



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Trinity College Dublin

Textual Sentiment Analysis in International Financial Markets

by

Sha Liu

A dissertation submitted in fulfillment of the requirements for the
Degree of Doctor of Philosophy in Business Studies

in the

School of Business

Trinity College, Dublin

22 October, 2013

Declaration of Authorship

I declare that this thesis has not been submitted as an exercise for a degree at this or any other University and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Signed: She Lin

Date: 5/11/2013



Abstract

Textual sentiment refers to the degree of positivity or negativity in texts. Compared with investor sentiment, which has been extensively studied by behavioral finance researchers, textual sentiment includes the more objective reflection of conditions within firms, institutions and markets. This thesis extracts negative sentiment from news stories by using the dictionary-based approach (based on the Loughran and McDonald's (2011) finance negative words), and studies its effects in both equity and debt markets. The main motivation behind this thesis is to attempt to open up a new avenue of research by constructing consecutive firm-specific sentiment and performing time-series analysis in addition to panel data analysis. Moreover, this thesis aims to investigate whether the well-documented role of textual sentiment in equity markets also applies to debt markets.

This thesis contains a systematic survey of the textual sentiment analysis literature in finance. It compares and discusses the information sources, content analysis methods, and financial models that have been used, innovatively summarizes their contributions from several perspectives, and suggests new paths for future research. This is the first comprehensive review of textual sentiment literature since the emergence of textual analysis in finance around the year of 2005.

Consecutive firm-specific sentiment from news stories is generated for 20 large firms traded in the U.S. market, most of whom have 10-year daily observations. Panel regressions, vector autoregression (VAR) models, and rolling-window regressions are employed to test the hypotheses regarding the interrelations between firm-specific sentiment, firm-level equity returns and trading volumes. The main findings are that at the panel data level, negative words have an immediate negative impact on firm-level equity returns, while higher returns predict lower next-day negative sentiment; higher returns predict lower next-day trading volumes. For individual firms, the most evident relations are that trading volumes cause sentiment, and that returns cause trading volumes. In addition, firm-specific textual sentiment does not always forecast firm-level returns, but has a

varying impact over time. The predictability tends to be concentrated during certain periods. There is strong evidence of the ‘indirect’ effects of sentiment on equity returns: sentiment causes trading volumes, whilst trading volumes cause returns. This research has filled a gap in the literature where nobody has hitherto examined the time-series pattern of sentiment’s role for individual stocks. Besides, the time-varying impact of textual sentiment and its ‘indirect’ effects on equity returns have not been previously studied.

This thesis investigates if textual sentiment is a determinant of sovereign bond yield spreads. No previous study has examined the role of textual sentiment on bond pricing. The study period of this research incorporates the European sovereign debt crisis, and the sample consists of the four most affected countries: Greece, Ireland, Portugal, and Spain. The results show that the increase in the percentage of negative words in the country-specific news stories within the latest three trading days leads to the increase of the difference of today’s and yesterday’s sovereign yield spreads. Meanwhile, sovereign bond yield spreads are found to be significantly directly related to negative textual sentiment. The increases of the change in yield spreads are associated with higher negative sentiment.

Overall, this thesis provides innovative ways to examine the role of firm-level sentiment on individual stocks, and the connection between textual sentiment and sovereign debt yield spreads. The primary results, which are consistent with the literature, are that negative words have a significant impact on asset prices, both equity and debt. Overall, the analysis constitutes strong evidence for the consideration of textual sentiment as a potentially important factor in asset pricing models.

Acknowledgements

First and foremost, I would like to thank my supervisor Professor Colm Kearney. Colm has been a source of encouragement, enlightening ideas and constant support. After he moved to Australia, he still acted as my primary supervisor and offered me timely and detailed comments on my research. He has given me critical advices on methodology and writing. I would also like to thank my co-supervisor, Professor Khurshid Ahmad, who has given me essential help on the content analysis part of my thesis. He recommended relevant books and articles to me, taught me technical concepts and provided me with the content analysis program.

I would like to express my sincere gratitude to my boyfriend, Jingguang Han, who has greatly helped me in the process of data collection and processing. Without the custom computer programs he created for me, it would take me much more time to finish the thesis. I would also like to thank Professor Brian Lucey, Dr. Jenny Berrill and Dr. Rachael Carroll, who helped me in different ways during the course of my research.

I am very grateful for the postgraduate research studentship I received from Trinity College Dublin, and the research funding from Enterprise Ireland grant #CC-2011-2601-B.

My deepest gratitude goes to my parents and my grandmother, who have been very supportive in all my personal and professional pursuits. Without their support, I would not have had the opportunity to pursue the Master's and Doctoral degrees outside my home country. They always stand by my side whenever I meet difficulties in life, and encourage me to move on. I truly thank them for their unconditional love.

*I dedicate this thesis to my parents
for their endless love, support, and encouragement.*

Contents

Declaration of Authorship	ii
Abstract	iii
Acknowledgements	v
List of Figures	x
List of Tables	xi
Chapter 1 Introduction	1
1.1 Introduction	1
1.2 Background, motivation and research questions	3
1.2.1 Research background	3
1.2.2 Motivation	5
1.2.3 Main research questions	6
1.3 Structure	7
1.4 Main findings and novel contributions	8
1.5 Summary and conclusions	10
Chapter 2 A survey of textual sentiment analysis literature	11
2.1 Introcution	11
2.2 Information sources	13
2.2.1 Corporate-expressed sentiment	13
2.2.2 Media-expressed sentiment	14
2.2.3 Internet-expressed sentiment	15
2.2.4 Comparing and contrasting the three sentiment sources	16
2.3 Content analysis methods and sentiment measures	18
2.3.1 Content analysis methods	18
2.2.2 Comparing and contrasting the content analysis methods	23

2.2.3 Sentiment measures	24
2.4 Econometric models and trading strategies	25
2.5 Main findings to date	33
2.6 Summary and conclusions	38
Chapter 3 Methodology	48
3.1 Introduction	48
3.2 Content analysis method	49
3.3 Models and tests	50
3.3.1 VAR model and related tests	50
3.3.2 Panel regression estimation and related tests	54
3.4 Summary and conclusions	62
Chapter 4 Results – case study of equity	63
4.1 Introduction	63
4.2 Sentiment and equity data descriptions	65
4.2.1 Generating firm-specific sentiment and descriptive statistics of sentiment data	65
4.2.2 Descriptive statistics of equity data	67
4.3 Main hypotheses	69
4.4 Empirical testing on panel data	70
4.4.1 The data used in the panel estimates	70
4.4.2 Panel estimation	71
4.5 Empirical testing on individual firms	75
4.5.1 Empirical testing over full data period	75
4.5.2 Empirical tests over rolling-window periods	77
4.6 Summary and conclusions	80
Chapter 5 Results – case study of sovereign debt	111
5.1 Introduction	111
5.2 Review of sovereign yield spread determinants literature	113
5.3 Description and analysis of the data	117
5.3.1 Data description and variable construction	117
5.3.2 Hypotheses	121
5.4 Results	122
5.4.1 Does sentiment determine yield spreads	124
5.4.2 Determinants of country-specific sentiment	128
5.5 Summary and conclusions	130

Chapter 6 Conclusions and future work _____ **143**

6.1 Introduction _____ **143**

6.2 Main findings and contributions _____ **144**

6.3 Limitation in thesis and future work _____ **146**

 6.3.1 Limitations _____ 146

 6.3.2 Future work _____ 148

6.4 Summary and conclusions _____ **149**

Appendix _____ **151**

Bibliography _____ **168**

List of Figures

2.1 Sentiment extraction (dictionary-based approach)	42
2.2 Summary of Tetlock (2007)	43
2.3 Summary of Tetlock et al. (2008)	44
2.4 Summary of Loughran and McDonald (2011)	45
4.1 Searching for qualified news stories	82
4.2 Plots of <i>FN</i> series	83
4.3 Equity returns	86
4.4 Trading volumes	88
4.5 Predictability of sentiment on return (1-year sliding windows)	90
5.1 Textual sentiment series	132
5.2 Sovereign 10+-year bond yield spreads against Germany	134

List of Tables

2.1 Qualitative information sources	46
2.2 Content analysis methods	47
2.3 Summary of Tetlock et al. (2008)	44
2.4 Summary of Loughran and McDonald (2011)	45
4.1 Average number of articles and number of words per day (firm-specific news)	95
4.2 Descriptive statistics of <i>FN</i> series	96
4.3 Summary statistics of firm-level returns and trading volumes	97
4.4 Equity data and summary statistics	98
4.5 Panel unit root tests	99
4.6 Panel estimation: do sentiment and volumes predict returns? (Sentiment measure: First lag of <i>FN</i>)	100
4.7 Panel estimation: do sentiment and volumes predict returns? (Sentiment measure: Change in <i>FN</i>)	101
4.8 Panel estimation: do returns and volumes cause sentiment?	102
4.9 Panel estimation: do sentiment and returns cause volumes?	103
4.10 VAR modeling over full sample period	104
4.11 Summary of causality tests	106
4.12 The direct and indirect effects of <i>FN</i> on returns	107
4.13 The direct and indirect effects of firm-specific sentiment on firm-level equity returns (rolling regressions)	108
4.14 Robustness analysis: the direct and indirect effects of firm-specific sentiment on firm-level equity returns (rolling regressions)	109
4.15 Robustness analysis: winsorizing returns, volumes and sentiment data	110
5.1 Number of country-specific news stories	135
5.2 Descriptive statistics of sentiment	136
5.3 Panel data unit-root tests	137
5.4 Fixed-effects regression with Driscoll-Kraay (1998) standard errors	138

5.5 Robustness analysis using pooled OLS estimation with Driscoll-Kraay (1998) standard errors	139
5.6 Robustness analysis using the fixed effects models with Driscoll-Kraay (1998) standard errors	140
5.7 Determinants of the country-specific negative sentiment	141
5.8 Determinants of the country-specific negative sentiment (robustness analysis)	142

CHAPTER 1

INTRODUCTION

1.1 Introduction

The effect of ‘sentiment’ on stock market pricing has been frequently studied by researchers during the past three decades. Broadly speaking, two types of sentiment have been studied. The first is investor sentiment, which refers to beliefs about future cash flows and investment risks that are not justified by the facts at hand (Baker and Wurgler (2007)). A substantial body of this literature focuses on finding and quantifying the overall effects of investor sentiment on individual stocks and the market, using various ways to measure investor sentiment. The second type of sentiment that has been studied over the past decade is text-based or ‘textual’ sentiment. This refers to the degree of positivity or negativity in texts. In some studies, particularly those using corporate disclosures as the information source, the term ‘tone’ (positive or negative) is used to refer to sentiment. In broader terms, however, textual sentiment may also include affects other than positivity/negativity, such as strong/weak, active/passive, and the degree of uncertainty.

The fundamental difference between investor sentiment and textual sentiment is that the former captures the subjective judgments and behavioral characteristics of investors, while the latter can include the former, but also includes the more objective reflection of

conditions within firms, institutions and markets. The connection between textual sentiment and investor sentiment is complex, and the extent to which they are causally related has not yet been thoroughly examined or understood. It is also unclear how investors interpret textual sentiment. Existing studies on textual sentiment do not make assumptions about investor rationality, or about the relationship between textual sentiment and investor behaviors.

Textual sentiment researchers extract sentiment from verbal/qualitative information, such as corporate disclosures and news articles, and investigate its role in the market. Hitherto, only quantitative fundamental information has been analyzed by researchers to test market efficiency. However, qualitative information should also be considered to be equally important as it may contain hard-to-quantify information which has incremental explanatory power on asset prices. Doing a thorough analysis of qualitative information is challenging, but quantifying sentiment in texts and tracing its effects is a natural start point. The foremost issue in studying textual sentiment is computational linguistic analysis, and this makes textual sentiment analysis in finance a multi-disciplinary research effort.

Textual sentiment analysis in finance usually centers on the following questions. Does sentiment (hereafter, ‘sentiment’ is synonymous with ‘textual sentiment’ unless otherwise stated) have an impact on future firm performance such as stock prices, returns, volatilities, trading volumes, and other firm-level and market-level variables? What are the determinants of textual sentiment? Does textual sentiment have consistent effects over time? This thesis aims to answer these questions from both similar and new perspectives compared to the existing literature. It focuses specifically on examining the time-varying effects of firm-level sentiment on stock returns and trading volumes, and on the role of textual sentiment in sovereign debt market. The main source of qualitative information is news stories. Negative words in these news stories are the sentiment category analyzed in the thesis.

The remainder of this chapter is structured as follows. Section 1.2 describes the research background, the motivation of the thesis, and the main research questions. Section 1.3 outlines the structure of the thesis on a chapter by chapter basis. Section 1.4 presents the

main findings and outlines the novel contributions. Section 1.5 provides a brief summary of this chapter.

1.2 Background, motivation and research questions

This section first describes the research background of this thesis, including classical financial theory, the development of behavioral finance, and the emergence of textual sentiment analysis. It then explains the motivation of the thesis and how it contributes to the literature. Finally, it outlines the main research questions of the thesis.

1.2.1 Research background

The dominant theory of financial markets for the past half century has been the efficient markets hypothesis (EMH) of Fama (1970), which assumes that prices incorporate all available information about the underlying asset. The EMH relies on a number of important assumptions. *First*, investors are rational, they value each security for its fundamental value (firm's discounted risk-adjusted cash flows conditional on investors' information sets), and they make choices consistent with the principles of subjective expected utility (SEU)¹. *Second*, even if some investors are not fully rational, arbitrageurs will bring prices to their fundamental values. *Third*, the costs of getting prices to reflect information are always zero (Grossman and Stiglitz (1980)). The EMH has, however, been heavily challenged during the 1980s and early 1990s. Evidence shows that the EMH is difficult to sustain in the real world. Empirically, excess equity volatility (Shiller (1981)), the calendar effect (e.g. Keim (1983)), differences in returns of past winners and losers (De Bondt and Thaler (1985)), the book-to-market effect (Fama and French (1992)), the momentum effect (Jegadeesh and Titman (1993)), and the stock market crash of October 1987 are all evidence against the EMH.

The stylized findings of anomalies in the EMH indicate a research program that is under strain. An alternative approach to asset pricing is the behavioral finance (BF) perspective

¹ Subjective expected utility is a method in decision theory in the presence of risk, promoted by L. J. Savage in 1954.

which is derived from cognitive psychology and decision theory. BF is closely related to behavioral economics, and is underpinned by the decision science research of Kahneman and Tversky (1979). BF seeks to understand investors' decisions and their effects on financial markets with recourse to psychological studies of the cognitive and emotional biases of market participants. BF does not make the assumption that all economic agents are rational, but rather that they are prey to individual and/or collective biases which are psychological in nature (e.g. heuristic-driven bias). BF also recognizes frame dependence (i.e. decision-making depending on the form used) when a person has difficulty seeing through an opaque frame. These create anomalies in market prices and returns that cannot be explained by the EMH.

BF has itself also been critiqued. There are two main streams of criticisms. The first is the *joint hypothesis problem* (Fama (1970)). This problem can be summarized as stating that market efficiency is not testable without some equilibrium model, an asset pricing model: 'we can only test whether information is properly reflected in prices in the context of a pricing model that defines the meaning of 'properly'' (Fama (1991): 1576). Therefore, the findings of anomalies in equity pricing may be due to either market inefficiency or to an inappropriate pricing model. Following this idea, Fama and French (1993) incorporated size-based and book-to-market factors into the basic Capital Asset Pricing Model (CAPM), which uses a single factor, beta, to compare the excess returns on a security with the excess returns on the market as a whole. The Carhart (1997) four-factor model, the extension of the Fama-French three-factor model, contains an additional momentum factor. Another criticism of BF is concerned with psychological robustness. Hirshleifer (2001) notes a common objection to psychological models of asset pricing: 'it is too easy to go theory fishing for psychological biases to match data ex post' and 'confused investors will learn their way to good decisions' (p.1,535).

We have already observed that in the past decade, researchers have begun to study textual sentiment. One hypothesis about textual sentiment is that it contains new information about fundamentals that will be reflected in prices, either temporarily or permanently. This hypothesis is raised within the EMH framework. An alternative hypothesis is that

positivity or negativity contained in texts is associated with investor sentiment and behaviors; textual sentiment can reflect or cause investors' biases, which help explain the 'anomalies' in the market. This hypothesis is built within the BF framework. The most likely scenario is that elements of truth are contained in both hypotheses.

1.2.2 Motivation

The motivation of this thesis is to extend the existing literature that analyzes the effect of sentiment on asset price returns and trading volumes.

Firstly, the effect of daily 'firm-specific' sentiment on individual stocks has not been extensively studied. Both Tetlock (2007) and Garcia (2012) used daily sentiment, but the sentiment is extracted from general economics and finance commentaries, and is considered to be 'market-level' sentiment, and the authors examined its effect on market returns and trading volumes. It is usually difficult to get consecutive firm-level sentiment at daily frequency because most firms do not have news stories every day. Previous studies that have studied firm-level sentiment have used different methods to deal with this issue. For example, Tetlock et al. (2008) employed panel data models and were able to examine sentiment's effects on the next day's returns. Other studies, such as Davis et al. (2008), Engelberg (2008), Henry (2008), Feldman et al. (2008), Demers and Vega (2010) and Doran et al. (2010), generated firm-level sentiment from quarterly or annual corporate disclosures. Because of the low-frequency and characteristics of this type of sentiment, these researchers employed standard event-study methods to investigate the relationship between sentiment on the announcement day and the asset returns in the event windows.

The main motivation behind this thesis is to attempt to open up a new avenue of research by constructing consecutive firm-level sentiment and performing time-series analysis in addition to panel data analysis. Moreover, textual sentiment analysis has rarely been carried out on markets other than equity. This thesis also aims to extend the existing literature into debt markets in order to find out if the well-documented role of sentiment in equity markets also applies to debt markets.

1.2.3 Main research questions

Three main research questions are addressed in this thesis. *First*, does firm-specific sentiment cause individual stock returns and trading volumes, and *vice versa*? As is known from the literature, there exists a mutual impact between market-level textual sentiment (especially negative words) and market returns. The relationship between sentiment and market trading volumes is also significant. When it comes to firm-specific sentiment, firm-level equity returns and trading volumes, it is possible that the interrelations also exist, but this has not yet been researched. In addition, firm-specific textual sentiment could act as a previously unconsidered source of information that drives firm-level returns by first working through trading volumes.

Second, does the effect of firm-level sentiment on individual stocks evolve over time? Garcia (2012) finds that controlling for other well-known time-series patterns, the predictability of stock returns using news content (general economics and finance) is concentrated in recessions. It is reasonable to assume that the effect of firm-level sentiment on individual stocks is time-varying, although it is not necessarily concentrated in recessions.

Finally, does textual sentiment play a role in debt markets? It has not yet been investigated whether sentiment has any effect on bond yield spreads or interest rates. This thesis examines if sentiment is a determinant of sovereign bond yield spreads in the context of European sovereign debt crisis. Traditionally, credit risk, liquidity risk and general risk aversion have been identified as the three determinants of sovereign yield spreads. This thesis aims to discover if textual sentiment has any additional explanatory power on sovereign yield spreads in addition to these three factors.

1.3 Structure

Chapter 2 surveys the fast-growing ‘textual sentiment analysis’ literature. It describes and compares the qualitative information sources used in the literature to date. The three most popular information sources, corporate disclosures, news articles and internet messages, have some unique advantages and disadvantages relative to one another. The most frequently used content analysis methods, the dictionary-based approach and machine learning, are also reviewed in detail. The econometric models that are employed in the literature are presented, and the general forms of these models are described. The chapter summarizes the main findings and contributions of the literature from varied perspectives, and suggests new ideas for future research.

Chapter 3 presents the methodology of the thesis. It first introduces the content analysis method and the dictionary used in the thesis, and briefly describes the text corpora for separate studies of equity and debt. The second part of this chapter presents the major econometric models, methods, and tests used in the thesis, including the rationale for using them and under what assumptions they work well.

Chapter 4 examines the effects of firm-specific textual sentiment (extracted from company-specific news stories) on firm-level stock returns and volumes, and *vice versa*. The sample consists of 20 large corporations traded in the U.S. market. Hypotheses regarding the interrelations among sentiment, returns and volumes are formulated, and they are tested by employing panel regressions, vector autoregression (VAR) models, and rolling-window regressions in sequence. In this way, the role of textual sentiment is examined over the full data period and on sliding periods, from the perspective of both individual companies and panel data. The main findings are that at the panel data level, negative words have an immediate negative impact on firm-level equity returns, while higher returns predict lower next-day negative sentiment; higher returns predict lower next-day trading volumes. For individual firms, the most evident relations are that trading volumes cause sentiment, and that returns cause trading volumes. In addition, firm-specific

textual sentiment does not always forecast firm-level returns, but has a varying impact over time.

Chapter 5 examines whether country-specific textual sentiment is a determinant of sovereign bond yield spreads and whether sovereign bond yield spreads are directly related to negative textual sentiment. It first presents a review of the literature that discusses the determinants of sovereign yield spreads in the euro zone. The study period of this chapter incorporates the European sovereign debt crisis, and the sample consists of the four most affected countries: Greece, Ireland, Portugal, and Spain. Textual sentiment has never been considered as a potential determinant of sovereign yield spreads, probably because textual sentiment analysis is relatively new and not widely known. By employing panel regressions, this chapter shows that the well-documented role of textual sentiment in equity markets also applies to bond markets. Negative words have a positive non-negligible impact on sovereign yield spreads, after controlling for the traditionally recognized variables.

Chapter 6 summarizes the main findings and implications of the thesis and draws together the important conclusions. It also describes the limitations of the thesis and discusses future research.

1.4 Main findings and novel contributions

Chapter 2 of this thesis provides a systematic survey of the textual sentiment analysis literature. It compares and discusses the information sources, content analysis methods, and financial models that have been used, and innovatively summarizes their contributions from several perspectives. To my knowledge, there has been no such comprehensive review of textual sentiment literature since the emergence of textual analysis in finance around the year of 2005. This chapter has been prepared as a research discussion review paper for submission to an international academic journal.

