</i>	<i>HANDELS</i>	<i>LEDIGA JOBB</i>
<i>AVANZA</i>	<i>HANDELS</i>	<i>NÄTVERK</i>	<i>AVTAL</i>	<i>HANDELSBOLAG</i>	<i>LÖN</i>
<i>AVKASTNING</i>	<i>HANDELSBOLAG</i>	<i>PENGAR</i>	<i>BRIST</i>	<i>HYRA BIL</i>	<i>PENGAR</i>
<i>BANK LÅN</i>	<i>HYRA BIL</i>	<i>PRODUCENT</i>	<i>BYGGLOV</i>	<i>INFLATION</i>	<i>RESURS BANK</i>
<i>BONUS</i>	<i>KOPPAR</i>	<i>RESURS BANK</i>	<i>CHECK</i>	<i>INNEBOENDE</i>	<i>RÄNTA LÅN</i>
<i>BOOM</i>	<i>KRONOFOGDEN</i>	<i>STENA</i>	<i>DAGENS INDUSTRI</i>	<i>JOBB</i>	<i>STARTA EGET</i>
<i>BRIST</i>	<i>LEDIGA JOBB</i>	<i>TJÄNA PENGAR</i>	<i>DOLLAR TILL SEK</i>	<i>KAPITAL</i>	<i>TJÄNA</i>
<i>CSN</i>	<i>LÅN</i>	<i>UPPSÄGNINGSTID</i>	<i>DONATION</i>	<i>KOPPAR</i>	<i>TJÄNA PENGAR</i>

Tabell A.3: Lista över de 30 mest positiva och 30 mest negativa sökorden för respektive aktiemarknad.

Positiva ord			Stockholmsbörsen			Negativa ord			Positiva ord			First North Sweden			Negativa ord		
T-värde	P-värde	Sökord	T-värde	P-värde	Sökord	T-värde	P-värde	Sökord	T-värde	P-värde	Sökord	T-värde	P-värde	Sökord	T-värde	P-värde	Sökord
1	3,0162	0,0029	DONATION	-2,3653	0,0190	NEDGÅNG	3,5847	0,0004	HANDELSBOLAG	-2,2293	0,0270	HYRA BIL					
2	2,8350	0,0051	HANDELS	-1,6064	0,1098	HYRA BIL	3,0703	0,0024	BYGGLOV	-1,8157	0,0710	DOLLAR TILL SEK					
3	2,7672	0,0062	BRIST	-1,5485	0,1231	HYRA STUGA	2,9972	0,0031	LÖN	-1,6929	0,0921	KÖPA IPHONE					
4	2,4713	0,0143	LÄNSFÖRSÄKRINGAR BANK	-1,4596	0,1460	FÖRDEL	2,8528	0,0048	AVANZA	-1,6899	0,0927	INVESTERA					
5	2,3677	0,0189	RESURS BANK	-1,4220	0,1566	DOLLAR SEK	2,8071	0,0055	LEDIGA JOBB	-1,5034	0,1344	CASINO					
6	2,3655	0,0190	AVKASTNING	-1,4123	0,1595	EURO	2,6982	0,0076	EKONOMISK	-1,4346	0,1530	NEDGÅNG					
7	2,3369	0,0205	HANDELSBOLAG	-1,2901	0,1986	NYTTA	2,6728	0,0082	TJÄNA PENGAR	-1,4259	0,1555	FÖRDEL					
8	2,3120	0,0218	DAGENS INDUSTRI	-1,2526	0,2119	INDUSTRIELL	2,6463	0,0088	HANDELS	-1,4205	0,1571	LÅNA					
9	2,2121	0,0281	AVANZA	-1,2415	0,2159	LÄGENHETER	2,5487	0,0116	JOBB	-1,2594	0,2094	KÖP					
10	2,0416	0,0426	ÄGA	-1,2008	0,2313	BIDRAG	2,5268	0,0123	RÄNTA LÅN	-1,2152	0,2258	BIDRAG					
11	2,0343	0,0433	LÄNSFÖRSÄKRINGAR	-1,0908	0,2767	HUR MYCKET SKATT	2,5109	0,0129	CHECK	-1,1405	0,2555	GULD					
12	1,9741	0,0498	NÄTVERK	-1,0577	0,2915	INVESTERA	2,5080	0,0130	INFLATION	-1,1142	0,2666	FOREX					
13	1,8790	0,0618	PENGAR	-0,9870	0,3249	PRIS GULD	2,4501	0,0152	INNEBOENDE	-1,0147	0,3115	SILVER					
14	1,8643	0,0638	AFFÄRER	-0,8802	0,3798	JOBB MALMÖ	2,4199	0,0165	BRIST	-0,9881	0,3244	REA					
15	1,8052	0,0726	PRODUCENT	-0,8712	0,3847	LEVERANTÖR	2,3854	0,0180	KRONOFOGDEN	-0,9776	0,3295	EURO					
16	1,7718	0,0780	KRONOFOGDEN	-0,8454	0,3990	SEK TILL DOLLAR	2,3180	0,0215	FÖRSÄKRINGSKASSAN	-0,9714	0,3326	NORDEA BANK					
17	1,7643	0,0793	LEDIGA JOBB	-0,8226	0,4117	RIK	2,2618	0,0248	DONATION	-0,9407	0,3481	KÖPA AKTIER					
18	1,7309	0,0851	BONUS	-0,8125	0,4175	BITCOIN	2,2255	0,0272	TJÄNA	-0,9004	0,3690	PRISJAKT					
19	1,7167	0,0877	KOPPAR	-0,7808	0,4359	GULD	2,2063	0,0286	RESURS BANK	-0,8894	0,3749	TILLVERKARE					
20	1,7107	0,0888	LÅN	-0,7715	0,4413	DOLLAR TILL SEK	2,1678	0,0314	DAGENS INDUSTRI	-0,8550	0,3936	VAGABOND					
21	1,6617	0,0982	UPPSÄGNINGSTID	-0,7676	0,4436	FÖRSÄKRING LÄNSFÖRSÄKRINGAR	2,1430	0,0334	KOPPAR	-0,8497	0,3965	DOLLAR SEK					
22	1,6251	0,1058	AKTIER	-0,7676	0,4437	ZINK	2,0560	0,0411	ARBETSFÖRMEDLINGEN JOBB	-0,8425	0,4006	BÖTER					
23	1,6248	0,1059	BANK LÅN	-0,7567	0,4502	KÖPA LÄGENHET	2,0470	0,0420	STARTA EGET	-0,7496	0,4544	HYRA STUGA					
24	1,6078	0,1095	TJÄNA PENGAR	-0,7432	0,4583	KR DOLLAR	1,9709	0,0502	KAPITAL	-0,7000	0,4848	HUR MYCKET SKATT					
25	1,6046	0,1102	AKTIE	-0,7130	0,4767	PARTNER	1,9234	0,0559	PENGAR	-0,6835	0,4951	DOLLAR EURO					
26	1,5911	0,1132	CSN	-0,7074	0,4802	TILLVERKARE	1,9191	0,0565	AKTIER	-0,6475	0,5181	KONTANTER					
27	1,5699	0,1181	FYND	-0,6121	0,5412	RÅKNA UT SKATT	1,8518	0,0656	AVTAL	-0,5743	0,5665	ÖVERSKOTT					
28	1,5669	0,1188	STENA	-0,5832	0,5605	CASINO	1,8170	0,0708	LEDA	-0,5641	0,5733	BITCOIN					
29	1,5518	0,1224	BYGGLOV	-0,5510	0,5823	CSN BIDRAG	1,8125	0,0715	TIMLÖN	-0,5274	0,5985	PARTNER					
30	1,5427	0,1246	KAPITAL	-0,5086	0,6117	JOBB STOCKHOLM	1,7654	0,0791	POLICY	-0,5212	0,6028	IMPORTERA					