Few studies have performed time-series analysis on firm-specific textual sentiment, and no one has examined the time-varying pattern of its role in stock markets. This is because it is difficult to get consecutive sentiment for individual companies. In this thesis, consecutive sentiment from news stories was generated for 20 large firms, most of whom have 10-year daily observations. This analysis has filled a gap in the literature where nobody has hitherto examined the time-series pattern of sentiment's role for individual stocks. In addition, this thesis creatively examines the time-varying 'direct' and 'indirect' effects of sentiment on stock returns. The primary conclusion to emerge is that firm-specific sentiment directly causes returns at least some of the time, and the predictability tends to be concentrated during certain periods. This suggests promising future research on whether these periods are associated with important firm-specific news and events. The indirect effects of sentiment on equity returns (sentiment causes trading volumes, whilst trading volumes cause returns) are most evident in the panel data. This chapter has been prepared as a research paper that has been accepted by the 2013 Annual Meetings of the European Financial Management Association. It is also being prepared for submission to a high-impact academic journal.

No previous study has examined the role of textual sentiment on bond pricing. This thesis investigated if textual sentiment is a determinant of sovereign bond yield spreads. The novel contribution is that it discovers textual sentiment has additional explanatory power on sovereign yield spread over credit risk, liquidity and international risk aversion, and the well-documented role of textual sentiment in stock market also applies to bond market. Meanwhile, sovereign bond yield spreads are found to be significantly directly related to negative textual sentiment. This analysis suggests another perspective of studying the determinants of bond yield spreads or interest rates. It forms the thesis paper that has been accepted by the 20th Annual Conference of the Multinational Finance Society, and will be submitted for journal publication.

1.5 Summary and conclusions

This chapter has introduced the relevant research background, including the EMH and BF frameworks, and recent trends in textual sentiment analysis. The description of the research background provides a clear picture of where textual sentiment analysis is positioned and how it is related with the relevant topics. Subsequently, this chapter presented the motivation for the thesis and the main research questions. The structure of the thesis was introduced on a chapter by chapter basis. The main findings and novel contributions were highlighted.

Overall, this thesis aims to construct novel time series using new textual sentiment analysis on firm-level equity returns and on country-level debt markets. The thesis contains novel analysis that contributes to our understanding of how information and sentiment obtained from news stories interacts with financial markets, and *vice versa*. The main ambition of the thesis is to show how such analysis can open up awareness for related research using textual sentiment in many other applications.

CHAPTER 2

A SURVEY OF TEXTUAL SENTIMENT ANALYSIS LITERATURE

2.1 Introduction

Many thousands of empirical studies in accounting and finance have examined the roles of quantitative financial information in the process of asset price discovery. These studies have shown that quantitative information about market fundamentals cannot fully explain asset price movements (e.g. Shiller (1981), Mitchell and Mulherin (1994)). According to Fama (1991), this does not necessarily suggest market inefficiency; it may be due to the lack of an appropriate asset pricing model. There is growing realization that to develop a ‘good’ asset pricing model, one has to incorporate not only the conventional quantitative measures of firm performance, but also qualitative information (Feldman et al. (2008)).

The inclusion of qualitative information into equity asset pricing models provides another perspective and potentially complementary information to quantitative informational measures in the price formation process. Qualitative information in publicly available documents or media articles may contain additional hard-to-quantify information. Li (2006) suggests that text-based information can potentially provide a more independent test of

market efficiency than the number-based measures, because many of the latter are highly correlated so different anomalies may reflect the same empirical regularity.

An increasing body of textual sentiment analysis since 2005 extracts sentiment from corporate disclosures, news stories or internet postings. Corporate disclosure studies usually aim to discover the fundamental relation between sentiment and future firm performance or other quantitative variables. News stories and internet posting studies focus on the short-term effects of sentiment on market variables such as stock prices, returns, trading volumes and volatility. Every information source and linguistic method has its unique advantages and disadvantages which influence the research focus and limitations of each study. Because of the use of different types of sentiment and varied research focuses, they employ a variety of models and methods to test hypotheses and to arrive at results.

This Chapter is structured as follows. Section 2.2 describes and discusses the qualitative information sources used in the literature. Each of the most popular information sources – public corporate disclosures, news articles and internet messages – are described along with their unique features that are advantageous to others. Section 2.3 introduces and compares the most frequently used textual analysis methods: the dictionary-based approach and machine learning, and introduces the sentiment measures derived from the different linguistic analysis methods. Section 2.4 presents the models employed in the existing literature, including contemporaneous linear regressions and univariate and multivariate time-series models. The general forms of models are summarized. This section also describes the ex-post sentiment-based trading strategies that are discussed in the literature. Section 2.5 summarizes the main findings and contributions of the literature to date. Section 2.6 summarizes the chapter, draws together the most important conclusions and suggests future research directions.

2.2 Information sources

The qualitative information that has been analyzed by textual sentiment researchers in finance comes predominantly from three sources: public corporate disclosures, news articles and internet messages. The sentiment expressed in these texts conveys market participants' and commentators' information and opinions about many aspects of developments in firms, institutions and markets. It also reflects how sentiment responds to these events. The most important work on sentiment analysis within corporate disclosures includes Davis et al. (2008), Feldman et al. (2008), Henry (2008), Henry and Leone (2009), Demers and Vega (2010), Doran et al. (2010), Li (2010), Loughran and McDonald (2011), Jegadeesh and Wu (2012), and Price et al. (2012). The literature that extracts sentiment from news articles or analyst's reports includes Tetlock (2007), Tetlock et al. (2008), Sinha (2010), Garcia (2012), Huang et al. (2012), and Rees and Twedt (2012). A small group of studies examines internet-posting sentiment, foremost amongst them being Antweiler and Frank (2004), Das and Chen (2007) and Chen et al. (2012). A summary of the information sources used in each of these papers is presented in Table 2.1. We now discuss each information source in turn.

2.2.1 Corporation-expressed sentiment

Public corporate disclosures are a natural source of textual sentiment for researchers insofar as they are official releases that come from insiders who have better knowledge of the firm than outsiders. The linguistic style and tone in these texts might convey useful information about expected future firm performance over and above the numbers in financial statements. This corporation-expressed sentiment is particularly useful in examining the role of qualitative information about individual firm performance and stock pricing. Offsetting this, however, is the low frequency of the data, because firms usually make these disclosures on a quarterly or annual basis.

The corporation-expressed sentiment literature mainly studies corporate periodic reports, 10-Ks, 10-Qs, or earnings press releases and earnings conference calls. The former

category includes Li (2006), Feldman et al. (2008), Li (2010), Loughran and McDonald (2011) and Jegadeesh and Wu (2012). In particular, Feldman et al. (2008) and Li (2010) focus only on the MD&A sections of 10-Ks and 10-Qs. Loughran and McDonald (2011) also use MD&A texts in only one part of their analysis. The earnings disclosures category includes Davis et al. (2008), Henry (2008), Henry and Leone (2009), Demers and Vega(2010), Doran et al. (2010), and Price et al. (2012). In these studies, textual sentiment is considered as a new factor in addition to the usual firm-level fundamentals that are analyzed in typical event studies. Feldman et al. (2008) show how MD&As might be a superior information source relative to earnings announcements because the preliminary earnings announcements were typically not filed with the SEC prior to 2003, hence not routinely scrutinized as periodic reports with the SEC. Moreover, MD&As contain more qualitative information from management than earnings announcements, because the purpose of MD&A is to provide a management perspective within the narrative on their firms' past performance, their current financial positions, and their future prospects.

2.2.2 Media-expressed sentiment

Media-expressed sentiment is the positivity or negativity contained in news stories, in-depth commentaries or analyst reports. These texts are relevant to general economic conditions, to overall financial market conditions and prospects, and to individual industries and firms. Obviously, sentiment in media articles about general economic conditions or financial markets is an appropriate choice for studying market-wide price patterns and activities, while industry-specific or firm-specific news stories and analyst reports are more appropriate for analyzing firm-level stock prices, returns, trading volumes and other performance indicators and attributes.

In the media-expressed sentiment literature, Tetlock (2007) and Garcia (2012) study general economics and finance news stories or commentaries in two major U.S. newspapers, the *Wall Street Journal* and the *New York Times*, respectively. Engelberg (2008), Tetlock et al. (2008), and Sinha (2010) employ much wider news sources, and they all research firm-specific news. Engelberg (2008) collects news articles addressing

earnings announcements of 4,700 unique firms; Tetlock et al. (2008) use the *Dow Jones News Service* in addition to the *Wall Street Journal*; and Sinha (2010) analyzes firm-level sentiment in 587,719 news articles provided by the *Thomson Reuters NewsScope* service. Using wider news sources rather than a small number of particular newspapers is generally more appropriate and preferred, because it decreases the probability of getting biased sentiment from some sources that do not represent the overall market. Huang et al. (2012) and Rees and Twedt (2012) are the only researchers to date that have examined sentiment in analyst reports relating to specific firms. Rees and Twedt (2012) argue that financial analysts perform a pivotal role in evaluating accounting data and disseminating their analysis to the public, and that this ensures their continued interest from investors and academic researchers.

2.2.3 Internet-expressed sentiment

Internet postings are a potentially useful source of textual sentiment because many people spend a considerable amount of time every day reading and writing internet postings about stocks. The message flows comprise potentially valuable insights, market sentiment, manipulative behavior, and reactions to other sources of news (Das and Chen (2007)), possibly causing a significant impact on financial markets. Internet-expressed sentiment is potentially ‘noisier’ than corporation-expressed or media-expressed sentiment, because it contains more views from individual traders. This can make it a potentially powerful source from which to extract small investor sentiment. Even if a high proportion of these messages contain noise or irrational sentiment, Black (1986) argues that the narratives around a large number of small events could generate more potent causal sentiment factors than the narratives around a smaller number of large events if the former have greater proportions of noise within them.

Antweiler and Frank (2004), Das and Chen (2007) and Chen et al. (2012) have conducted computational textual sentiment analysis on internet messages. Antweiler and Frank (2004) analyzes 1.5 million messages posted on *Yahoo!Finance*¹ and *Raging Bull*² about the 45

¹ <http://finance.yahoo.com/>

companies in the *Dow Jones Industrial Average* index and the *Dow Jones Internet Index*. Das and Chen (2007)'s sample consists of all messages posted on the *Yahoo!* message boards from July to August 2001 that relate to the 24 technology-sector stocks presented in the *Morgan Stanley High-Tech Index* (MSH), resulting in a total of 145,110 messages. Chen et al. (2012) analyze the single-ticker articles published between 2006 and 2010 on the *Seeking Alpha*³ website, covering 3,030 distinct United States common stocks.

2.2.4 Comparing and contrasting the three sentiment sources

The advantages and disadvantages of each information source are not difficult to identify. *First*, to the extent that corporate disclosures convey sentiment from management – the insiders who know most about their firms – they are a potentially valuable textual source, but it is unlikely that management reveals ‘the truth, the whole truth, and nothing but the truth’, and it may be tempted to seek to manipulate investors’ judgments. Because these narratives are firm-specific, they are particularly relevant to studying the role of qualitative information in individual firm performance and stock pricing. In contrast, textual sentiment from news stories can be used in the context of either market-level or firm-level analysis depending on the subject matter of the news stories. So although news stories are a more flexible information source than corporate disclosures, sentiment from outsiders is less likely to capture insiders’ views and perspectives. The qualitative information in analyst reports are likely to sit somewhere between corporate disclosures and news stories. Analysts may have some insider information that is communicable to investors and other market participants and commentators through sentiment.

Second, news stories cover many different kinds of events. From the event-study perspective, researchers could examine the impact of a whole range of events on firm earnings and stock returns. In this case news, stories are superior to corporate disclosures insofar as the latter contain limited event types such as earnings conference calls and annual reports. Tetlock et al. (2008) argue that ‘examining all newsworthy events simultaneously limits the scope for dredging for anomalies’ (Fama (1998): 1438).

² <http://www.ragingbull.com/>

³ <http://seekingalpha.com/>

Third, news stories predominantly reflect hindsight rather than foresight. They are usually written about what has happened or what is scheduled to happen rather than what might happen in the future. In contrast, corporate disclosures such as the MD&A sections of 10-Ks, the transcripts of conference earnings calls, and analyst reports tend to contain more forward-looking statements. Sentiment in these narratives has more potential predictability on future outcomes such as firm performance and stock prices or returns.

Fourth, because corporate releases are usually on a quarterly or annual basis, the sentiment series extracted from them is not ideal for time-series modeling because of its low frequency. It is, however, appropriate for cross-sectional analysis and event studies that examine the effect of sentiment on stock returns around corporate releases. Once again, media articles are a more flexible information source for extracting textual sentiment that can be used in modeling monthly, weekly, daily and even intra-day levels, as long as the relevant texts are available at the required frequencies.

Fifth, because the online media tends to be open and unregulated, internet-expressed sentiment is likely to be noisier than corporation-expressed or media-expressed sentiment. It is also likely to contain little new information that is incremental to published public news. A large proportion of internet messages are written and posted by noise traders or uninformed investors who might be susceptible to particular opinions and sentiments, and their 'information' is likely to be less accurate or reliable. In other words, their comments and advice may actually have a low ratio of information to noise. It follows that internet postings are not an ideal source of information for testing market efficiency. Insofar as they better represent small investor sentiment, they might provide more evidence for the behavioral finance framework.

Sixth, the pre-processing of internet postings prior to content analysis is more costly than corporate disclosures and media articles, because people tend to write less accurately, clearly and formally on internet boards, and the meaning of the texts can be ambiguous. Corporate documents and media articles are written more professionally and require less

pre-processing time. This is probably the reason why internet postings have been the least frequently studied source of textual sentiment analysis.

Since each of the information sources has its own advantages and disadvantages, the best practice may be to employ as many information sources as possible. Kothari et al. (2009) investigate an extensive set of sources from the print media, including a combination of corporate disclosures, analyst reports and news stories in financial press. Their sample comprises all available disclosure texts for 887 firms, amounting to 326,357 texts in total.

2.3 Content analysis methods and sentiment measures

This section introduces and discusses the content analysis methods used in the literature, and the sentiment measures that have been adopted.

2.3.1 Content analysis methods

Almost half a century ago, Stone *et al.* (1966) realized that words and sentences are quintessential human artifacts, the products of social constructs and experience, and that they provide an important source of evidence about individual and collective processes such as economic and financial activities. He defined content analysis as ‘any research technique for making inferences by systematically and objectively identifying specified characteristics within text.’ Sentiment is articulated in human discourse, public speeches, news reports, blogs and other forms of written, spoken and visual communication. By analyzing the raw data of words and sentences, behavioral scientists can uncover the evidence of sentiment or affect within text. The most common content analysis methods in the textual sentiment literature are the *dictionary-based approach* and *machine learning*. Table 2.2 summarizes the methodology employed in each study in the textual sentiment literature in finance.

The dictionary-based approach

The dictionary-based approach uses a mapping algorithm in which a computer program reads text and classifies the words, phrases or sentences into groups based on pre-defined dictionary categories (Li (2010)). This is often referred to as the ‘bag-of-words’ model in natural language processing. Documents are considered to be the bag of words, and the structure along with any linear ordering of words within the context is ignored (Manning and Schutze 1999). There are two important issues in the dictionary-based approach. The first is the word lists, i.e. the words contained in each sentiment category; and the second is how each word in the word lists should be weighted. We discuss each of these in turn.

Word Lists/Dictionaries

The most popular word lists are the General Inquirer (GI) built-in dictionaries developed and used by Philip Stone, a specialist in social psychology (Stone et al. (1966)). Most of its word lists come from the *Harvard IV-4* dictionaries. Appendix 1 and 2 contain excerpts from the introductory session on GI’s official website⁴ and examples of GI word lists. A number of financial researchers have used the GI program and word lists (or the GI word lists only) to derive sentiment in texts, including Tetlock (2007), Engelberg (2008), Feldman et al. (2008), Tetlock et al. (2008), Kothari et al. (2009), Price et al. (2012), and Rees and Twedt (2012).

Another frequently used dictionary is built from the popular textual analysis program, *DICTION* (Hart (2000)), which was developed by Roderick Hart, a specialist in politics and mass media. *DICTION 5.0* counts words based on 33 separate dictionaries and two sets of variables, and produces outputs of raw frequencies, percentages, and standardized scores for these lists of words. Included with the dictionary scores are scores for five master variables - *activity*, *certainty*, *commonality*, *optimism* and *realism* - and other calculated variables. A description of the *DICTION 5.0* software from its official website⁵ is attached in Appendix 3. Davis et al. (2008) employ *DICTION* to gauge managers’ tone in earnings press releases by counting words characterized by linguistic theory as optimistic or pessimistic. Henry (2008) also uses *DICTION*.

⁴ <http://www.wjh.harvard.edu/~inquirer/3JMoreInfo.html>

⁵ <http://rhetorica.net/diction.htm>

It is important to note that both the *GI* and *DICTION* are general English language linguistic dictionaries rather than dictionaries that are specific to the domain of financial disclosure. Li (2010) finds that the classification of tone based on the *GI* does not provide sufficient accuracy. This is not surprising, because general word lists omit words that are considered positive or negative in the context of financial disclosure, and include other words that would not (Henry and Leone (2009)). For example, words such as ‘tax’ and ‘liability’ are on the negative wordlist of the *GI*, but are not negative in the financial context. Loughran and McDonald (2011) find that almost three-fourths (73.8 percent) of the negative word counts in the *GI* list are attributable to words that are typically not negative in a financial context. To overcome this problem, dictionaries / word lists specific to the finance domain have been built by researchers so that more accurate and efficient sentiment scores can be generated. Studies that focus on comparing various dictionaries include Henry and Leone (2009) and Loughran and McDonald (2011). The former uses three word lists – *DICTION*, *GI* and Henry (2008) - to gauge the tone of earnings press releases. The authors find that the context-specific word list developed by Henry (2008) is more powerful than the general word lists used in previous research. Later research by Price et al. (2012) also concludes that the Henry (2008) dictionary is more powerful in detecting cumulative abnormal returns beyond the initial reaction window. Loughran and McDonald (2011) use *GI* negative words which have been expanded by inflecting each word to forms that retain the original meaning of the root word, and the finance-specific word lists developed by themselves (L&M lists) to assess sentiment in 10-Ks⁶. They argue that although the apparent power of the two negative word lists is similar, the use of the finance-specific list to avoid those words in the *GI* list that might proxy for industry or other unintended effects. The L&M lists have become increasingly popular in more recent studies by Doran et al. (2010), Jegadeesh and Wu (2012), Garcia (2012), and Chen et al. (2012).

Term Weighting

⁶ The L&M finance dictionary consists of all words that occur in 5% or more of the 10-Ks from 1994 to 2008 (Loughran and McDonald (2011)).

Most studies employ proportional weighting, which considers every word in the list to be equally important. Typical examples are the studies that use the *GI*, which calculates simple frequencies for words appearing in the text that fall within each category. Term weighting, however, acknowledges that raw word counts are not the best measure of a word's information content (Loughran and McDonald (2011)). Loughran and McDonald (2011) use two weighting schemes, a simple proportional weighting and one that weights each word inversely proportional to its document frequency, or the frequency with which each it appears in the sample of documents. Jegadeesh and Wu (2012) argue that in the context of inferring tone, there is no reason to expect why words that are found in fewer documents should be more powerful than those found in more documents. These researchers use a weighting scheme that is particularly suitable for finance applications. They assign weights for each word based on how the market has reacted to them in the past. Their results indicate that the appropriate choice of term weighting is more important than a complete and accurate dictionary to which the weighting scheme is applied.

Figure 2.1 shows a generic process of using the dictionary-based approach to extract sentiment from texts. The first step is to collect qualified texts to form the corpus. Pre-processing of the collected files is usually necessary in order to split the articles and group them by date. Subsequently, the sentiment categories (i.e. the word lists) in the content analysis program should be selected. Researchers can also write their own programs which allow the addition of customized word lists. Sentiment scores are obtained by selecting the text files and running the program. It is also possible to construct other sentiment measures based on the original sentiment scores. Finally, the sentiment measures and other variables can be used for financial modeling and hypothesis testing.

Machine learning

Pioneered by mathematicians and computer scientists, machine learning relies on statistical techniques to infer the content of documents and to classify them based on statistical inference (Li (2010)). The steps involved in this method are as follows. A proportion of the complete corpus of text to be analyzed is designated as the 'training set'. Each word in the

training set is manually classified as ‘positive’, ‘negative’, or some other dimension of sentiment (such as ‘strong’, ‘weak’, ‘active’, ‘passive’ etc). A selection of sentiment analysis algorithms (e.g. the Naïve Bayesian algorithm) is then trained on the training corpus. The algorithms ‘learn’ the sentiment classification rules (or ‘grammar’) from the pre-classified data set, and apply these rules out-of-sample to the whole corpus. When all words in the complete corpus are classified, sentiment measures can be derived using various combinations of the classifications used in the training corpus. In essence, machine learning involves one or more algorithms reading a training set and writing a ‘model’ containing its statistics, which are then applied to the whole corpus to derive textual sentiment scores.

Although machine learning can be carried out by customized programs, two established programs are worth mentioning. The first is the *Rainbow* package developed by McCallum (1996). *Rainbow* supports several alternative classification methods. The default is *Naïve Bayes*, but k-nearest neighbor, *TFIDF*, and *probabilistic indexing* are all available. Another text-processing tool is the *Reuters NewsScope Sentiment Engine*⁷. This engine has three major processes: pre-processing; the lexical and sentiment pattern identifier; and the sentiment classifier. In the first two processes, documents are split into sentences which are further split into words; the subject of the sentence is identified; and each word is identified as a noun, verb, adjective, adverb, and intensifier. The last process, sentiment classification, is done by a three layer back-propagation neural network. It follows the general steps of machine learning, and produces as output the probability of the text being positive, negative, or neutral.

Antweiler and Frank (2004) use the default method *Naïve Bayesian* algorithm within the *Rainbow* package to classify messages into one of three categories: *buy*, *sell* and *hold*. The *Naïve Bayesian* algorithm is also employed by Li (2010) to classify sentences into one of four tones: *positive*, *neutral*, *negative*, and *uncertain*. Das and Chen (2007) use five algorithms to classify internet messages into *bullish*, *bearish* or *neutral*. Sinha (2010) uses the *Reuters NewsScope Sentiment Engine* to calculate the probabilities of news articles

⁷ <http://www2.reuters.com/productinfo/newsscope/sentiment/>

being positive, negative and neutral, respectively. Huang et al. (2012) use the Naïve Bayes machine learning approach to extract textual opinions from analyst reports about S&P 500 firms.

2.3.2 Comparing and contrasting the content analysis methods

The dictionary-based and machine learning approaches have their advantages and disadvantages. *First*, the dictionary-based approach is probably the easiest for business, economic and financial analysts to handle because the well-established programs like *GI* and *DICTION* are readily available and have been most frequently used in the literature. As discussed previously, however, the general dictionaries are not appropriate for textual analysis in financial contexts. This problem can be largely mitigated by applying finance-specific word lists such as the *L&M* lists, and the key issue with this methodology lies in the choice of the most appropriate weighting scheme. The latter most likely depends on the nature of the corpus being analyzed and the hypotheses being tested.

Second, implementing machine learning is more time-consuming and costly than the dictionary-based approach, because the text in the ‘training set’ must be manually classified. In addition, to ensure the highest quality of classification, the criteria for selecting appropriate people to read the texts are strict (e.g. native speakers with economic and finance backgrounds). The accuracy rate of machine learning is usually higher than the dictionary-based approach (Li (2010)). For example, the *Reuters NewsScope Sentiment Engine* boasts a 75 percent accuracy rate compared to the average assessment of human analysts (Sinha (2010)). Huang et al. also document that their classification accuracy achieved using the Naïve Bayes machine learning approach is 80.9% in the in-sample validation and 76.9% in the out-of-sample validation, which is substantially higher than that achieved using the dictionary-based approaches based on the general dictionaries (e.g. 48.4% for *GI* and 54.9% for *DICTION*).

2.3.3 Sentiment measures

Having extracted sentiment from a corpus of text, the construction of sentiment measures is relatively straightforward. Due to differences in the methods of content analysis, however, the characteristics of the resulting sentiment measure are determined by the content analysis method used. We now summarily describe and compare the most common types of sentiment measures.

Measures in the dictionary-based approach

In studies using the dictionary-based approach with the *GI*, the most common measure is the percentage of the number of words in a given sentiment category to the total number of words in the text (e.g. Kothari et al. (2009)), or the standardized percentage (Z-score) (e.g. Tetlock et al. (2008)). Standardization is necessary if the raw frequency of matched words in the total number of words is not stationary, which can happen when regime changes occur over time in the distribution of words in the text (Tetlock et al. (2008)). This can happen, for example when the writing style changes with the author. The raw percentage of words is always greater or equal to 0, but the standardized percentage can be either greater or less than 0. Positive/negative Z-scores indicate that the degree of optimism/pessimism or other affects in the text is above/below average. Another relative measure in contrast to the absolute measure (i.e. raw percentage) is the count of positive words minus the count of negative words, divided by the sum of positive and negative word counts. With the relative measure, one can more easily identify whether a piece of text is relatively positive (if greater than 0) or relatively negative (if less than 0), and the magnitude of sentiment relative to the length of the text. In studies using *DICTION*, sentiment measures are also raw frequencies and standardized scores. Because there is no direct sentiment category of ‘positive’ or ‘negative’, Davis et al. (2008) calculate the optimistic measure as the frequencies of words of praise, satisfaction and inspiration, and the pessimistic measure as the frequencies of words of blame, hardship and denial.

Tetlock (2007) constructs his sentiment measure in another way, employing principle component analysis (PCA) to extract the most important semantic component from the (77×77) variance-covariance matrix of the categories in the *GI* dictionary. By doing this, he extracts an underlying media factor as the linear combination of all *GI* categories which accounts for the maximum variation in the total variance of all the categories. In fact, the first factor is approximately equal to a linear combination of only 4 *GI* categories: *Negative*, *weak*, *fail* and *fall*. The advantage of PCA is that it does not decide *ex-ante* what type of sentiment to get from the text, but objectively extracts the measures of sentiment that account for the greatest variation in the total variance of all categories. This principle factor thus represents the overall style of the text. Doran et al. (2010) also use PCA to define sentiment measures. However, if one wants to focus on a certain type of sentiment, PCA would not be appropriate.

Measures in machine learning

Studies that perform machine learning construct the sentiment measure based on the classification of the texts. For example, Li (2010) defines the affect of a sentence to be 1 if the learning algorithm predicts the sentence to be positive, 0 if neutral, and -1 if negative or uncertain. For each document, the author defines the overall affect as the average score of individual sentences. Das and Chen (2007) construct their daily sentiment series in a similar way. Sinha (2010) defines the sentiment score as the difference between the probability of the news article being positive and negative.

2.4 Econometric models and trading strategies

Having decided upon the information source from which to construct the measure of sentiment (most commonly from corporate disclosures, news articles or internet messages), and having decided upon the content analysis method (usually some variant of the dictionary-based or machine learning approaches), the researcher constructs one or more measures of sentiment. This is a particularly interesting part of the research program – and

it is about to get even more enthralling. Armed with a relatively new kind of data, the next step is to investigate any number of hypotheses about how the sentiment interacts with the financial variables and *vice versa*. The range of modeling approaches is very wide, and in this section we present a summary review of what has been investigated to date. We shall see that this research agenda is very new, and that a great deal of innovative approaches, models and hypotheses remain to be examined.

Linear regression models

The most common approach has been to employ the linear regression model on time series data comprising general market-level sentiment and stock index performance. Where cross-sectional sentiment and firm-level performance variables are available, some researchers have used panel regression methods to control for the individual heterogeneity of firms (e.g. Henry (2008), Demers and Vega (2010)). We can describe and summarize the general approach that has been adopted in single-equation studies within the sentiment-finance literature in Equation 2.1 below.

$$y_t = \alpha_0 + \sum_{j=1}^J \alpha_{1,j} y_{t-j} + \sum_{j=1}^J \sum_{f=1}^F \alpha_{2,j}^f X_{t-j}^f + \sum_{j=1}^J \sum_{a=1}^A \alpha_{3,j}^a S_{t-j}^a + \varepsilon_t \quad (2.1)$$

In Equation 2.1, y is the dependent variable, X is a vector of $i = 1 \dots N$ control variables, and S is a vector of various measures of affect or sentiment. Equation 2.1 states that firm-level or market-level performance in period t depends on a constant term, α_0 ; on $j = 1 \dots J$ lagged dependent terms, $\sum_{j=1}^J \alpha_{1,j} y_{t-j}$; on $j = 1 \dots J$ lagged terms of $f = 1 \dots F$ control variables in the X vector, $\sum_{j=1}^J \sum_{f=1}^F \alpha_{2,j}^f X_{t-j}^f$; on $j = 1 \dots J$ lagged terms of $a = 1 \dots A$ affect or sentiment terms in the S vector, $\sum_{j=1}^J \sum_{a=1}^A \alpha_{3,j}^a S_{t-j}^a$; and on a residual unexplained component, ε_t . In practice, the number of lagged terms can differ for each term on the right hand side of the equation, but we have set them all to J here for notational convenience.

The dependent variable in the textual sentiment – finance literature is typically some type of firm-level or market-level performance measure such as future earnings (Li (2010)), future earnings changes (Li (2006), Li (2010)), or future returns on assets (Davis et al. (2008)). Tetlock et al. (2008) employ standardized unexpected earnings (SUE)⁸ as the dependent variable. Other dependent variables have also been employed. Kothari et al. (2009) investigate the impact of textual sentiment on three dependent variables: firms' costs of capital, the standard deviations of stock returns, and the standard deviations of analysts' forecast errors. Li (2010) employs a future liquidity measure⁹ as the dependent variable. Tetlock (2007) tried to use negative sentiment to predict the Small-Minus-Big (SMB) factor, the idea being that because small stocks have the highest individual investor ownership, if the pessimism factor proxies for small investor sentiment, it should predict the returns on small stocks. In Antweiler and Frank (2004), 15-minute interval stock returns, price volatilities, two measures of trading volumes¹⁰ and a liquidity measure¹¹, are each employed in turn as dependent variables.

Researchers using the event study methodology have also defined the dependent variable in Equation 2.1 as the cumulative abnormal return (CAR) over some event window, usually a 3-day period centered on the event date. Davis et al. (2008), Engelberg (2008), Feldman et al. (2008), Henry (2008), Demers and Vega (2010), Doran et al. (2010) and Price et al. (2012) all employ the standard event study methodology to examine the extent to which sentiment in corporate disclosures (or news articles about disclosures) impacts on firms' cumulative abnormal returns around the 'event' or during a post-event period. The calculation of abnormal returns is critical in event studies. The abnormal return is the difference between the actual return of the security over the event window and the normal return of the firm over the event window. There are different choices for modeling normal return. For example, Davis et al. (2008), Engelberg (2008), Feldman et al. (2008) and Demers and Vega (2010) follow the approach of Fama and French (1992) to construct the

⁸ $UE_t = E_t - E_{t-4}$, $SUE_t = (UE_t - \mu_{UE_t})/\sigma_{UE_t}$, where E_t is the firm's earnings in quarter t , UE_t is the unexpected earnings. The mean (μ) and standard deviation (σ) of unexpected earnings are calculated on the basis of previous 20 quarters data.

⁹ i.e. Cash flow from operations in the next four quarters scaled by the book value of current liabilities at the end of that quarter

¹⁰ The log number of small, medium and large trades and the log number of traded shares

¹¹ The daily average of the bid-ask spread

book-to-market and size-sorted portfolios. Henry (2008) chooses the return of the equal-weighted market portfolio as the normal return to calculate abnormal returns for each stock. Tetlock et al. (2008), Chen et al. (2012), Rees and Twedt (2012) and Huang et al. (2012) apply a method similar to event studies. Tetlock et al. (2008) and Chen et al. (2012) set the news stories date as the ‘event’, while Rees and Twedt (2012) and Huang et al. (2012) set the analyst reports publication date as the ‘event’. They test whether the negative sentiment measure has real impact on the firm’s next-day return or abnormal return over a certain holding period.

The $\mathbf{X}$ vector of control variables has included firm characteristics and previous market variables such as the cash flow from operations, the book-to-market ratio, the market value of equity, accruals and leverage, current earnings surprises, analyst earnings forecast revisions and dispersions, recent and more distant stock prices, returns and volatilities, and stock market index returns and trading volumes. The idea is to test if the sentiment measure still has significant predictability on the dependent variable measuring firm-level or market-level performance after controlling for the vector of fundamental variables and previous market activities. In studies that examine textual sentiment in analyst reports (Rees and Twedt (2012) and Huang et al. (2012)), control variables include the deviation of analyst’s earnings per share forecast, the deviation of analyst’s stock recommendation, the revisions of recommendation, earnings forecast and target price, and the indicator variables for recommendation upgrades, reiterations and downgrades.

Some researchers have also considered and examined the possibility of reverse causation, whereby firm-level or market-level performance can impact on future measures of sentiment. Equation 2.2 captures these possibilities in general form.

$$s_t = \beta_0 + \sum_{j=1}^J \beta_{1,j} s_{t-j} + \sum_{j=1}^J \sum_{f=1}^F \beta_{2,j}^f X_{t-j}^f + \sum_{j=1}^J \sum_{p=1}^P \beta_{3,j}^p Y_{t-j}^p + \varphi_t \quad (2.2)$$

Here, s is the dependent variable obtained as some measure of affect or sentiment from the $\mathbf{S}$ vector in Equation 2.1, and $\mathbf{Y}$ is a vector of firm-level or market-level performance variables from which the dependent variable in Equation 2.1 is obtained. Equation 2.2 states that sentiment in period t depends on a constant term, β_0 ; on $j = 1 \dots J$ lagged dependent terms, $\sum_{j=1}^J \beta_{1,j} s_{t-j}$; on $j = 1 \dots J$ lagged terms of $f = 1 \dots F$ control variables in the $\mathbf{X}$ vector, $\sum_{j=1}^J \sum_{f=1}^F \beta_{2,j}^f X_{t-j}^f$; on $j = 1 \dots J$ lagged terms of $p = 1 \dots P$ firm-level or market-level performance variables. As in Equation 2.1, we have set the number of lagged terms to J for notational convenience.

Because corporate documents contain much information about firm-level fundamentals, it is natural to hypothesize that firm fundamentals also have impact on the sentiment of corporate disclosures. Li (2006) and Li (2010) apply versions of Equation 2.2 to test whether firm characteristics in the $\mathbf{X}$ vector (e.g. book-to-market equity ratio, the market value of equity, earnings scaled by book value of assets, accruals, and stock return volatility) have explanatory power on their sentiment measure and changes in this. Das and Chen (2007) use variations of Equation 2.1 and 2.2 to examine the interrelation between the *Morgan Stanley High Tech Index* (MSH) and their sentiment measures¹². Using the same methodology, they also examine the interactions between firm-level stock prices and firm-specific sentiment. Antweiler and Frank (2004) also employ these equations. Their dependent variables in Equation 2.1 include daily stock returns, price volatilities, two measures of trading volume and one liquidity measure. Each of these is regressed on two lags of sentiment. These authors then let the sentiment measures be dependent variables as in Equation 2.2, and regress each of them in turn on lags of the market variables. In both sets of regressions, the contemporaneous stock market index is employed as a control variable. Garcia (2012) examines the feedback effect from the *Dow Jones Industrial Average* (DJIA) index returns to media sentiment, and how market-level sentiment affects future returns. This author also incorporates a dummy variable to distinguish between how the interactions differ between recession and non-recession periods. Overall, the results of

¹² Constructed based on the classification of texts by five algorithms of machine learning

these studies show that firm size, book-to-market ratio, change in earnings and accruals are amongst the most powerful variables that determine sentiment in corporate annual and quarterly reports.

Vector autoregression models

The vector autoregression (VAR) model has also been used to capture the evolution and interdependencies between firm-level or market-level performance, control variables and sentiment. The VAR is a more advanced time-series model than single equation models insofar as all variables in the VAR are determined endogenously by their own history and by the history of all the other variables in the model. The standard p -th order VAR is written as Equation 2.3.

$$A(L)Z_t + BX_t = u_t \quad (2.3)$$

with

$$A(L) = 1 - A_1L - A_2L^2 - \dots - A_pL^p,$$

$$E(u_t) = 0, \quad E(u_t u'_t) = \Sigma, \quad E(u_t u'_s) = 0, \text{ for } t \neq s, \quad E(x_t u_t) = 0,$$

$$Z_t = (Y_t, S_t),$$

and X_t being the vector of controls in Equation 2.1 and 2.2.

This is a standard VAR representation in which Z is a $(1 \times n)$ vector of endogenous variables, with n depending on how many measures of firm-level or market-level performance and semantic measures are included from the $\mathbf{Y}$ and $\mathbf{S}$ vectors in Equations 2.1 and 2.2. A is an $(n \times n)$ matrix of coefficients, u is an $(n \times 1)$ vector of white noise disturbance terms, and L denotes the lag operator (for example, $L^j x_t = x_{t-j}$). It is well established that this VAR model can be estimated by ordinary least squares, which yields consistent and asymptotically efficient estimates of the A matrix because the

right-hand-side variables are predetermined and are the same in each equation of the model.

Tetlock (2007) sets up a VAR model with 5 lags to investigate the interdependencies between negative sentiment, the *Dow Jones Industrial Average (DJIA)* returns, and the NEW York Stock Exchange (*NYSE*) trading volumes. By employing the VAR system, the author attempts to predict the *DJIA* returns using negative sentiment, to predict negative sentiment using the *DJIA* returns, and to predict *NYSE* volumes using negative sentiment. He finds that negative sentiment causes *DJIA* returns, which have feedback effects on negative sentiment, and that negative sentiment also causes trading volumes. There remains a great deal of research to be done on extending the range of variables included in VAR models, and in investigating more generally how the many variables within the $\mathbf{Y}$, $\mathbf{S}$ and $\mathbf{X}$ vectors interact.

Volatility models

Antweiler and Frank (2004) employ volatility models to test the effects of sentiment or sentiment-related variables on the second moment of stock returns. They employ both GARCH-type models and the ‘realized volatility’ approach, but only report results of the latter, as both results are consistent with each other. In their GARCH-type approach, they let the sentiment measure ‘bullishness’ enter the mean equation first, and they insert the ‘agreement’ measure, amongst other exogenous variables, into the volatility equation, making the model into GARCH-X, GJR-GARCH-X and EGARCH-X forms¹³. In the ‘realized volatility’ approach, the authors investigate the extent to which the internet postings have an impact on the realized volatility (proxied by the log of the standard deviation of intra-day returns) by incorporating the ‘agreement’ measure, the number of messages and the number of trades as explanatory variables. They find that the number of

¹³ The GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model, introduced by Bollerslev (1986), is a process of the conditional variance (volatility) of the dependent variable (stock returns, in this case). The conditional variance is expressed as its own lags and lags of the squared error term. Compared with GARCH, the GJR-GARCH model (Glosten et al.(1993)) also models asymmetry in the volatility process. The EGARCH model (Nelson (1991)) models the process of log conditional variance. In the GARCH-X, GJR-GARCH-X and EGARCH-X models, exogenous variable X is also incorporated in the volatility process.

messages is a predictive factor of volatility, but agreement among the posted messages is not significant.

Textual sentiment based trading strategies

Another intuitive way to examine the extent to which positive and negative verbal information leads to significant differences in risk-adjusted stock returns is to design trading strategies based on textual sentiment. This methodology can be considered complementary to econometric modeling to examine the extent to which textual sentiment has financial market impacts.

Tetlock et al. (2008) labeled all news stories with a fraction of negative words (Neg) in the previous year's top (bottom) quartile as negative (positive) stories. They considered a simple strategy that longs all firms with positive news stories on the prior trading day and shorts all firms with negative news stories. They modeled holding both the long and short portfolios for one full trading day and rebalancing at the end of the next trading day. They used the Fama-French three-factor (1993) and the Carhart four-factor (1997) models to adjust the trading strategy returns for the returns on the market, size, book-to-market, and momentum factors. Engelberg (2008) formed a similar trading strategy which longs firms with articles that have no negative words and shorts firms with articles that have above 5% negative words. This author modeled holding the portfolio for 80 days after the earnings announcement, and examined portfolio returns for the first and second 40 days separately. The portfolio returns were risk-adjusted by the Carhart four-factor (1997) model. Li (2006), Feldman et al. (2008), Sinha (2010) and Chen et al. (2012) also constructed similar portfolios and examined their abnormal returns. Their results generally show that the textual-sentiment based strategies can generate significant abnormal returns. It appears from this analysis that further work on the relationship between textual sentiment and financial market activity might be useful to market participants as well as researchers.

2.5 Main findings to date

The research outcomes of the textual sentiment literature include the impact of textual sentiment on firm performance, security prices and market activities. This literature also provides evidence on market efficiency. The essential findings can be summarized as follows.

Sentiment and security prices and returns

In the media-expressed sentiment literature, negative sentiment is the strongest influence. Negative sentiment or a large increase in negative sentiment causes downward pressure on market prices immediately. Tetlock (2007), Tetlock et al. (2008) and Garcia (2012) have all concluded this. Chen et al. (2012) found that the negative sentiment in internet articles is negatively associated with both contemporaneous and next-day abnormal returns. In addition, by using a new term-weighting scheme, Jegadeesh and Wu (2012) claimed that they find a significant relation between document tone and market reaction for positive words as well. Much evidence shows that optimism in today's sentiment leads to higher stock (or stock index) returns tomorrow, except Antweiler and Frank (2004), who used internet message information and discovered that a positive shock to message board postings predicts negative returns on the next day. From a portfolio perspective, Sinha (2010) discovered that negative tone in the past 6 months predicts negative returns. Similarly, positive tone in news articles in the past 6 months predicts positive returns.

Tetlock et al. (2007) discovered that pessimism measures significantly predict negative returns to the SMB factor over the following week. This provides evidence for the hypothesis that pessimism measures the sentiment of individual investors. Sinha (2010) found that there is a significant relationship between long-term sentiment and the momentum factor. The trading strategy to go short on stocks with negative past sentiment scores and to go long on stocks with positive past sentiment scores generates portfolios that are highly correlated with the momentum factor. This finding suggests that under-reaction to sentiment information is related to the momentum phenomenon.

Considering the possibility of reverse causation, do stock prices or returns predict textual sentiment? Tetlock (2007) discovered that negative returns predict more pessimism in the next day's WSJ column. Garcia (2012) found that stock returns are indeed important predictors of media sentiment, both the positive and the negative measures, and that as expected, positive market returns tend to increase positivity and decrease negativity in future news stories. However, Das and Chen (2007)'s finding is that the value of the stock index on a given day is not related to the sentiment level in internet messages on the next day.

The corporation-expressed-sentiment literature (Davis et al. (2008), Engelberg (2008), Demers and Vega (2010), Jegadeesh and Wu (2012), Price et al. (2012)) concludes that changes in the tone of corporate disclosures from the recent past are significantly correlated with short window contemporaneous returns around the date that the disclosures are made, or drift excess returns¹⁴, even after controlling for firm financial information and earnings surprises. Abnormal market returns are higher as the tone of the press release becomes more positive.

Sentiment and other market variables

Studies have also found that textual sentiment has significant impacts on trading volumes. Tetlock (2007) found that unusually high or low values of pessimism lead to temporarily high market trading volume. Antweiler and Frank (2004) discovered that internet board postings can help predict trading volumes, and using daily data showed that the effect from message posting activities to trading volumes is more significant than the reverse direction. Das and Chen (2007) concluded that there is a strong contemporaneous relationship between sentiment and trading volume. Although these authors examined internet-posting sentiment as Antweiler and Frank (2004), they conclude that sentiment has no predictability on future trading volumes.

Kothari et al. (2009) found that both positive and negative news disclosures in the media have impact on firms' cost of capital, return volatility, and analyst forecast dispersion.

¹⁴ Excess returns in a long period (e.g. 80 days) following the event.

Positive news disclosures reduce all three variables, while negative news disclosures increase them. The authors also concluded that positive news statements by management do not materially affect the firm's cost of capital.

Sentiment and firm fundamentals; information and market efficiency

Does textual sentiment relate to firm fundamentals? The literature to date shows that textual sentiment is positively correlated with future firm performance. Li (2006) found that stronger emphasis on risk in firms' annual reports is associated with lower future earnings. Li (2010) discovered that the tone of the forward-looking statements is positively correlated with future performance and has explanatory power incremental to other variables. Tetlock et al. (2008) showed that negative words in news stories forecast low subsequent firm earnings. Chen et al. (2012) found that negative sentiment 30 days to 3 days prior to earnings announcements is inversely associated with the earnings surprise. This suggests that textual sentiment is not purely noise, but represents value-relevant information.

Does the linguistic style of corporate disclosures, news articles and internet messages contain any useful incremental information over quantitative financial information? The answer is generally yes, although it must be noted that 'linguistic style' is not solely represented by sentiment. Feldman et al. (2008) found that market participants seem to behave as if they use qualitative information from MD&A disclosures (or other information that is correlated with it) in addition to the quantitative financial information provided routinely by firms. This indicates that the MD&A sections do have information content. Demers and Vega (2010) also found that management-conveyed verbal information plays a significant role in the price discovery process.