Tabell A.4: T-test över differens för antal hämtningar i Google Trends

t	Frihetsgrader	p	Medelvärde	95 % konfidensintervall över differens	
				Nedre	Övre
-0,66618	263	0,5059	-0,02155656	-0,08527148	0,0415835

Tabell A.5: Deskriptiv statistik. Notera att detta är statistik för variablerna i regressionsmodellerna, och inte statistik för de faktiska aktieindexkurserna, volatiliteterna och handelsvolymerna.

Variabel	Definition	Medelvärde	Std. avvikelse	Min	Max
$FEARS_{OMXS}$	Index för investerarsentiment på Stockholmsbörsen baserad på sökmotordata	0,3036	0,2837	-0,828	1,256
$FEARS_{FNSAS}$	Index för investerarsentiment på First North Sweden baserad på sökmotordata	0,2543	0,3181	-1,218	1,45
$OMXS_t$	Aktiekursutveckling på Stockholmsbörsen	0,0010	0,0198	-0,068	0,054
$FNSAS_t$	Aktiekursutveckling på First North Sweden	0,0036	0,022	-0,061	0,075
$Vol_{OMXS,t}$	Volatilitet på Stockholmsbörsen	0,0176	0,0099	0,0047	0,0864
$Vol_{FNSAS,t}$	Volatilitet på First North Sweden	0,0166	0,0103	0	0,06
$HVOL_{OMXS,t}$	Förändring i handelsvolym på Stockholmsbörsen	-0,0237	0,2556	-0,834	0,511
$HVOL_{FNSAS,t}$	Förändring i handelsvolym på First North Sweden	-0,0656	0,3984	-1,593	0,66