What implications does this research have for market efficiency? Tetlock et al. (2008) state that the stock market is relatively efficient with respect to firms' hard-to-quantify fundamentals, although Tetlock (2007) found little support for the hypothesis that textual sentiment represents additional fundamental information. By and large, it has been clear that sentiment in texts does convey incremental information over quantitative financial

information, and that it might have at least some potential power in predicting market movements. Sentiment from qualitative information is mostly publicly available, or possibly partly private in some cases. In either case, according to the strongest form of EMH, a good asset pricing model might well incorporate textual sentiment as a pricing factor, in addition to risk premium and firm characteristics.

Overall contributions

Although all studies that have been reviewed in this Chapter perform sentiment analysis, they have varied focuses and contribute to the accounting, economics and finance literatures from several perspectives.

Engelberg (2008), Feldman et al. (2008), Henry (2008), Demers and Vega (2010) and Doran et al. (2010) all contribute to the analysis of announcement effects on security prices, as they apply the standard event-study method to study the impact of textual sentiment in earnings announcement on asset prices. Engelberg (2008) and Demers and Vega (2010) focus on comparing qualitative information and quantitative information, trying to examine the complementarity relation between the two types of information on stock valuation. Henry (2008) dives deep into qualitative information analysis. Besides optimism and pessimism, he also examines other styles of earnings press release: length, numerical intensity, complexity, and so forth.

Li (2006), Tetlock (2007), Davis et al. (2008), Kothari et al. (2009), Tetlock et al. (2008), Li (2010), and Garcia (2012) all concentrate on the interrelations between sentiment and market activities. Tetlock (2007) and Tetlock et al. (2008) also contribute to research on information and market efficiency, as well as the role of media in financial markets. Henry and Leone (2009), Loughran and McDonald (2011) and Jegadeesh and Wu (2012) endeavor to improve the accuracy of the dictionary-based approach to content analysis. The first two studies construct word lists that are appropriate in the finance domain, while the second and third studies question the use of the proportional term-weighting scheme that has been extensively used in earlier studies. Antweiler and Frank (2004), Das and Chen (2007) and Chen et al. (2012) concentrate on the connection between internet board

posting activities and the stock market. They study not only sentiment in these messages, but also ‘agreement’ of opinions, ‘article attention’, the number of messages, and the number of words. The latest studies by Doran et al. (2010), Garcia (2012), Rees and Twedt (2012) and Huang et al. (2012) have all made novel contributions to the literature. Doran et al. (2010) investigate the effects of sentiment on REITs, Garcia (2012) studies the interrelations between sentiment and market patterns in recessions and expansions, trying to unveil the difference, and Rees and Twedt (2012) and Huang et al. (2012) extract textual sentiment from financial analysts’ reports.

The most important studies in the literature should be Tetlock (2007) and Tetlock et al. (2008), the two earliest researches on textual sentiment, and Loughran and McDonald (2011), which provides word lists in the finance domain for the most popular linguistic method, the dictionary-based approach. Figure 2.2 presents a flow-chart summary of Tetlock’s (2007) methodological approach. *First*, daily news stories from the Wall Street Journal ‘Abreast of the Market’ column over the 16-year period 1984-1999 are collected. *Second*, the principal components of sentiment in the news are identified by implementing PCA on the 77 sentiment categories in the Harvard IV-4 psychosocial dictionary. *Third*, three standardized sentiment measures (the media factors) are constructed. *Fourth*, VAR modeling is performed, incorporating the media factors, DJIA index returns, NYSE volumes and the SMB factor. *Fifth*, robustness and sensitivity analysis is performed by altering the return window and using after-morning trading volumes amongst others.

Figure 2.3 summarizes Tetlock et al. (2008). *First*, all Wall Street Journal and Dow Jones News Service stories regarding individual S&P 500 firms from 1980 to 2004 are retrieved and pre-processed. *Second*, the standardized negative sentiment measure is constructed. *Third*, use negative words from news stories of 30 to 3 trading days prior to an earnings announcement to predict earnings. Also firms’ negative sentiment in a day is utilized to predict stock returns on the following day. *Fourth*, textual sentiment based trading strategies are implemented, in order to examine the portfolio’s risk-adjusted return. *Fifth*, examine whether negative words in stories containing the word ‘earn’ predict earnings better than negative words in other stories.

Figure 2.4 summarizes Loughran and McDonald (2011). *First*, the corpus is constructed including 50,115 firm-year 10-Ks between 1994 and 2008. *Second*, the Harvard dictionary negative word list is expanded on by inflecting each word to forms that retain the original meaning, of the root word. A list of 2,337 words is created which typically have negative implications in a financial sense. The two word lists are then compared. *Third*, event study is then implemented in order to examine the effect of both negative sentiment measures in turn. Both proportional weighting and another word-weighting scheme (td.idf) are employed. Sentiment extracted from all the words in 10-Ks and from only the MD&A sections are examined separately. *Fourth*, the effects of additional finance word lists (e.g. finance positive, uncertainty) are examined using the same models established in the previous step. Besides excess return another two variables, event-period abnormal volume and post-event return volatility are employed as the dependent variable. *Fifth*, regress earnings surprise on both negative measures in turn.

2.6 Summary and conclusions

This Chapter reviews the textual sentiment literature in finance. Previous studies are categorized and discussed mainly from three aspects: information sources, content analysis methods, and financial models. The textual sentiment literature has made new contributions to different areas of finance research. It contributes to the analysis of announcement effects, as sentiment in corporate disclosures was examined rather than the financial information disclosed. It analyzes the complementarity relation between qualitative and quantitative information regarding their effects on future firm performance. It improves the linguistic analysis method so as to get more accurate and efficient sentiment measures. Finally, it studies information and market efficiency from a new perspective, in contrast to the early work of, inter alia, Mitchell and Mulherin (1994) and DeGennaro and Shrieves (1997)) who used easily quantifiable aspects of news as a proxy for the news itself, such as the timing of news, the volume of news (number of words), and the type of news (periodic announcements and general publicly available news).

What is agreed and what remains controversial in the literature, and what remains to be done? It is generally agreed that textual sentiment has a strong influence on stock returns and trading volumes. The media-expressed-sentiment literature demonstrates that textual sentiment has contemporaneous or short-term effects on stock prices, returns/abnormal returns, and trading volumes (e.g. Antweiler and Frank (2004), Tetlock (2007), Tetlock et al. (2008), Garcia (2012), Jegadeesh and Wu (2012), Chen et al. (2012)). The exception is Das and Chen (2007), who conclude that sentiment has no predictability on future trading volumes. Event studies in the corporation-expressed sentiment literature normally show that there is a significant market response to sentiment in corporate disclosures, in a short window (1 day or 3 days) around the event date (Davis et al. (2008), Engelberg (2008), Demers and Vega (2010), Jegadeesh and Wu (2012)).

In the dictionary-based approach literature, particularly the studies which use the GI dictionary or L&M word lists, negative words are found to be the most powerful sentiment category in explaining market activities. Tetlock (2007) emphasizes that negative words summarize common variation in the entire set of GI word categories better than any other single category including positive words. Tetlock et al. (2008) find that using positive words as the sentiment measure produces much weaker results than using negative words. However, the more recent study by Jegadeesh and Wu (2012) has pointed out that by using a new term-weighting scheme, a significant relation between document tone and market reaction is discovered for positive words as well, so they emphasize the importance of choosing the appropriate term-weighting scheme.

Stock returns are found to predict textual sentiment by Tetlock (2007) and Garcia (2012), who use news stories as the information source. Yet Das and Chen (2007) reveal that the value of the stock index on a given day is not related to the sentiment level in internet messages on the next day. The work by Li (2006), Tetlock et al. (2008), Feldman et al. (2008), Demers and Vega (2010), Li (2010), Garcia (2012), Chen et al. (2012), Huang et al. (2012), Price et al. (2012), and Rees and Twedt (2012) consistently shows that textual sentiment contains new value-relevant information. The exception is Tetlock (2007), who

rejected the hypothesis that media content contains new information about fundamental asset values.

In light of what has been done, what is agreed and what remains controversial in the literature, the following directions for future research are suggested. *First*, the content analysis process can be further improved. Although the L&M word lists have been increasingly popular in the latest research, for the dictionary-based approach, the publishing of complete authoritative field-specific dictionaries is desirable for future studies. Moreover a better term-weighting scheme should be developed. Other content analysis approaches can also be developed and refined by linguists, psychologists and computer scientists. The purpose of improving the content analysis procedures is to obtain more accurate and efficient sentiment measures, and to examine whether other sentiment categories can be as powerful as negative words.

Second, there are still some qualitative information sources that have not yet been widely studied. For example: speeches, blogs, twitter, TV programs and videos amongst others. Speeches from the most influential individuals in the economics and finance fields are particularly worth studying. *Third*, volatility models have rarely been used. For example, the GARCH-X, GARCH-M-X models can be applied to examine the direct and indirect relationships between sentiment and volatilities. They may also assist in discovering the mechanism through which sentiment in texts gets transmitted to asset prices and returns. *Fourth*, the extant textual sentiment studies concentrate on the stock market. The same methodology can be applied on other financial assets, such as commodities and bonds. Meanwhile, it is interesting to study global financial markets other than the U.S., especially emerging markets, and to analyze qualitative information in other languages, as different markets may display differing patterns. *Fifth*, with the help of computer programs which have the ability to source a virtually continuous flow of news and analyze sentiment automatically, researchers will be able to examine sentiment's effect in real time. This is exciting because one can view and analyze the price movement and trading volume immediately following a piece of extremely positive or negative news. *Finally*, the relationship between textual sentiment and investor behaviors requires further research.

Tetlock (2007) states that 'it is unclear whether the financial news media induces, amplifies, or simply reflects investors' interpretations of stock market performance.' (pp.1139). It is of great importance to add to the literature on whether textual sentiment predicts investor sentiment, or whether it reflects past investor sentiment and how investors interpret and react to textual sentiment.

Figure 2.1: Sentiment extraction (dictionary-based approach)

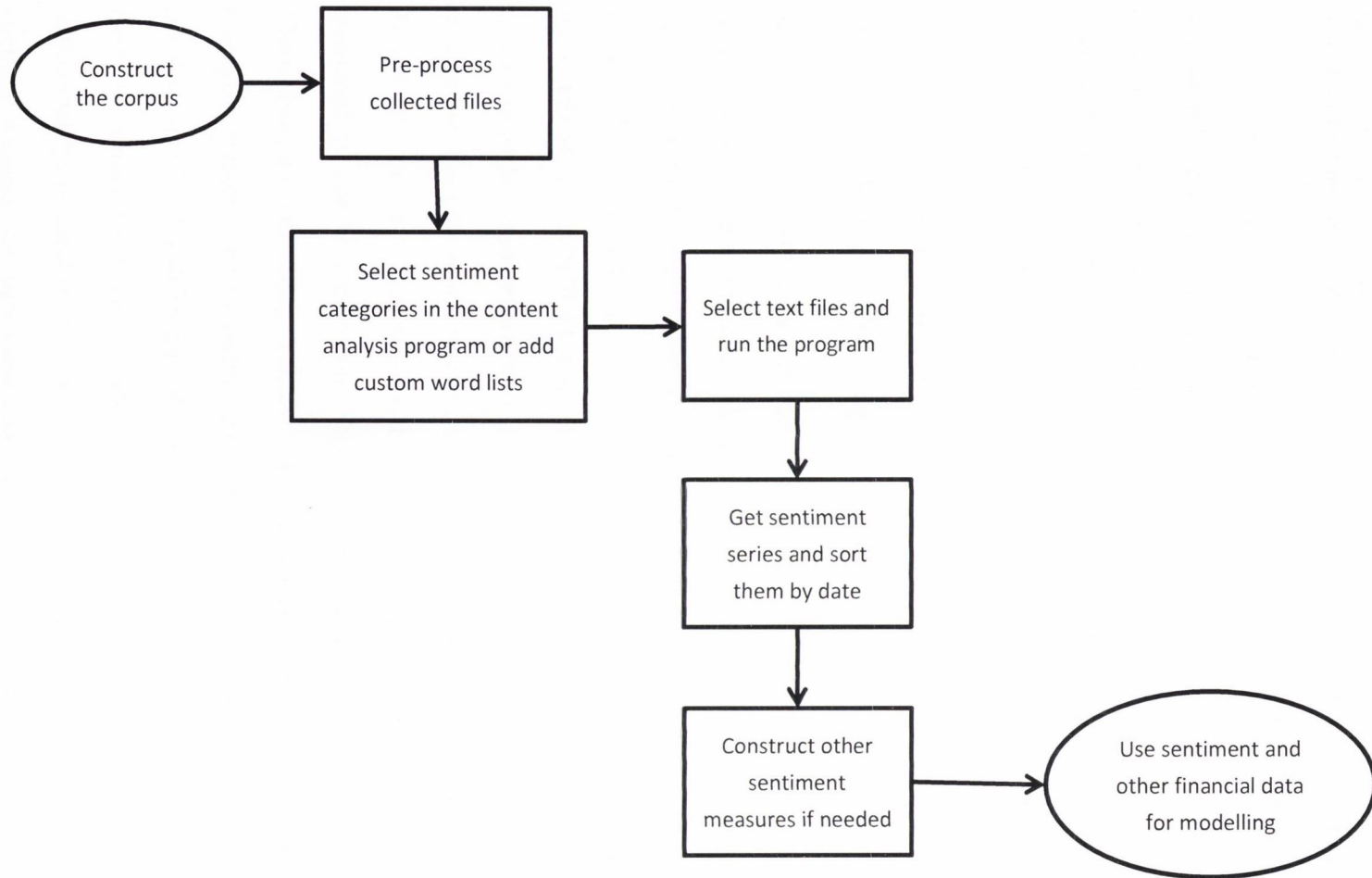
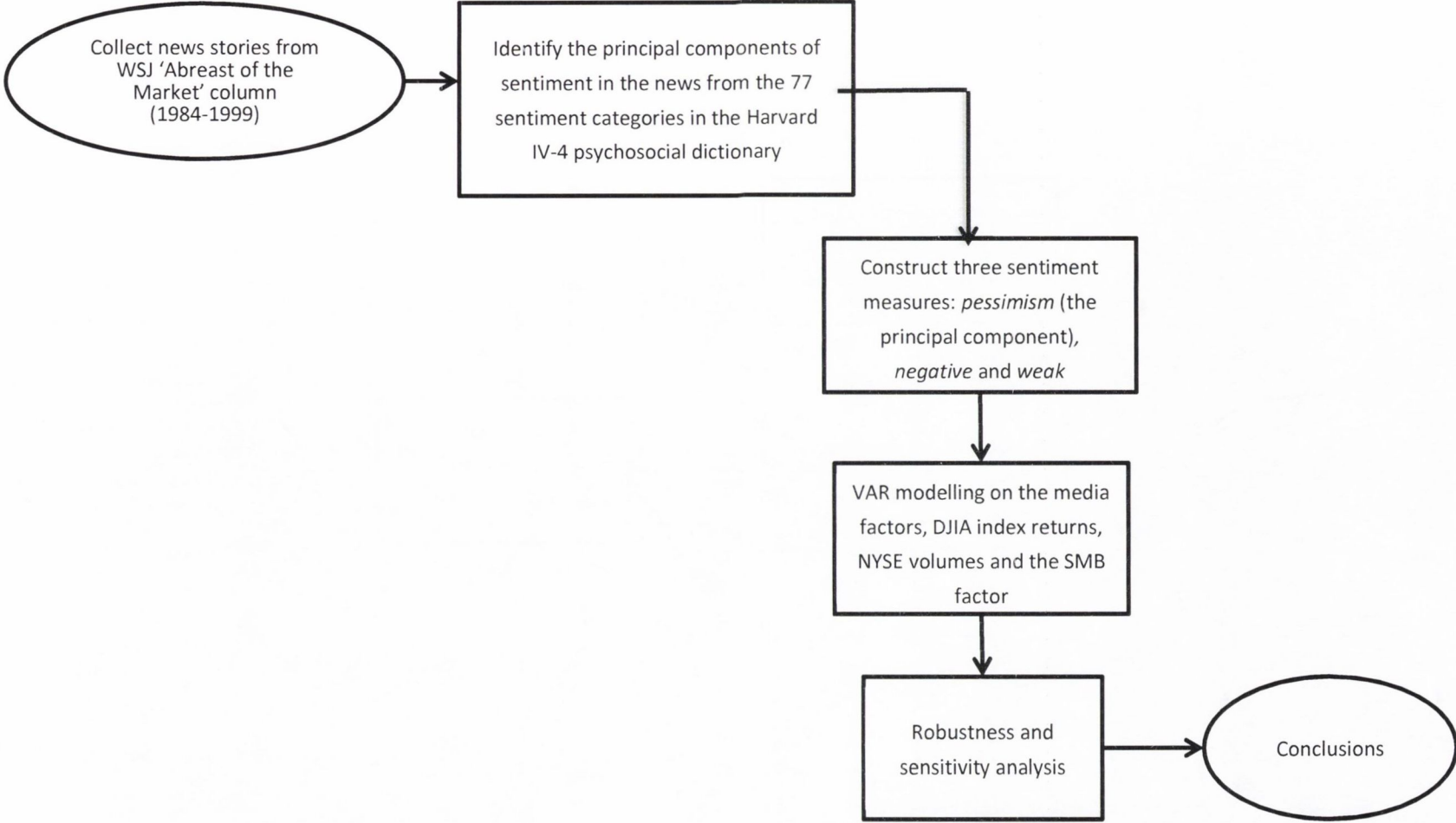


Figure 2.2: Summary of Tetlock (2007)



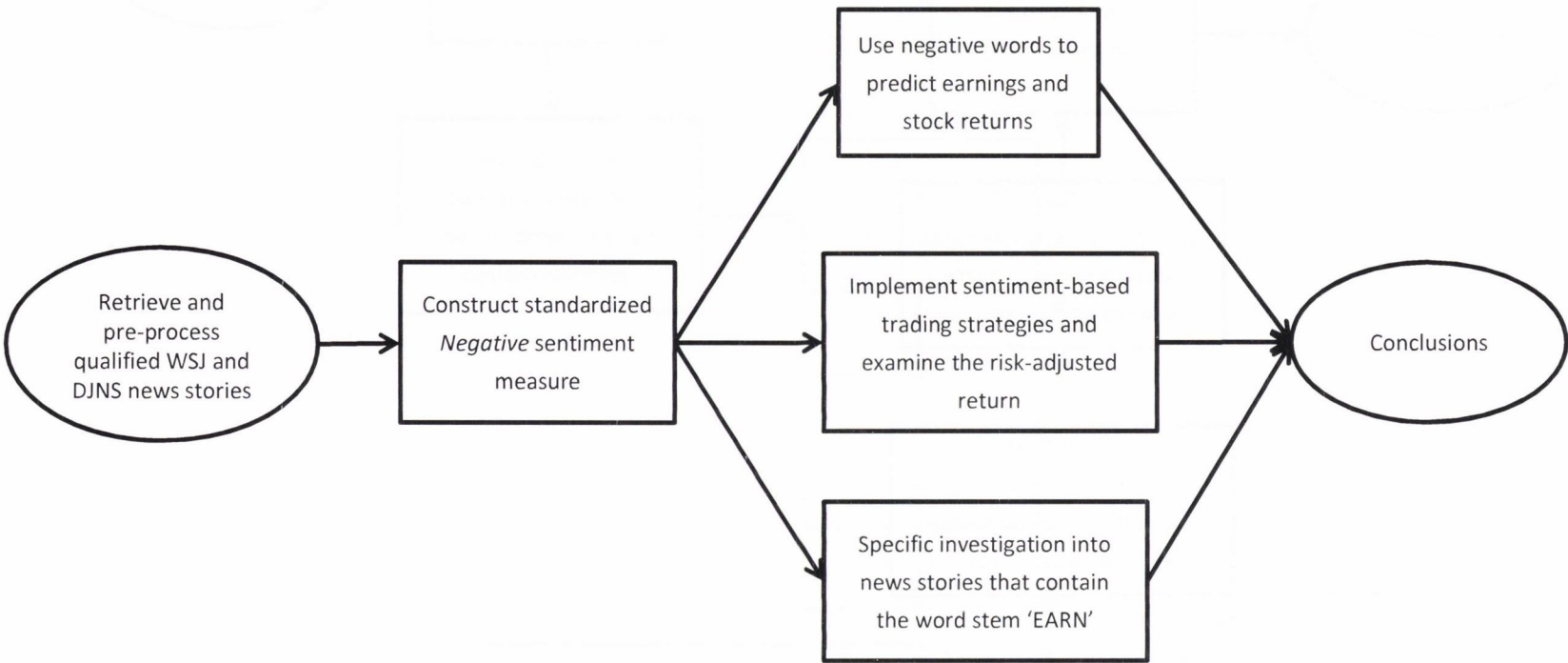


Figure 2.3: Summary of Tetlock et al. (2008)

Figure 2.4: Summary of Loughran and McDonald (2011)

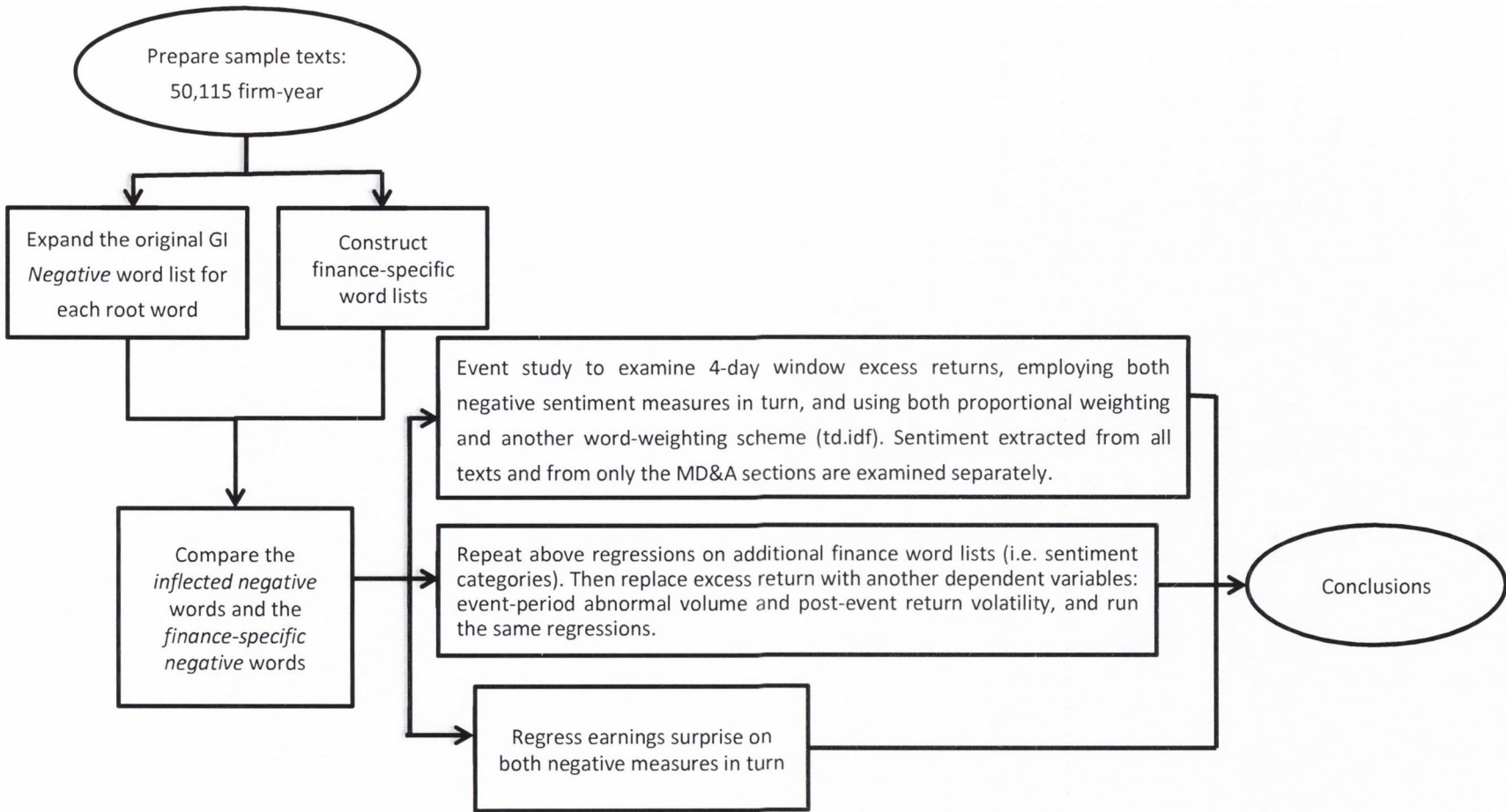


Table 2.1: Qualitative information sources

Corporation-expressed Sentiment		Media-expressed Sentiment (News stories, comments, analysts' reports)	Internet-expressed Sentiment (Board postings)
<i>Annual Reports/10-Ks</i>	<i>Earnings press Releases/ Earnings conference calls</i>		
Li (2006) Feldman et al. (2008) Li (2010) Loughran and McDonald (2011) Jegadeesh and Wu (2012)	Davis et al. (2008) Henry (2008) Demers and Vega (2010) Henry and Leone (2009) Doran et al. (2010) Price et al. (2012)	Tetlock (2007) Tetlock et al. (2008) Engelberg (2008) Sinha (2010) Garcia (2012) Rees and Twedt (2012) Huang et al. (2012)	Antweiler and Frank (2004) Das and Chen (2007) Chen et al. (2012)
Combination: Kothari et al. (2009)			

Notes: This table summarizes the information sources used in the literature. They are corporate disclosures, news stories and analyst reports, and internet board postings.

Table 2.2: Content analysis methods

Dictionary-based approach	<i>DICTION</i>	Davis et al. (2008); Henry (2008)
	<i>GI</i>	Engelberg (2008); Feldman et al. (2008); Tetlock (2007); Tetlock et al. (2008); Kothari et al. (2009); Price et al. (2012); Rees and Twedt (2012)
	<i>Custom dictionary</i>	Henry and Leone (2009); Loughran and McDonald (2011); Garcia (2012); Doran et al. (2010); Jegadeesh and Wu (2012); Chen et al. (2012); Price et al. (2012)
Machine learning	Antweiler and Frank (2004); Das and Chen (2007); Li (2010); Sinha (2010); Huang et al. (2012)	
Other	Li (2006)	

Notes: This table displays the content analysis methods employed in the literature.

METHODOLOGY

3.1 Introduction

The methodology of this thesis contains two parts. The first part is content analysis, which is employed to quantify the level of sentiment in the texts. Content analysis first requires an appropriate text corpus to be formed and a linguistic analysis method to be used to gauge the sentiment. In this thesis, the dictionary-based approach (also referred to as the ‘bag-of-words’ model) is used in the content analysis, and the text corpora for analysis in equity and sovereign debt markets consist of firm-specific or country-specific news stories and media articles respectively.

The second part of the methodology of this thesis is performing economic and econometric analysis to test hypotheses, obtain results and draw conclusions about how the sentiment interacts with financial market activities. The economic and econometric methodology includes univariate time-series models, vector autoregressions, rolling-window regressions, panel data regressions and related hypotheses testing.

The rest of this chapter is organized as follows. Section 3.2 introduces the content analysis method. It describes the model, text corpora and the dictionary. Section 3.3 presents the

major econometric models, methods, and tests used in the thesis. Section 3.4 summarizes the chapter.

3.2 Content analysis method

This thesis uses the dictionary-based approach to extract sentiment from the texts. This approach uses a ‘mapping’ algorithm and assigns each word (or phrase) in a document into different categories based on some predefined dictionaries (Li (2010)). This is also known as the ‘bag-of-words’ model in natural language processing. Documents are considered as a bag of words, and all the structure and linear ordering of words within the context is ignored (Manning and Schutze 1999: 237).

In order to generate firm-specific or country-specific sentiment, the foremost issue is to collect a corpus that is a representative sample of the population of interest. A sample is representative if what we find for the sample also holds for the general population (Manning and Schutze 1999:119). For special language such as finance and economics (the case of this thesis), as a rule of thumb, the representative corpus usually requires tens of thousands of words (Wright and Budin (2001): 736). The corpora for this thesis have well met this requirement: in Chapter 4, the corpus for each firm ranges from 7 to 52 million words, and in Chapter 5, the corpus for each country ranges from 2 to 6 million words. These corpora were created according to the name identity of firms or countries, and keywords (e.g. firm names, public finance terms). The detailed descriptions of creating corpora and generating the firm-specific or country-specific sentiment are presented in Chapter 4 Section 4.2.1 and Chapter 5 Section 5.3.1.

The dictionary used in the thesis is Loughran and McDonald’s (2011) (L&M) finance dictionary¹, because it has been shown to be more appropriate in the business and finance contexts than general dictionaries, and have already become popular in the latest studies (e.g. Loughran and McDonald (2011), Jegadeesh and Wu (2012), Garcia (2012), Rees and Twedt (2012)). The L&M finance dictionary consists of all words that occur in 5% or more

¹ http://nd.edu/~mcdonald/Word_Lists.html

of the 10-Ks from 1994 to 2008 (Loughran and McDonald (2011)). This thesis uses their ‘finance negative’ (*FN*) word list only. The *FN* word list contains 2,349 words that are typically negative in financial contexts. Examples of words include: abandon, barrier, cancel, confuse, dispute, overestimate, preclude, and resign. The complete list of *FN* words is attached in Appendix 4. This thesis employs proportional weighting, which considers every word in the list to be equally important.

3.3 Models and Tests

The major models used in the thesis are Vector Autoregression (VAR) models (in Chapter 4) and panel regression models (both Chapter 4 and 5). This section introduces these models, how they are estimated, and inference procedures.

3.3.1 VAR Model and Related Tests

A VAR is a k -equation, k -variable model in which each variable is in turn explained by its own lagged values, plus (current) and past values of the remaining $k-1$ variables. It is a system of univariate autoregression models. A VAR model is used to capture the evolution and the interdependencies between multiple time series.

The VAR model used in this thesis is written as follows:

$$A(L)x_t + B\mu_t = \varepsilon_t \tag{3.1}$$

with

$$A(L) = 1 - A_1L - A_2L^2 - \dots - A_pL^p$$

$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t \varepsilon_t') = \Sigma, \quad E(\varepsilon_t \varepsilon_s') = 0 \quad \text{for } t \neq s,$$

$$E(\varepsilon_t \varepsilon_t) = 0,$$

and

$$\mu_t = \text{constant}$$

This is a standard (or unstructured, reduced-form) VAR representation where x is a $k \times 1$ vector of endogenous variables, A is a $k \times k$ matrix of coefficients, ε is a $k \times 1$ vector of white noise error terms, μ is a $k \times 1$ vector of constants, L is the lag operator (e.g. $L^i x_t = x_{t-i}$), and p is the lag length.

Rearranging Equation 3.1, it becomes

$$x_t = -B\mu_t + A_1x_{t-1} + A_2x_{t-2} + \dots + A_px_{t-p} + \varepsilon_t \quad (3.2)$$

The standard VAR representation does not show explicitly any concurrent dependence between the component series. As there are no time- t endogenous variables on the right hand side of Equation 3.2, no variable has a direct contemporaneous effect on other variables in the model. However in a structural VAR, contemporaneous values of endogenous variables are incorporated as regressors, and the error term (also referred to as the structural innovation) is mean zero and serially uncorrelated. Because this thesis does not use a structural VAR, its equation is not displayed here.

Equation 3.1 or 3.2 can be estimated using Ordinary Least Squares (OLS), which yields consistent and asymptotically efficient estimates of the A matrix because the vector x contains only predetermined variables, and the error terms are assumed to be serially uncorrelated and to have constant variance (Enders (1994): 301).

The advantage of the VAR system in contrast to individual univariate autoregressive models is that the system incorporates feedback since each endogenous variable is allowed to affect each other. Particularly using impulse response functions², one can assess how long and how much a one-time shock to one endogenous variable could impact the other endogenous variables. Together with the Granger causality test (see below), a VAR could be used to test if one or more variables have predictive ability to forecast the variable(s) of interest. Moreover, the standard VAR can be easily estimated using the classic OLS, and the OLS estimates have the merits that they are consistent and asymptotically efficient.

² Impulse response function is not used in this thesis, so it is not described in details in this section.

VAR models also have weaknesses. *First*, although a VAR can be a good forecasting model, it is ‘a-theoretical’ because the specification of the standard (reduced-form) VAR is determined purely on the information contained in the available data. Generally, the system of equations is not implied by any economic or financial theories. *Second*, there are a large number of parameters to be estimated. Each matrix A_i contains k^2 coefficients, and together with a total of k constants, $k + pk^2$ terms need to be estimated. *Third*, the standard methods of statistical inference may give misleading results if some of the variables are highly persistent (Stock and Watson (2001)). *Fourth*, ‘standard VARs miss nonlinearities, conditional heteroskedasticity and drifts or breaks in parameters.’ (Stock and Watson (2001): 110)

Granger causality test

The Granger causality test is used to examine whether the lags of one variable enter into the equation for another variable. The null hypothesis is that all the coefficients of the lags of the former variable are equal to zero. The F-statistic is calculated and compared with the critical value. For example, a three-equation VAR(p) system is written in the following form:

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} A_{11}^1 & A_{12}^1 & A_{13}^1 \\ A_{21}^1 & A_{22}^1 & A_{23}^1 \\ A_{31}^1 & A_{32}^1 & A_{33}^1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \dots + \begin{bmatrix} A_{11}^p & A_{12}^p & A_{13}^p \\ A_{21}^p & A_{22}^p & A_{23}^p \\ A_{31}^p & A_{32}^p & A_{33}^p \end{bmatrix} \begin{bmatrix} x_{t-p} \\ y_{t-p} \\ z_{t-p} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix} \quad (3.3)$$

To test if variable y causes x , the null hypothesis is:

$$A_{12}^1 = A_{12}^2 = \dots = 0$$

The F-statistic is calculated as:

$$F = \frac{(R_{UR}^2 - R_R^2)/q}{R_{UR}^2/(T-n)} \quad (3.4)$$

where R_{UR}^2 is R^2 from the unrestricted regression model, R_R^2 is R^2 from the restricted regression model, q is the number of restrictions (e.g. if test $A_{12}^1 = A_{12}^2 = A_{12}^3 = 0$, then $q=3$), T is the number of observations, and n is the number of explanatory variables in

the unrestricted regression (including the intercept). If the F-statistic is greater than the critical value, then the null is rejected, and we conclude that y causes x .

How can we decide the proper specification of a VAR model? The first factor to consider is how many lags of the endogenous variables should be included in the system. Specifically in this thesis, although a system containing three endogenous variables (i.e. return, volume and sentiment) is expected, based on our intention to examine sentiment's effects in the system, there is a need to test whether sentiment should enter into the model, namely whether sentiment has significant effects on either returns or volumes, and vice versa. This leads to the lag-length test and the block exogeneity test.

Lag-length test

There are two ways to decide the optimal number of lags to be included in a VAR model. The first is the likelihood ratio (LR) test. The LR test statistic is:

$$LR = (T - m)(\ln|\Sigma_r| - \ln|\Sigma_u|) \sim \chi^2(q) \quad (3.5)$$

Where T is the number of usable observations, m is the number of parameters estimated in each equation of the unrestricted system, including the constant. $\ln|\Sigma_r|$ is the natural log of the determinant of the covariance matrix of residuals of the restricted system. q is the total number of restrictions in the system. The null hypothesis is that the restricted system is true. If the LR statistic is greater than the critical value, one should reject the null, and the unrestricted system (the one including more lags) is better.

The other way to find the optimal number of lags in VAR models is using the Akaike information criterion (AIC) or the Schwarz criterion (SBC).

$$AIC = T\ln|\Sigma| + 2N \quad (3.6)$$

$$SBC = T\ln|\Sigma| + N\ln T \quad (3.7)$$

Here, Σ is the determinant of the covariance matrix of residuals of the system, and N is the total number of regressors in the system. The researcher should choose the number of

lags that minimizes the information criteria. The two criteria should only be used as complements to the LR test, because they are not tests, but mainly indicate the goodness of fit of each system. In this thesis, we use LR tests to decide the lag length of VAR models in Chapter 4.

Block exogeneity test

The block exogeneity test is used to determine whether lags of certain variables enter each equation. It can be thought as a ‘block exclusion’ test. If the lags of a variable can be excluded from any equation except the one with itself as the dependent variable, the variable should not be included in the VAR system. The way to perform the test is to estimate a VAR with the lags of the variable or variables of interest (unrestricted system) and another without the lags (restricted system). The test is in fact a likelihood ratio test, with test statistics the same as Equation 3.5. If the LR statistic is greater than the critical value, then we reject the null of restricted system, and the lags of that variable should enter the equations. We use this test in Chapter 4.

3.3.2 Panel regression estimation and related tests

Panel data refers to multi-dimensional data. It is multi-dimensional in the sense that it contains data of different entities (e.g. countries, firms, households, schools) over time. The datasets used in both Chapter 4 and Chapter 5 of this thesis can be used as panel data because they contain cross-sectional observations of different companies or countries over a period of time.

There are several advantages to using panel data over conventional cross-sectional or time-series datasets.

First, panel data includes a larger number of observations than a conventional time-series dataset. This is particularly useful when the time-series data for individual entities does not have sufficient observations. For example, a country has only 60 monthly observations for each variable. The dataset of a single country is far from enough for robust statistical

analysis. Panel data analysis has greater capacity for capturing the complexity of economic relationship and generating more accurate predictions for individual outcomes by pooling the data (Hsiao (2007)). Particularly, in Chapter 4 of this thesis, it will be impossible to examine the role of textual sentiment from a relatively broader perspective if panel data is not used. Conventional time-series models are only able to investigate individual stocks, which are only a very small part of the market and have little implications for the whole market. Similarly, in Chapter 5, individual countries have limited data points. Using panel data greatly increases the number of observations.

Second, panel data analysis provides more accurate inference of model parameters (Hsiao et al. (1995)). Compared with cross-sectional or time-series data, panel data generally has more degrees of freedom, hence improving estimation efficiency.

Third, panel data models can control for individual heterogeneity (Hsiao (2003)). Different firms, countries, individuals and other entities are heterogeneous. The entity-specific characteristics should be controlled for, otherwise the obtained results would be biased.

Fourth, besides entity-specific characteristics, there may be a few entity-invariant or time-invariant factors that also affect the dependent variable, no matter whether they are observable. Panel data analysis is able to control for these variables whereas a cross-sectional or time-series study cannot (Baltagi (2005)).

Fifth, panel data simplifies computation and statistical inference for non-stationary time series (Hsiao (2007)). If time-series data are non-stationary, the large sample distributions of the least-squares and maximum likelihood estimators are no longer normally distributed (Dickey and Fuller (1979)). In a panel dataset, even if the time series of certain variables are not stationary, the central limit theorem indicates that the limiting distributions of many estimators remain asymptotically normal (Levin et al. (2002); Im et al. (2003)) as long as observations among cross-sectional units are independent.

Panel data analysis also has a few limitations. Firstly, it is obviously more costly to collect panel data than to collect cross-sectional or time-series data, because it will take more time

to design the sample constitution, find and obtain the required data. The problem is more serious if the collection of data requires interviewing people, issuing surveys and manually imputing data. Another possible limitation is that ‘macro panels on countries or regions with long time series that do not account for cross-country dependence may lead to misleading inference’ (Baltagi 2005: 8). In this thesis, this problem is solved by using the Driscoll-Kraay (1998) standard error in appropriate panel regressions, which is robust to cross-sectional dependence. The other limitations of panel data identified in the literature (e.g. Baltagi 2005, Chapter 1) are associated with surveys, responses and informants’ biases. These issues are irrelevant to this thesis, because the thesis does not use surveys or interviews to collect data.

There are several models/estimation methods for panel regressions. The three models used in this thesis are pooled ordinary least squares (OLS), the fixed-effects and the random-effects models. I describe each of them in turn.

Pooled OLS estimation

Compared with the fixed-effects and random-effects models, estimation by pooled OLS does not control for unobserved cross section heterogeneity. So technically it is not a panel data method. Consider the following linear panel model:

$$y_t = X_t\beta + u_t, \quad t = 1, 2, \dots, T \quad (3.8)$$

where y_t is a $N \times 1$ vector of dependent variables at time t , X_t is a $N \times K$ matrix of explanatory variables, β is a $K \times 1$ vector of coefficients and u_t is a $N \times 1$ vector of innovations. N is the number of entities (e.g. firms, countries), T is the time periods, and K is the number of explanatory variables.

There are several assumptions for using pooled OLS (POLS) to estimate Equation 3.8 (Wooldridge 2002, Chapter 7).

POLS.A1: $E(X_t'u_t) = 0, t = 1, 2, \dots, T$

POLS.A2: $\text{rank}[\sum_{t=1}^T E(X_t'X_t)] = K$

POLS.A3: (a) $E(\mathbf{u}_t^2 X_t' X_t) = \sigma^2 E(X_t' X_t)$, $t = 1, 2, \dots, T$, where $\sigma^2 = E(\mathbf{u}_t^2)$ for all t ;

(b) $E(\mathbf{u}_t \mathbf{u}_s' X_t' X_s) = 0$, $t \neq s$, $t, s = 1, 2, \dots, T$

POLS.A1 states that innovations are uncorrelated with the regressors for all t , but says nothing about the relationship between x_s and u_t , for $s \neq t$. **POLS.A2** essentially rules out perfect linear dependencies among regressors in order to ensure that β is identified and can be estimated. Part (a) of **POLS.A3** is a fairly strong homoscedasticity assumption. It stipulates that not only the conditional variance of residuals does not depend on X_t , but also that the unconditional variance is the same for all t . Part (b) of **POLS.A3** is essentially that the conditional correlation of residuals is zero. Together with the homoscedasticity assumption, this implies that the covariance of the residuals will be $\sigma^2 \mathbf{I}_T$.

Under these assumptions, the pooled OLS estimator is consistent and asymptotically normal. However, if some of the conditions in the assumptions do not hold, there will be significant problems for statistical inference. In these cases, the robust variance (standard errors) estimates should be used. For example, if the homoscedasticity assumption does not hold, the coefficient estimates β are not affected, but the standard errors are usually unreliable estimates, so the t-statistics will be unreliable. In this case, the Huber-White standard errors (Huber (1967) and White (1980, 1982)) should be used to calculate the t-statistics. Similarly, if the assumption of the conditional correlation of residuals does not hold, clustered standard errors (such as Liang and Zeger (1986)) will be more appropriate in calculating t-statistics so as to make valid statistical inference. In this thesis, both the Huber-White standard errors and the clustered standard errors are used with the pooled OLS estimation.

The fixed effects model

Consider the following panel data regression model presented in a different way:

$$y_{it} = X_{it} \beta + u_{it} \quad i=1, 2, \dots, N; \quad t=1, 2, \dots, T \quad (3.9)$$

where i denotes different entities (e.g. firms, countries), and t denotes time. The i subscript therefore denotes the cross-sectional dimension, while t subscript denotes the time-series dimension. X_{it} is a $1 \times K$ vector of explanatory variables, and β is $K \times 1$. Equation 3.9 and 3.8 are essentially the same, but 3.9 more clearly indicates a particular cross section observation of individual i . It is not necessary to use subscript i in Equation 3.8 because the pooled OLS method does not control for any cross section heterogeneity.

Most of the panel data applications use the following one-way error component model for the innovations:

$$u_{it} = \mu_i + v_{it} \quad (3.10)$$

with μ_i denoting the unobservable individual-specific effect or individual heterogeneity, and v_{it} denoting the remaining errors. For the fixed effects (FE) model, μ_i is assumed to be a fixed parameter and it accounts for any individual-specific effect that is not included in the regressors. The remaining error term v_{it} varies with entities and time, thus it can be thought of as the usual error term in the regression. The FE model allows μ_i to be correlated with the regressors X_{it} .

The fixed effects model is appropriate if we are focusing on a specific set of entities, such as firms or countries. Substitute the innovations given by Equation 3.10 into 3.9, we could get

$$y_{it} = X_{it}\beta + \mu_i + v_{it} \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T \quad (3.11)$$

Because μ_i is unobservable, it cannot be directly controlled for. However, μ_i can be eliminated by using the *within* transformation, which is demeaning each variable. Because μ_i is constant, its effect is eliminated by demeaning. This yields the *within model*:

$$\ddot{y}_{it} = \ddot{X}_{it}\beta + \ddot{v}_{it} \quad (3.12)$$

where $\ddot{y}_{it} = y_{it} - \bar{y}_i$, $\ddot{x}_{itk} = x_{itk} - \bar{x}_{it}$ and $\ddot{v}_{it} = v_{it} - \bar{v}_{it}$.

Then the FE estimator of β , denoted as $\hat{\beta}_{FE}$ is obtained by using OLS estimation.

In this thesis, the FE model is used in Chapter 5 to examine if textual sentiment is a determinant of sovereign bond yield spreads, controlling for the fixed country effects.

The random effects model

If μ_i is treated as a random variable, then Equation 3.9 becomes a random effects (RE) model. Some believe that the RE model is appropriate if we are drawing N individuals randomly from a large population. However, Wooldridge (2002) argues that this is wrongheaded for microeconomic panel data applications, because with a large number of random draws from the cross section, it always makes sense to treat the unobserved effects as random draws from the population (pp.286). The key issue is that whether μ_i is uncorrelated with the explanatory variables X_{it} . If μ_i is assumed to be uncorrelated with X_{it} , then the model is referred to as RE.

The RE model is estimated by the method of feasible general least squares (GLS). The assumptions for RE model are as follows (Wooldridge 2002, Chapter 10):

RE.A1: (a) $E(v_{it}|X_i, \mu_i) = 0, t = 1, 2, \dots, T$; (b) $E(\mu_i|X_i) = E(\mu_i) = 0$

where $X_i \equiv (X_{i1}, X_{i2}, \dots, X_{it})$.

Part (a) of **RE.A1** is a strict exogeneity assumption. This is an assumption in addition to the orthogonality assumption (part (b)) as with pooled OLS. Under **RE.A1**, we could get $E(u_{it}|X_i) = 0$, which is a strict exogeneity assumption required by GLS.

RE.A2: $\text{rank } E(X_i' \Omega^{-1} X_i) = K$.

where Ω is the unconditional variance matrix of u_i .

Under **RE.A1** and **RE.A2**, GLS and feasible GLS (FGLS) estimators are consistent, but a standard random effects analysis adds assumptions on the idiosyncratic errors v_{it} that give Ω a special form.

RE.A3: (a) $E(v_i v_i' | X_i, \mu_i) = \sigma_v^2 \mathbf{I}_T$; (b) $E(\mu_i^2 | X_i) = \sigma_\mu^2$.

Part (a) of **RE.A3** implies that 1) v_{it} has a constant unconditional variance σ_v^2 across t , and 2) v_{it} is serially uncorrelated. Combination of part (b) of **RE.A1** and part (b) of **RE.A3** indicates that $\text{Var}((\mu_i | X_i) = \text{Var}(\mu_i)$, which is a homoscedasticity assumption on the unobserved individual effect μ_i .

Under the above assumptions, the RE model is estimated by feasible GLS. The feasible GLS estimator, denoted as $\hat{\beta}_{RE}$, is both consistent and efficient. In this thesis, the RE model is used in Chapter 4, together with the pooled OLS estimation, to investigate the immediate effects of textual sentiment on individual stock returns and trading volumes.

Hausman's specification test

How can we decide between the FE and the RE model? Hausman (1978) suggests a specification test where the null hypothesis is that $E(u_{it} | X_{it}) = 0$, which is true for RE but not for FE. This is a critical hypothesis given that the innovations u_{it} contain individual effects (μ_i) which are unobserved and may be correlated with X_{it} .

The test is performed by comparing $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$ for the subset of coefficients of time-varying variables. The test statistics is:

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' [Var(\hat{\beta}_{FE}) - Var(\hat{\beta}_{RE})]^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}) \quad (3.13)$$

Under H_0 , the statistics is asymptotically distributed as χ_K^2 , where K denotes the dimension of β . Under H_0 , $\hat{\beta}_{RE}$ is consistent and efficient and $\hat{\beta}_{FE}$ is consistent but inefficient; under H_A , $\hat{\beta}_{RE}$ is inconsistent but $\hat{\beta}_{FE}$ remains consistent. If the p-value of the test is smaller than 0.05 (i.e. significant), FE model should be chosen over RE. In this

thesis, the Hausman's specification tests are used in both Chapter 4 and 5, in order to decide between the FE and the RE model.

Panel unit root tests

Panel unit root testing emerged from time series unit root testing. It is employed on panel data to detect if the variables are stationary and appropriate for modelling so as to avoid invalid statistical inferences. This is particularly important because there is no evidence that the cross-sectional units in our datasets are independent, hence central limit theorem could not be invoked, and the conventional Wald test statistics cannot be approximated well by t or chi-square distributions (e.g. Dickey and Fuller (1979)). The major difference between panel unit root tests and unit root tests on conventional time series is that we have to consider the asymptotic behaviour of both the time-series dimension T and the cross-sectional dimension N . The way in which N and T tend to infinity is crucial for determining asymptotic properties of estimators and tests proposed for nonstationary panels (Baltagi (2008): 275). Because panel unit root testing is not a major part of the thesis' methodologies, this section briefly summarizes the differences of the tests which are used in this thesis instead of describing each of them in detail.

Consider the following simple panel-data model for a single variable with a first-order autoregressive component:

$$y_{it} = \rho_i y_{i,t-1} + \mathbf{z}'_{it} \gamma_i + \epsilon_{it} \quad (3.14)$$

where i denotes the individuals, t denotes the time as usual. y_{it} is the variable being tested, and ϵ_{it} is a stationary error term. The $\mathbf{z}_{it}$ could represent the panel-specific mean, or panel-specific mean plus time trend, or nothing at all.

Panel unit root tests are used to test the null hypothesis $H_0: \rho_i = 1$ for all i against the alternative $H_a: \rho_i < 1$. Equation 3.14 is often written as

$$\Delta y_{it} = \phi_i y_{i,t-1} + \mathbf{z}'_{it} \gamma_i + \epsilon_{it} \quad (3.15)$$

with the null hypothesis $H_0: \phi_i = 0$ for all i against the alternative $H_a: \phi_i < 0$.

There are several panel unit root tests. The four tests that are used in this thesis are the Levin-Lin-Chu (2002) (LLC) test, the Breitung (2000) test, the Harris–Tsavalis (1999) (HT) test and the Im–Pesaran–Shin (2003) (IPS) test. The LLC, Breitung and HT tests make the simplifying assumption that all panels share the same autoregressive parameter so that $\rho_i = \rho$ for all i . The IPS test, however, allows the autoregressive parameter to be panel specific. The major difference between panel unit root tests and time-series unit root tests is that the asymptotic behaviour of the time-series dimension T and the cross-sectional dimension N should be considered. The four aforementioned tests make differing assumptions about the rates at which the number of panels, N , and the number of time periods, T , tend to infinity or whether N or T is fixed. For example, a standard³ LLC test requires $N/T \rightarrow 0$ for the test statistics to have a well-defined asymptotic distribution; the HT test and the IPS test without the accommodations for serial correlation assumes $N \rightarrow \infty$ and T is fixed; a standard⁴ Breitung test works best with ‘large’ T and at least ‘moderate’ N .

3.4 Summary and Conclusions

This chapter describes the methodologies used in this thesis. Firstly the content analysis method and the dictionary used are described. Next, the major econometric models and methods used in Chapters 4 and 5 are described, including the VAR model, the VAR lag length test, the block exogeneity test, panel estimation by pooled OLS, the fixed effects and random effects models, the Hausman’s specification test, and the panel unit root tests. The rationale of using the models and the tests, their strengths and limitations, and under what assumptions they work well are also elaborated.

Overall, the choice of data, model and testing methodologies are a fundamental part of any research programme. Each choice involves compromises between advantages and disadvantages, and I have summarized them in this chapter.

³ With a panel-specific mean term and no linear time trend in the model

⁴ With a panel-specific mean term and no linear time trend in the model

CHAPTER 4

RESULTS – CASE STUDY OF EQUITY

4.1 Introduction

Most papers that use news stories as the information source have studied market-level sentiment (e.g. Tetlock (2007), Garcia (2012)); they extract sentiment from general economics and finance news stories and assess its effect from a market-wide perspective. Few studies have looked into news stories concerning individual firms' firm-specific sentiment. This raises interesting questions: What are the overall effects of firm-specific sentiment on individual stocks? Is the role of firm-specific sentiment similar to that of market-level sentiment? Does the relationship between sentiment, equity returns and trading volumes evolve over time? Is there an indirect relationship between sentiment and equity returns? What is the implication of the role of firm-specific sentiment on market efficiency?

In this chapter, we explore these questions by examining the interrelations among firm-specific textual sentiment, firm-level equity returns and trading volumes for 20 large non-financial firms from the *Fortune 500* list over the 10-year period from January 2001 to December 2010. The firm-specific sentiment is extracted from firm-specific news stories and media articles. We first employ panel data regressions to test six hypotheses regarding the relation between firm-specific sentiment, firm-level equity returns and trading volumes.

Next, we employ VAR models to test the hypotheses on individual firms separately. Finally, we use 1-year rolling-window regressions to examine the time-varying effects of firm-specific sentiment on firm-level equity returns, and further we explore the ‘indirect’ effects of sentiment on returns. Robustness analysis is completed by using different estimation methods, winsorizing the data, or changing sentiment measures.

This research contains some novel features. *First*, by generating daily, firm-specific, textual sentiment, we show how it relates to firm-level trading volumes and equity returns. This research fills a gap in the literature as nobody has hitherto examined the time-series pattern of sentiment’s role for individual stocks. *Second*, we study the relation between firm-specific textual sentiment, trading volumes and firm-level returns in a manner consistent with the mixture of distributions hypothesis (Clark (1973)) that allows us to shed light on the extent to which firm-specific textual sentiment acts as a previously unconsidered source of information that drives firm-level returns by first working through trading volumes. *Third*, we extend previous research by showing that firm-specific textual sentiment has time-varying effects on firm-level trading volumes and returns that have not been previously examined.

Amongst the main findings of this chapter are that negative firm-specific textual sentiment predicts future firm-level returns; better firm-level returns predict less subsequent negative firm-specific textual sentiment; and greater trading volumes are associated with greater subsequent negative textual sentiment. Moreover, there is strong evidence of ‘indirect’ effects of sentiment on equity returns: sentiment causes trading volumes, whilst trading volumes cause returns. Besides, firm-specific textual sentiment has time-varying effects on firm-level returns that tend to be concentrated during discrete periods that most likely align with significant newsworthy episodes for each firm. Overall, our analysis constitutes strong evidence for the consideration of firm-specific textual sentiment as a potentially important time-varying factor in equity pricing models.

The rest of the chapter is organized as follows. Section 4.2 describes the process of creating corpora and generating the firm-specific sentiment, and contains the descriptions of the textual sentiment and equity data. Section 4.3 outlines the hypotheses regarding the

interrelations between our measures of firm-specific textual sentiment, trading volumes and firm-level returns. Section 4.4 tests the hypotheses by treating the dataset as a panel. Three groups of regressions are studied. The first group examines whether sentiment and trading volumes cause equity returns, the second examines whether returns and trading volumes cause sentiment, and the third examines whether returns and sentiment cause trading volumes. Section 4.5 examines each of the 20 individual firms, testing the hypotheses both over the full sample period and over rolling-window samples. Section 4.6 summarizes our main findings and draws together our conclusions.

4.2 Sentiment and equity data descriptions

4.2.1 Generating firm-specific sentiment and descriptive statistics of sentiment data

The first step in sentiment extraction is to select the sample of firms. The firms being considered are large multinational corporations (MNCs), which are more likely to have a stable news flow at daily frequency for an intended 10-year research period between 01/01/2001 and 31/12/2010. The most common database of large MNCs used in international business and financial research is the Fortune 500 list. We select 20 firms from the top 50 non-financial public firms on the 2011 Fortune 500 list¹. We omit financial firms because of their differing fundamentals compared with non-financial firms, and because their news stories may contain much information regarding transactions with their client companies, making it more difficult to separate the firm-specific firm sentiment. There are a total of 17 companies meeting the criteria (i.e. stable daily news flow², non-financial firm, top 50 Fortune 500 firm): Apple, AT&T, Boeing, Cisco Systems, Dell, Ford Motor, General Electric, Hewlett Packard, Home Depot, IBM, Intel, Johnson & Johnson, Merck, Microsoft, Pfizer, Verizon Communications, and Wal-mart Stores. Another 3 companies have sufficient news stories only in the latest five or six years. They

¹ http://money.cnn.com/magazines/fortune/fortune500/2011/full_list/

² Although for some companies there are a few days, usually less than 10 days throughout the 10-year period, having no articles

are ConocoPhillips and Chevron (ranges from 01/01/2006 to 31/12/2010), and Exxon Mobil (ranges from 01/01/2005 to 31/12/2010). These 3 firms are also incorporated in the sample. Therefore, the final sample consists of 20 large MNCs.

The next step is to search for relevant articles during the pre-specified time range to form a text corpus from which to extract the firm-specific sentiment. The electronic news database used is ‘*LexisNexis News and Business*’, a popular news database. Figure 4.1 shows the power search page of the *LexisNexis* electronic news database. We search for articles that contain the firm name in the headline, and at least 5 mentions of the firm name in the body. By choosing the option ‘Strong references only’, the retrieved articles are ensured to be highly relevant to the firm³. The source we choose is ‘All English Language News’, which includes articles from newspapers, newswires and press releases, magazines, journals and web-based publications. The purpose is to retrieve as many qualified articles as possible to ensure there is at least one article each day, and preferably multiple articles to get unbiased sentiment scores. Choosing all available English-language articles also lowers the possibility of getting biased sentiment from a small number of pre-specified sources. Unlike Tetlock (2007) and Garcia (2012) who use only one or two news commentary columns from the *Wall Street Journal* or the *New York Times* (as described in Chapter 2) as the information source, our texts have the advantage of being more objective because they are based on a much wider set of news article. We omit articles with high similarity to stories that have been previously published. The downloaded files are individual text files containing up to 500 articles.

A custom ‘splitter’ program is then used to split the text files into individual articles. This program also extracts date and time information for each article. We move all the articles on weekends to Friday, and when there are multiple stories in one day, they are considered to be one story per day. Articles on holidays are omitted because few market participants pay attention to the very few, if any, firm-specific news stories that appear on holidays. In addition, our examinations confirm that barely any articles are collected on holidays. Table 4.1 reports the average number of articles and words collected per day for each firm in our

³ The returned results are at least of 80% relevance, indicated by the relevance score

sample. Because the 20 firms are large firms, they each have at least 10 news articles per day. Microsoft and Boeing even have more than 40 articles every day.

We use *Rocksteady*, a content analysis program developed by Treocht Ltd⁴, to calculate the daily sentiment scores by counting the frequency of words in the Loughran and McDonald's (2011) 'Finance Negative' (*FN*) category relative to the total words in a day. By setting the sentiment score to be zero on days where there is no article, the final *FN* series are obtained. Figure 4.2 presents plots of the daily time series of *FN* sentiment for each firm. Through visual examination a few possible spurious outliers⁵ were found for each firm. A possible explanation for these outliers is that on some days there was little news so very few words were collected resulting in extreme fractions of 'finance negative' words in the texts.

Table 4.2 summarizes the descriptive statistics of the *FN* series, including mean, variance, skewness, excess kurtosis, the Jarque-Bera (J-B) test of normality, the Ljung-Box Q test for autocorrelation, and the Dickey-Fuller unit-root test. The skewness and excess kurtosis of a normal distribution are expected to be zero. The J-B column denotes the p-value of tests against the null hypothesis that the *FN* data is normally distributed. Clearly, none of the *FN* series is normally distributed. The Q10 column displays the p-value of tests for autocorrelation up to the 10th lag. All results indicate that the *FN* series are serially correlated. Dickey-Fuller unit-root tests were also calculated to test that whether each of the *FN* series is stationary. With the absolute values of all t-statistics greater than 25, the existence of a unit root is easily rejected for each series.

4.2.2 Descriptive statistics of equity data

We collected the daily closing prices and trading volumes of each of our 20 sample stocks from *Datastream*. The data range is from 2001-2010, with exceptions for Exxon Mobil from 2005-2010 and for Chevron and ConocoPhillips from 2006-2010. The closing prices

⁴ Treocht Ltd.,(www.treocht.com) is an Ireland-based company developing a web-based system that uses fusion analytics to deliver instant, accurate financial predictions.

⁵ To avoid the effect of outliers, sentiment series are winsorized at the 1 percent level in a few cases of our robustness analysis reported as below

are adjusted for dividends and splits, and the trading volumes have also been adjusted for capital events and divided by 1,000. The time series plots of the log returns and volumes for each company are presented in Figures 4.3 and 4.4. Based on visual examination of the plots, several possible spurious outliers were observed. Further examination shows that there is no evidence that these outliers are errors. In addition, the outliers mostly exist immediately after the ‘9.11’ event or in October 2008 due to the subprime crisis. Winsorization is used in a few cases⁶ to avoid the possible effect of outliers and for robustness analysis.

Table 4.3 provides descriptive statistics of the daily firm-level stock returns and trading volumes, including skewness, excess kurtosis, the Jarque-Bera (J-B) test of normality, the Ljung-Box Q test for autocorrelation and the Dickey-Fuller unit-root test. The skewness and excess kurtosis of a normal distribution are both expected to be zero. The J-B column denotes the p-value of tests against the null hypothesis that the data is normally distributed. The results reveal that none of the return or volume data is normally distributed. The Q10 column displays the p-value of tests for autocorrelation up to the 10th lag. The results indicate that all the volume series are serially correlated. All of the return series except Merck is shown to be serially correlated at the 10 percent significance level. The Dickey-Fuller unit-root tests for each of the return and volume series, with the critical value of the 1 percent significance level at -3.438, show that the existence of a unit root is easily rejected for each series.

Table 4.4 shows the market cap⁷, annual revenue⁸, average annual return (simple return), average daily trading volume, and average daily *FN* scores for each firm. The firms are sorted by the average *FN* scores, from high to low. This is only a rough examination of the relationship between negative sentiment and returns. Pfizer is the firm with the lowest average annual return. Its associated average *FN* score is the highest, which indicates that it had the most negative words in the firm-specific news, throughout the 10-year period. Wal-mart and Merck also have relatively low annual return whilst their associated *FN*

⁶ Sections 4.4.2 and 4.5.2

⁷ As on 15/11/2011

⁸ Data shown on 2011 Fortune 500 list

scores are among the highest.

4.3 Main hypotheses

As is known from the literature, market-level textual sentiment (in particular negative words) is associated with market returns, and the relationship between sentiment and market trading volumes is also significant. When it comes to firm-specific sentiment, firm-level returns and trading volumes, it is interesting to discover the extent to which these relationships also exist. In addition, it is interesting to examine whether firm-specific textual sentiment acts as a potential source of information that drives firm-level returns by first working through trading volumes. Therefore, the following hypotheses are tested:

$H_1(Null)$: Firm-specific sentiment does not cause firm-level equity returns.

$H_1(Alternative)$: Firm-specific sentiment does cause firm-level equity returns.

$H_2(Null)$: Firm-level equity returns do not cause firm-specific sentiment.

$H_2(Alternative)$: Firm-level equity returns cause firm-specific sentiment.

$H_3(Null)$: Firm-specific sentiment does not cause trading volumes.

$H_3(Alternative)$: Firm-specific sentiment causes trading volumes.

$H_4(Null)$: Firm-level trading volumes do not cause firm-specific sentiment.

$H_4(Alternative)$: Firm-level trading volumes cause firm-specific sentiment.

$H_5(Null)$: Firm-level trading volumes do not cause firm-level equity returns.

$H_5(Alternative)$: Firm-level trading volumes cause firm-level equity returns.

$H_6(Null)$: Firm-level equity returns do not cause trading volumes.

$H_6(Alternative)$: Firm-level equity returns causes trading volumes.

Tetlock (2007) has tested, at the market level, whether negative sentiment causes equity returns, whether equity returns cause negative sentiment, and whether negative sentiment causes trading volumes (corresponding to H_1 , H_2 , and H_3). Some other researches (e.g. Antweiler and Frank (2004), Das and Chen (2007), Tetlock et al. (2008), Garcia (2012), Chen et al. (2012)) have also examined versions of these three hypotheses. We examine

H_4 , H_5 , and H_6 in addition to H_1 , H_2 , and H_3 , and we consider a potential mechanism: sentiment indirectly causes returns by working through trading volumes. All the tests of these hypotheses are 2-tailed.

4.4 Empirical testing on panel data

One of the challenging issues in firm-specific sentiment analysis is that it is usually difficult to collect news stories for a large number of firms. Even if this were possible with the help of computer programs, time-series analysis at the daily frequency would be difficult as there would be many missing sentiment data points for individual firms, because it is rare that a firm would have news stories every day. In order to ensure there are minimum missing values for sentiment in this research, our sample consists only of large MNCs that usually attract large volumes of news. Combining the firm-level data and performing panel estimations proves a useful approach to exploring the overall market patterns. This section tests the hypotheses described in the previous section by running several panel regressions.

4.4.1 The data used in the panel estimates

To facilitate the construction of a strongly balanced panel, only firms with 10-year news data are incorporated in the dataset, resulting in a total of 17 firms and 2,515 observations per firm. This is the largest balanced panel that can be obtained from the whole datasets, yielding 42,704 observations. Besides returns and sentiment, several new variables are employed in the panel estimations. Firstly, volume is measured by ‘turnover’. Lo and Wang (2000) define turnover of stock j at time t as:

$$Turnover_{jt} = \frac{X_{jt}}{N_j} \quad (4.1)$$

where X_{jt} is the share volume of firm j that is traded at time t , and N_j is the total number of shares outstanding of firm j .

Three fundamental variables for each firm are also incorporated: the standardized unexpected earnings (SUE), the Book-to-Market ratio (B/M) and the market value (MV). Following Tetlock et al (2008), the SUE is transformed from each firm's quarterly earnings as follows:

$$SUE_t = \frac{UE_t - \mu_{UE_t}}{\sigma_{UE_t}} \quad (4.2)$$

$$UE_t = E_t - E_{t-4} \quad (4.3)$$

where E_t is the firm's earnings in quarter t , and μ_{UE_t} and σ_{UE_t} are the mean and standard deviation of the firm's previous 20 quarters of unexpected earnings data, respectively.

In order to ensure that the series are stationary, B/M and MV^9 , which are daily variables, are detrended in the following way: the variable is demeaned to obtain a residual, square this residual, and subtract the past 60-day moving average of the squared residual.

Three panel unit-root tests, the Levin-Lin-Chu (2002) (LLC) test, the Breitung (2000) test, and the Harris-Tsavalis (1999) (HT) test, are performed on FN , Turnover, SUE, B/M and MV. As described in Chapter 3 Section 3.3.2, each test has its own advantages and disadvantages to consider and there is no dominant performance of one particular test. In our panel dataset, N is moderate ($N=17$), T is large ($T=2,515$), and $N/T \rightarrow 0$. Under these conditions, the LLC test and the Breitung test work best. Table 4.5 reports the statistics and their significance levels (1 percent ***). The null hypotheses that the panels contain unit roots are rejected in all cases except one, the LLC test for SUE. Based on the overall results, we consider that all variables are appropriate for modelling.

4.4.2 Panel Estimations

In this section, three groups of models are tested in order to examine the relations between firm-specific sentiment, firm-level returns and trading volumes.

⁹ MV was divided by 10,000

Do sentiment and volumes cause returns?

Firstly, in line with Tetlock et al (2008), the pooled ordinary least squares (OLS) regressions using standard errors clustered by calendar quarter are performed. The model is written as:

$$R_{i,t} = \beta_{10} + \sum_{k=1}^2 \beta_{1k} R_{i,t-k} \text{ (or } AR_{i,t-k}) + \beta_{13} S_i + \beta_{14} V_{i,t-1} + \beta_{15} SUE_{i,t} + \beta_{16} (B/M)_{i,t} + \beta_{17} MV_{i,t} + \beta_{18} \cdot Exog_{t-1} + \varepsilon_{i,t} \quad (4.4)$$

where $i = 1, 2, \dots, 17$, β_{10} is the constant, R_i is the firm-level stock returns, S_i is the sentiment measure, AR_i is the abnormal returns, and V_i is the firm-level trading volumes. Firm characteristic variables SUE^{10} , B/M and MV (all described in Section 4.4.1) are employed as control variables. *Exog* includes three exogenous variables: volatility measure proxied by CBOE's volatility index (VIX)¹¹ and dummy variables for the Monday and January effects.

We begin by using the first lag of *FN* as the sentiment measure. The results show that *FN* negatively and significantly predicts returns on the following day at the 5 percent significance level in both the first and second regressions (Table 4.6 column (1) and (2)). The second regression replaces the lags of returns in column (1) by the lags of abnormal returns, which is the return adjusted for the three Fama-French (1993) factors. The impact of a 10 percent increase in the percentage of negative words on the next day's stock return is approximately -0.31 or -0.36 basis points. Volumes only marginally significantly predict returns on the following day in regression (2), and higher volumes are associated with higher returns on the following day.

We perform the following robustness tests. In our first robustness check, another panel

¹⁰ the value at the end of the preceding quarter

¹¹ Obtained from Datastream. The VIX index is detrended in the following way: we demean the variable to obtain a residual, square this residual, and then subtract the past 30-day moving average of the squared residual.

estimation method, the random effects model¹², is employed for the same regressions. The random effects model is the same as Equation 4.4 except that the innovations contain a random unobservable individual-specific effect besides the conventional white noise error terms. In regression (3), *FN* still negatively and significantly predicts returns on the following day at the 5 percent level (Table 4.6 column (3)). Yet in regression (4), where the lags of abnormal returns are used instead of the raw returns, sentiment's predictability becomes marginally significant. The magnitude of effects is the same as the pooled OLS regressions ((Table 4.6 column (4)). Volumes' effects on next day's returns become strongly significant, at the 5 percent level in regression (3) and 1 percent level in regression (4).

In our second robustness check, we use the change in *FN* (first difference) as the sentiment measure, and repeat regressions (1) to (4). In all regressions, a 10 percent increase of the differential of today's and yesterday's *FN* score has a negative impact on today's return at approximately 0.65 to 0.66 basis points, at the 1 percent level of significance. Higher volumes predict higher returns on the following day, at least at the 10 percent level. Table 4.7 reports the results.

Our third robustness analysis is completed by winsorizing *FN*, returns/abnormal returns, and trading volumes at the 1 percent level to avoid the impact of outliers, and re-do regressions (1) to (4). The estimated coefficients and t-statistics (not reported) of *FN*(-1), no matter whether the pooled OLS or random effects model are used, are qualitatively the same as in Table 4.6.

Do returns and volumes cause sentiment?

To examine the effects of firm-level returns and trading volumes on firm-specific sentiment, the following model is tested:

¹² The random effects model was selected in preference to the fixed effects model, according to the result of Hausman's specification test (1978) (described in Chapter 3 Section 3.3.2). A sample of the test results are reported in Table A.2 in the Appendix.

$$S_{i,t} = \beta_{20} + \sum_{k=1}^2 \beta_{2k} S_{i,t-k} + \beta_{23} R_{i,t-1} + \beta_{24} V_{i,t-1} + \beta_{25} SUE_{i,t} + \beta_{26} (B/M)_{i,t} + \beta_{27} MV_{i,t} + \beta_{28} \cdot Exog_{t-1} + \varepsilon_{i,t} \quad (4.5)$$

where all variables are defined in the same way as in Equation 4.4. Similarly, this model is first estimated using pooled OLS with standard errors clustered by calendar quarter. Then it is estimated using fixed effects model¹³ with Huber–White robust standard errors as a robustness check. The fixed effects model is the same as Equation 4.5 except that the innovations contain a fixed unobservable individual-specific effect besides the conventional white noise error terms. The results show that higher returns predict less negative words on the following day at the 1 percent level of significance, after controlling for the past sentiment, firm characteristics and exogenous variables (Table 4.8). One percent increase in returns would lead to around 0.6 percent decrease in negative sentiment on the following day. Trading volumes do not show any significant impact on next day's sentiment. We also winsorize *FN*, returns/abnormal returns, and trading volumes at the 1 percent level, and re-do the two regressions. The results (not reported) are qualitatively the same.

Do sentiment and returns cause volumes?

To examine the effects of firm-specific sentiment and firm-level returns on trading volumes, the following model is tested:

$$V_{i,t} = \beta_{30} + \sum_{k=1}^2 \beta_{3k} V_{i,t-k} + \beta_{33} R_{i,t-1} + \beta_{34} S_{i,t-1} + \beta_{35} SUE_{i,t} + \beta_{36} (B/M)_{i,t} + \beta_{37} MV_{i,t} + \beta_{38} \cdot Exog_{t-1} + \varepsilon_{i,t} \quad (4.6)$$

Again, this model is first estimated using pooled OLS with standard errors clustered by calendar quarter. Then it is estimated using fixed effects model¹⁴ with Huber–White robust standard errors as a first robustness check. The results show that higher returns predict lower trading volumes on the following day at the 1 percent level of significance, whatever

¹³ The fixed effects model was selected in preference to the random effects model, according to the result of Hausman's specification test, which is reported in Table A.2 in the Appendix.

¹⁴ The fixed effects model was selected in preference to the random effects model, according to the result of Hausman's specification test, which is reported in Table A.2 in the Appendix.

estimation method is used (Table 4.9). In the pooled OLS regression, negative sentiment is negatively associated with next day's trading volumes at the 1 percent significance level. However, in the fixed effects estimation, the impacts of negative sentiment become marginally significant. Our second robustness check is completed by winsorizing *FN*, returns/abnormal returns, and trading volumes at the 1 percent level, and re-do the two regressions (results not reported). This time, in the fixed effects estimation, the impacts of negative sentiment on trading volumes are significant at the 5 percent level. Other results remain qualitatively the same.

4.5 Empirical testing on individual firms

This section tests each of the hypotheses presented in Section 4.3 by examining each of the 20 firms in turn, using VAR models and rolling-window linear regressions, and the relevant diagnostic tests associated with these models.

4.5.1 Empirical testing over full data period

This part conducts tests for each firm on the interrelations among firm-specific sentiment, firm-level returns and trading volumes over the full sample period, using VAR models.

VAR lag length tests

Based on the assumption that returns, trading volumes and sentiment are interrelated with each other, the three variables are considered to be endogenous. An important factor to consider when using VAR models is to choose the right lag length. For this thesis, it is assumed that the maximum lag length is no more than 5 trading days as news articles published over a week ago are not expected to have significant impact on market activities today. Therefore, lag length 1 to 5 will be tested against each other using the LR test statistics (Equation 3.5), leading to a total of 10 tests – '5 vs 1, 2, 3 and 4', '4 vs 1, 2, and 3', '3 vs 1 and 2', and '2 vs 1'. In the LR tests, the unrestricted system is the VAR model with the greater lag length, while the restricted system is the model with the smaller lag

length. The null hypothesis is that the restricted system is true. If the test statistic is greater than the critical value, the null hypothesis is rejected.

Test results demonstrate that with the exception of Chevron, the optimal lag length for all other 19 stocks is 5 at the 1 percent level of significance. Chevron's optimal lag length is 4 at the same significance level. Table A.1 in the Appendix reports the LR statistics and significance levels for Apple and Chevron. Results of the other stocks are not reported because they are all similar to Apple's. Results based on the Akaike information criterion (AIC) are mainly consistent with the LR tests, although results based on Schwarz criterion (SBC) are mostly different. However, since the two criteria should only be used as complements to the LR tests, the computed AIC and SBC values are not reported in the thesis, and the lag length is selected according to the LR tests.

VAR modeling over the full sample

The specific equations in the VAR(5)¹⁵ model are presented as below:

$$R_t = \alpha_1 + \beta_1 \cdot L5(R_t) + \gamma_1 \cdot L5(V_t) + \delta_1 \cdot L5(S_t) + Exog_{t-1} + \varepsilon_{1t} \quad (4.7)$$

$$V_t = \alpha_2 + \beta_2 \cdot L5(V_t) + \gamma_2 \cdot L5(R_t) + \delta_2 \cdot L5(S_t) + Exog_{t-1} + \varepsilon_{2t} \quad (4.8)$$

$$S_t = \alpha_3 + \beta_3 \cdot L5(S_t) + \gamma_3 \cdot L5(R_t) + \delta_3 \cdot L5(V_t) + Exog_{t-1} + \varepsilon_{3t} \quad (4.9)$$

Where R , V and S represent firm-level equity returns, firm-level trading volumes and firm-specific sentiment. $L5$ is defined as the lag operator that transforms any variable x_t into a row vector $[x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}]$. $Exog$ includes three exogenous variables: volatility measure proxied by the CBOE's volatility index (VIX)¹⁶ and dummy variables for the Monday and January effects.

¹⁵ The model is VAR(4) for Chevron.

¹⁶ Obtained from Datastream. The VIX index is detrended in the following way: we demean the variable to obtain a residual, square this residual, and then subtract the past 30-day moving average of the squared residual.

Table 4.10 displays the results of estimating the VAR models, focusing on the F-statistics (with their significance levels in brackets). The column headed ‘Usable obs’ provides the number of usable observations for each firm. The three columns headed ‘Return’, ‘Volume’ and ‘Fin Neg’ denote the dependent variable in the vector of endogenous variables, and the three columns beneath are the explanatory variables in these equations. The F-statistics with their marginal significance levels indicate the joint significance of all lags of each explanatory variable in each equation. The most evident relations are that returns cause volumes (16 out of 20 cases are significant at the 5 percent level), and that volumes cause sentiment (15 cases are significant at the 5 percent level). The results are also relatively supportive for ‘Sentiment causes volumes’, with 8 cases being significant at the 5 percent level. The other relations are not very evident. Table 4.11 summarizes the number of significant cases in the causality tests.

It is important to note that although only 4 firms: Dell, GE, Verizon and Exxon Mobil, show that sentiment directly causes return, 3 firms: Ford Motor, GE and Home Depot, demonstrate indirect causality from sentiment to returns; that is to say sentiment causes volumes whilst volumes cause return (Table 4.12). A total of 6 firms or 30 percent, show that sentiment causes stock returns either directly or indirectly over the full sample period. Among them, GE reveals both a direct and indirect relationship.

4.5.2 Empirical tests over rolling-window periods

The previous two groups of models (i.e. panel models and VARs) are employed to test hypotheses H_1 to H_6 , from the perspectives of both the pooled sample and individual firms (Section 4.4.2 and Section 4.5.1). This section focuses on examining H_1 , and further explores the indirect relationship between sentiment and returns: sentiment causes volumes whilst volumes cause return (H_3 , H_5) by implementing regressions over 1-year (252-day) rolling windows (arbitrary choice). We have discovered in the previous analysis, especially in panel regressions, that firm-specific sentiment predicts equity returns. We now aim to investigate whether the effects of sentiment are time-varying. We propose another hypothesis:

$H_7(Null)$: Firm-specific sentiment does not cause firm-level equity returns in any periods.
 $H_7(Alternative)$: Firm-specific sentiment causes firm-level equity returns in some periods.

In line with Tetlock (2007), we assume that the error terms ε_t in Equation 4.7 to 4.9 are independent. This allows us to estimate each equation using the ordinary least squares (OLS) techniques separately. We first focus on Equation 4.7 to examine the time-varying direct effects of the five lags of FN on firm-level equity returns, and together with Equation 4.8 to examine the indirect effects from sentiment to stock returns.

Figure 4.5 displays the results of rolling regressions (Equation 4.7) for each firm, specifically whether the five lags of FN significantly forecast stock returns. The horizontal axis in each graph illustrates the date at which the individual 252-day sliding window ends. The vertical axis indicates whether the five lags of FN significantly forecast returns in the individual regression. A value of 1 on the vertical axis indicates significance at the 10 percent level, while a value of 2 indicates significance at the 5 percent level. From the graphs it is clear that with the exception of Chevron, sentiment predicts stock returns directly for at least some of the data period, although there is no fixed pattern of periods where predictability is concentrated.

Further analysis is performed to examine whether sentiment indirectly affects stock returns at some other data periods, namely, whether the lags of sentiment cause volumes and volumes cause returns. The results demonstrate that on the whole FN is more likely to have direct effects on returns, although this happens less than 10 percent (9.56 percent) of the time on average (Table 4.13 panel (a)). The indirect effects are negligible or completely absent for several firms, such as Apple, AT&T, Boeing, Chevron, ConocoPhillips, Dell, and Pfizer. On average, either or both of the direct and indirect effects appear 11.32 percent of time. Panel (a) of Table 4.13 also shows that the two effects do not usually overlap. For example, for HP, IBM, Intel, Johnson & Johnson, Merck and Verizon, the proportion of either direct or indirect effects is equal to or almost equal to the sum of the two respective proportions. Panel (b) further summarizes the descriptive statistics of the results, including the mean, standard deviation, minimum, first quartile, median, third

quartile and maximum. The null hypothesis that the mean equals zero is rejected as the T-statistics are all greater than the critical value 2.86 at the 1 percent level of significance (last column of panel (b)).

We perform the following robustness analysis. Firstly, the firm-level returns are replaced by firm-level abnormal returns, and the interrelations among textual sentiment, volumes and abnormal returns are examined. Abnormal returns are the firms' raw returns minus the return of a value-weighted portfolio with similar size/book-to-market-characteristics¹⁷. We then run the 252-day rolling-window regressions as before to examine the time-varying patterns, except that now the dependent variable in Equation 4.7 is replaced by firm-level abnormal returns. Compared with the raw returns, the direct effects of sentiment on the abnormal returns decrease slightly, but the standard deviation is significantly smaller and the minimum, median and maximum values are greater (Table 4.14 panel (a)-(b)). The direct effects increase greatly for Chevron and Pfizer. The two firms' sentiment has exposed little or no effect on raw returns. However, the overall indirect effects become weaker. By and large, the effects on abnormal returns and raw returns are similar. Since there is no major difference between sentiment's effects on raw returns and abnormal returns, only raw returns are employed as the dependent variable in the remaining robustness analysis.

Our second robustness analysis limits the effects of outliers by winsorizing the time series of returns, trading volumes and sentiment at the 1 percent level. Three exogenous variables are still included in the model to control for volatility and potential return anomalies, and only direct effects are examined. The results reveal that winsorizing the data does not lead to any major differences (Table 4.15). On average, the direct effects with the winsorized data increase slightly, from 9.56 percent to 9.98 percent of time. It is noteworthy, however, that the direct effects have risen for 16 of the 20 firms, while IBM has the largest increase by 66.7 percent.

Our third robustness analysis tests another sentiment measure: the count of positive words

¹⁷ The time series of the benchmark portfolios' (2×3) returns are obtained from Kenneth R. French's data library at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

minus the count of negative words, divided by the sum of positive and negative word counts. The daily sentiment is relatively positive if the measure is greater than 0, or relatively negative if less than 0. This measure replaces FN in Equation 4.7, and the direct effects of sentiment on returns are examined. According to the results (not reported), this measure is no better than FN .

4.6 Summary and Conclusions

The interrelations between stock market returns, trading volumes and textual sentiment have been frequently studied in the existing literature. The literature draws attention to the statistically significant mutual effects between market-level textual sentiment (especially negative sentiment) and market returns. It is also generally agreed that market returns predict textual sentiment, and that sentiment forecasts market-level trading volumes. In this Chapter, six hypotheses regarding the relations between firm-specific sentiment, firm-level equity returns and trading volumes are examined. We employ panel data regressions, VAR models and rolling regressions on individual firm data in sequence. We use Loughran and McDonald (2011)'s 'finance negative' words to proxy for sentiment.

Overall, the results of panel regressions show that higher negative sentiment predicts lower next-day returns and higher returns predict lower next-day negative sentiment; higher returns predict lower next-day trading volumes and higher trading volumes predict higher next-day returns; and that higher negative sentiment causes lower trading volumes on the following day. The effects of trading volumes on next day's sentiment are not significant. To summarize, at the panel data level, we reject H_1 to H_6 except H_4 .

VAR models are then employed for individual firms over the full data period. The most evident relations are that trading volumes cause sentiment, and that returns cause trading volumes. Around 75 percent of the results indicate that both H_4 and H_6 could be rejected at the 5 percent level of significance. Half of the results show that sentiment causes trading volumes at the 10 percent level of significance. Therefore, we tend to reject H_3 at the individual firm level. Based on the other results, we could not reject H_1 , H_2 and H_5 . In addition, it is important to note that a total of six firms (30 percent) show that sentiment

causes stock returns either directly or indirectly (through trading volumes) over the full sample period.

By employing 1-year rolling-window regressions, we find that with the exception of Chevron, sentiment predicts stock returns directly at least for some data periods (9.56 percent on average), although there is no fixed pattern of where the predictability is concentrated. Similar effects are observed when returns in the VAR system are replaced by abnormal returns. Therefore, H_7 is rejected. By and large, the indirect effects of sentiment on return are negligible for most firms. On average, either of the two effects emerged 11.32 percent of time, although they did not often coexist.

Aggregating the results, we conclude that the interrelations among firm-specific sentiment, firm-level equity returns and trading volumes are most evident in the pooled sample. The results regarding the role of textual sentiment on equity returns and trading volumes are consistent with previous studies (e.g. Tetlock (2007), Tetlock et al. (2008), Garcia (2012)). Negative words are found to have an immediate negative impact on equity returns, while lower returns forecast more negative words; higher returns predict lower next-day trading volumes.

Our first novel finding is that there is strong evidence of the indirect effects of sentiment on equity returns: sentiment causes trading volumes, whilst trading volumes cause returns, which are most evident in the panel data. In addition, we have examined the time-series pattern of sentiment's role for individual stocks. The results strongly support that trading volumes cause sentiment, and *vice versa*. Trading volumes were not found to cause sentiment in the panel data, however, probably because in the panel regressions, only the first lag (rather than five lags) of trading volumes was used as the explanatory variable. Our second novel finding is that sentiment does not cause equity returns all the time but that it has varying impacts over time. The discrete periods where the predictability is concentrated are likely to be associated with important firm-specific news and events. This should be a promising direction for future research. Overall, our results suggest that firm-specific textual sentiment is a potentially important time-varying factor in equity pricing models.

Figure 4.1 : Searching for qualified news stories

Search terms

Terms and Connectors

Natural Language

headline(Boeing) and at15(Boeing)

Search

Topics You Added

and

THE BOEING COMPANY

☒ Relevance: Strong References only

Use connectors to show relation of terms, e.g., election **and** opinion poll, car **or** vehicle, advertising **w/2** campaign Search tips

Add Index Terms:

Company | Industry | Subject | Geography

Sources

*All English Language News

More sources

+ Show options to search specific document sections

Duplicate Options

Specify date

On - High similarity

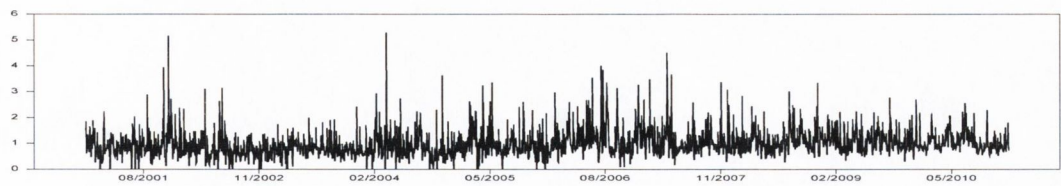
What's this?

Date is between... 1 Jan 2001 to 31 Dec 2010

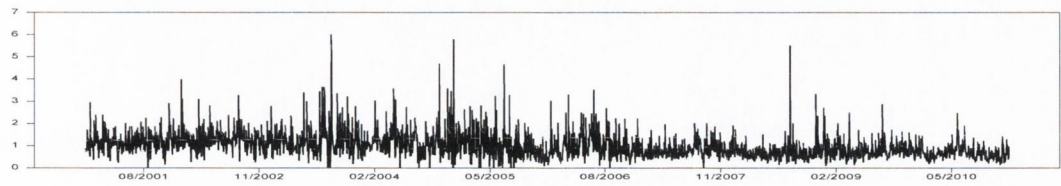
Search

Notes: Source is <http://www.lexisnexis.com.elib.tcd.ie/uk/nexis/indexing/RequestHandler.do>. This figure shows how the qualified firm-specific news are searched and retrieved from the *LexisNexis – News and Business* database. This figure shows the power search page of the database.

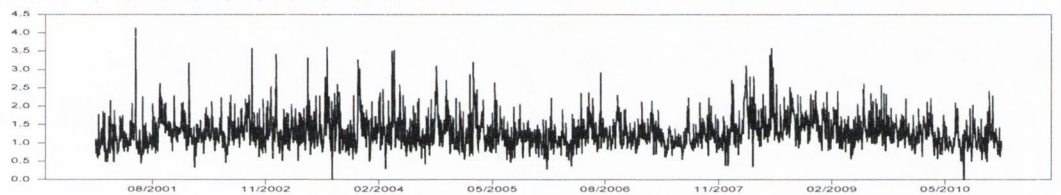
Figure 4.2: Plots of FN series



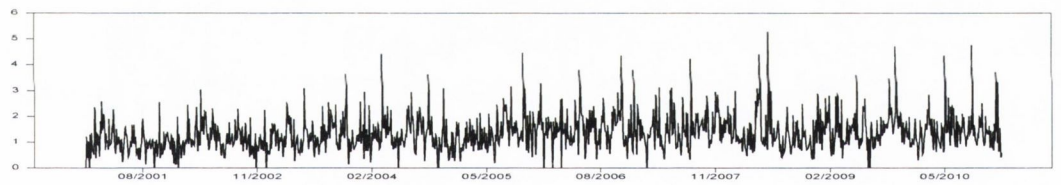
(a) Apple FN



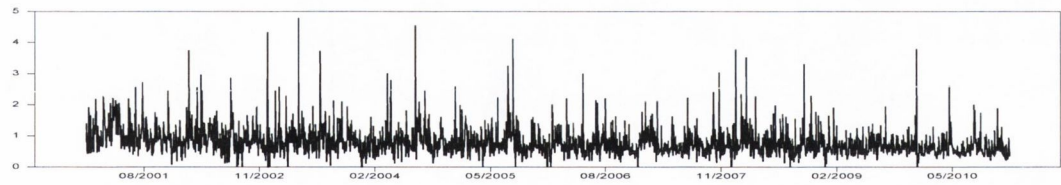
(b) AT&T FN



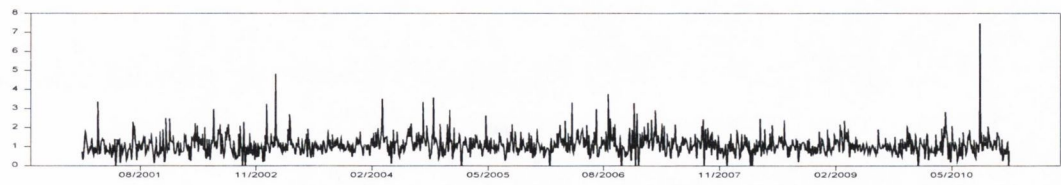
(c) Boeing FN



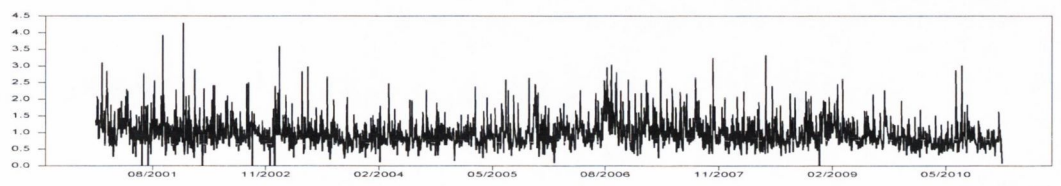
(d) Chevron FN



(e) Cisco FN

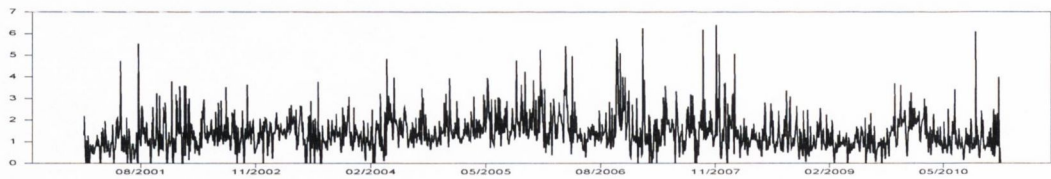


(f) ConocoPhillips FN

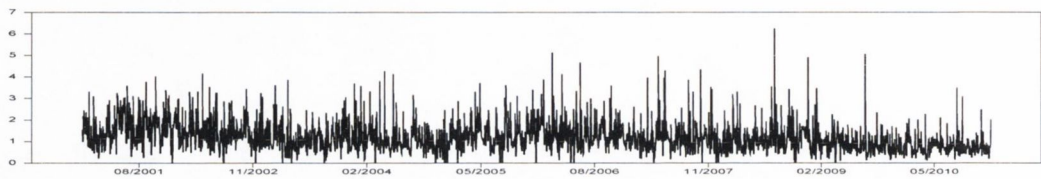


(g) Dell FN

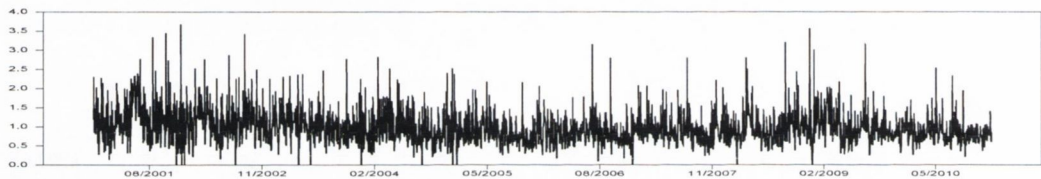
Figure 4.2 (Continued)



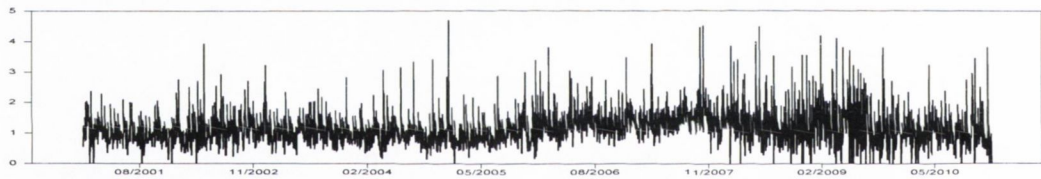
(h) ExxonMobil *FN*



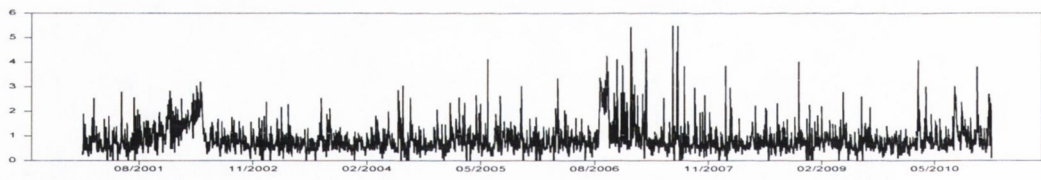
(i) Ford Motor *FN*



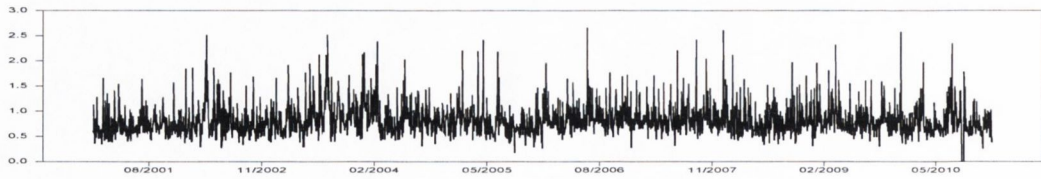
(j) General Electric *FN*



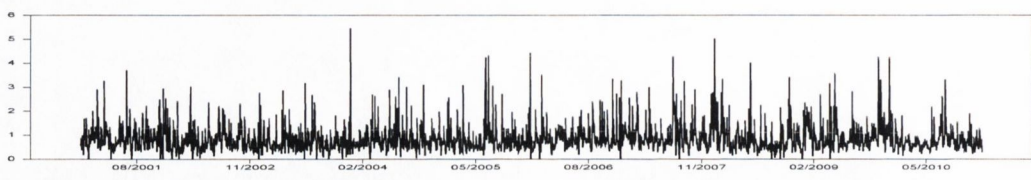
(k) Home Depot *FN*



(l) HP *FN*

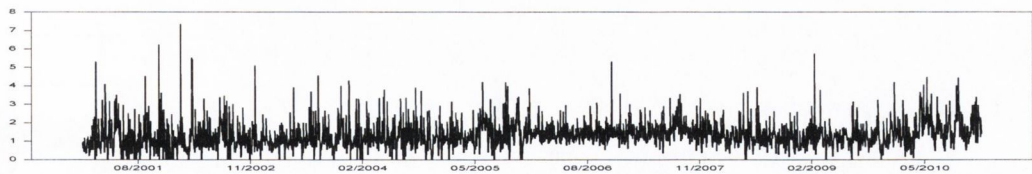


(m) IBM *FN*

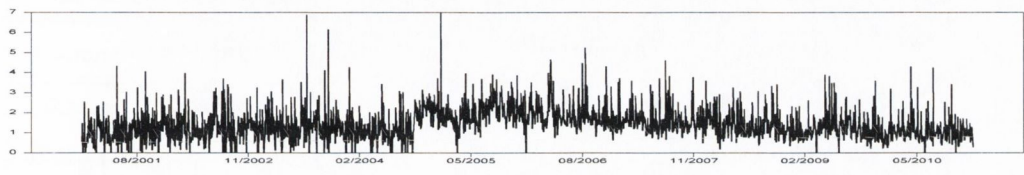


(n) Intel *FN*

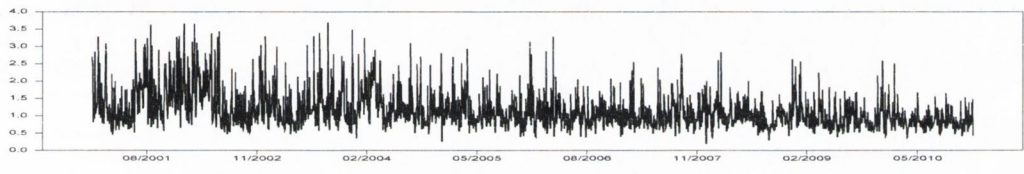
Figure 4.2 (Continued)



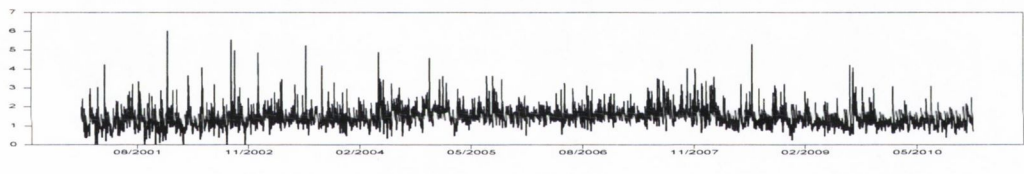
(o) Johnson & Johnson *FN*



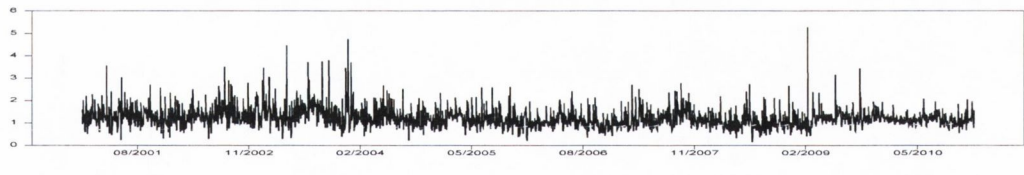
(p) Merck *FN*



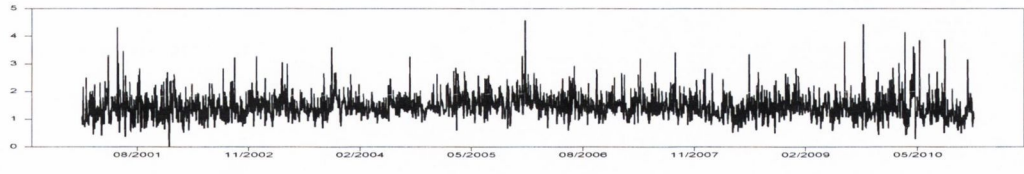
(q) Microsoft *FN*



(r) Pfizer *FN*



(s) Verizon *FN*



(t) Walmart *FN*

Figure 4.3: Equity Returns

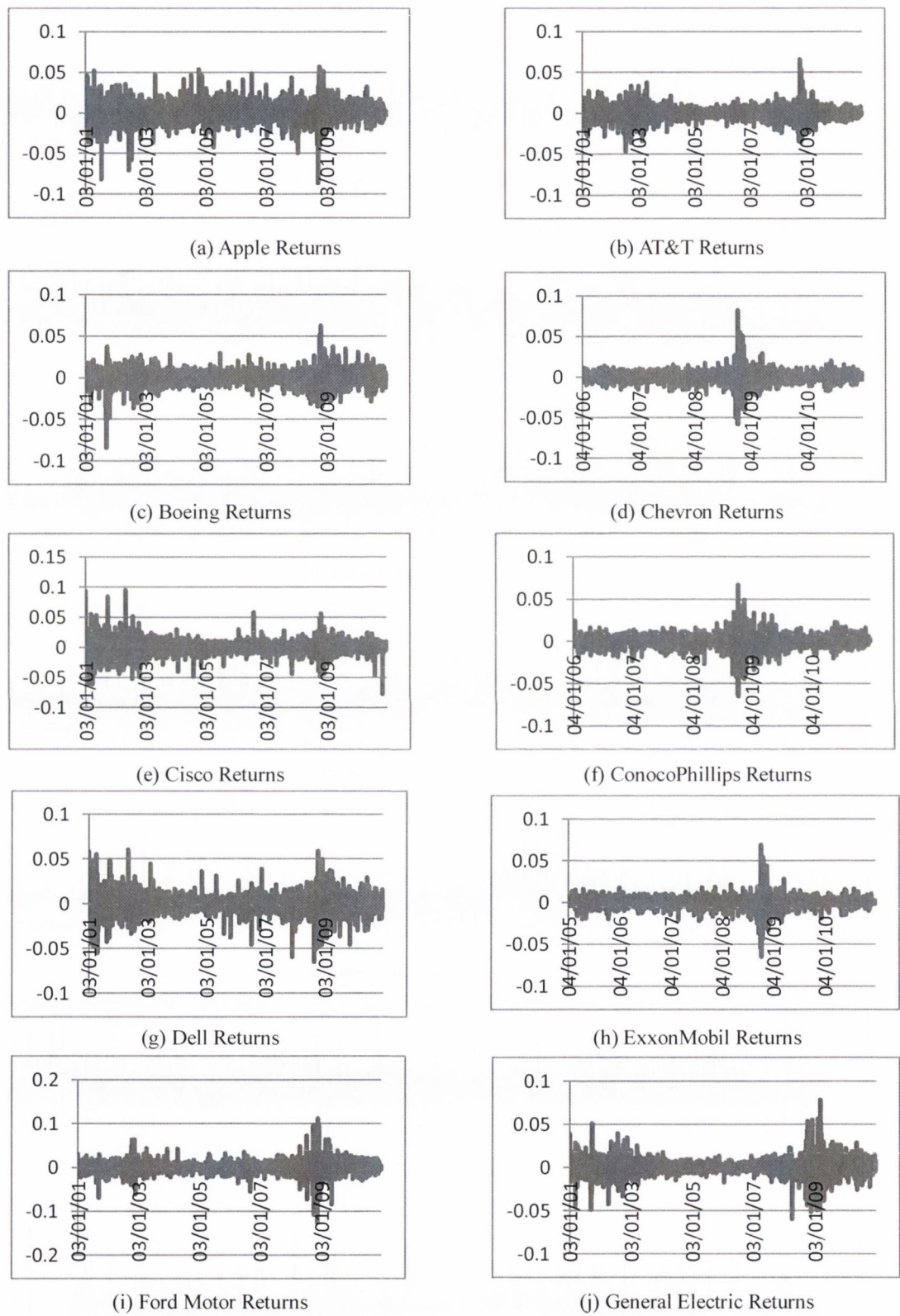
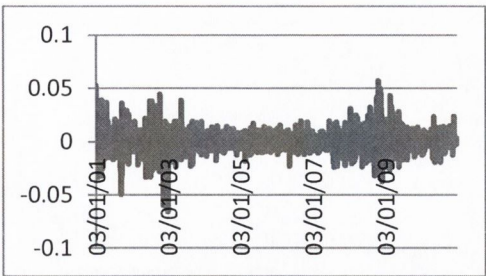
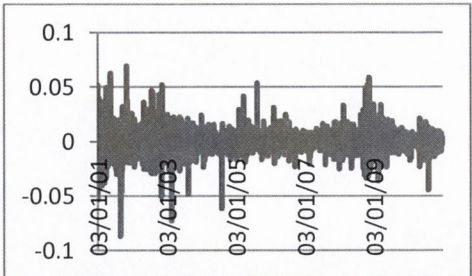


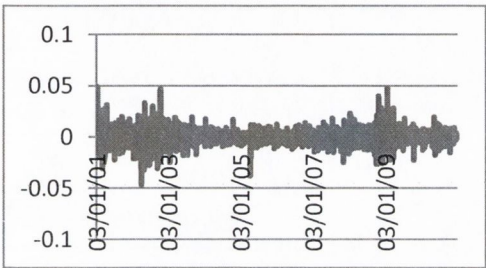
Figure 4.3: Equity Returns (continued)



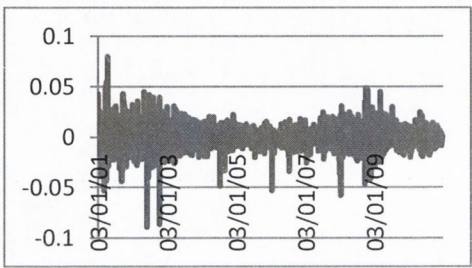
(k) Home Depot Returns



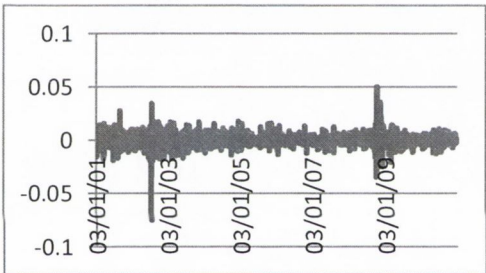
(l) HP Returns



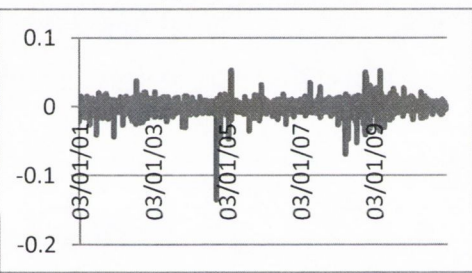
(m) IBM Returns



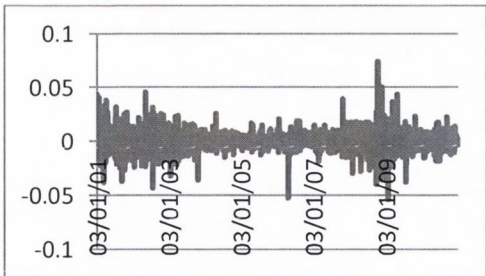
(n) Intel Returns



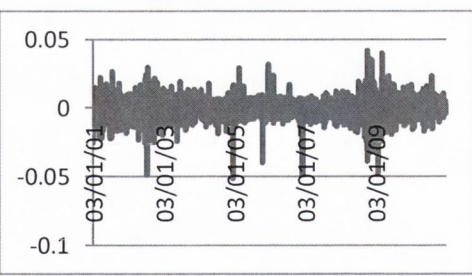
(o) Johnson & Johnson Returns



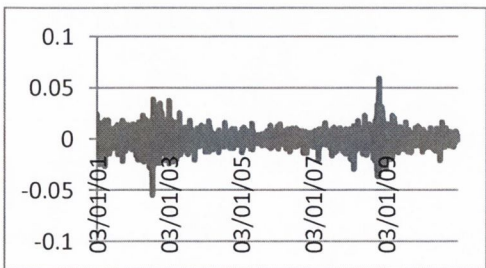
(p) Merck Returns



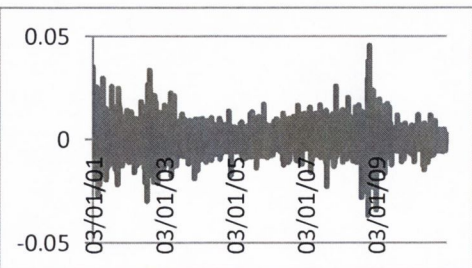
(q) Microsoft Returns



(r) Pfizer Returns

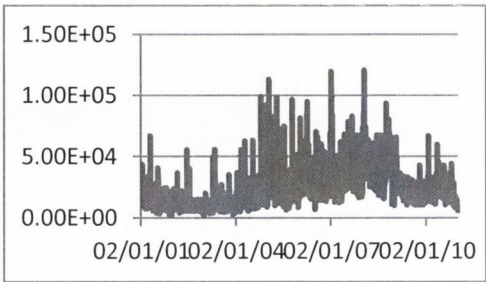


(s) Verizon Returns

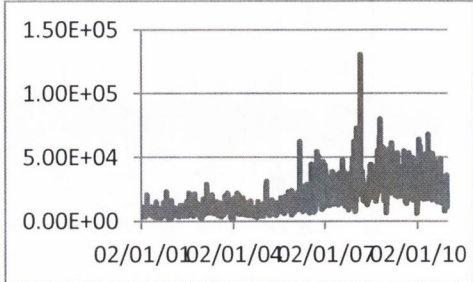


(t) Wal-mart Returns

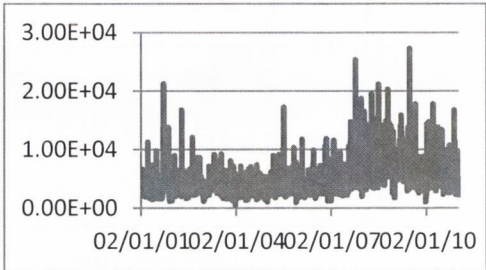
Figure 4.4: Trading Volumes



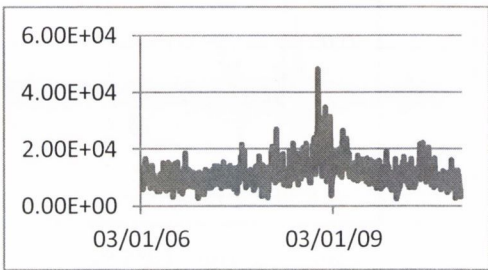
(a) Apple Volume



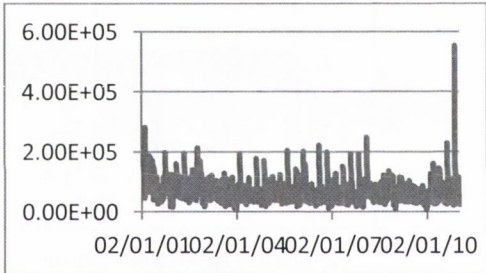
(b) AT&T Volume



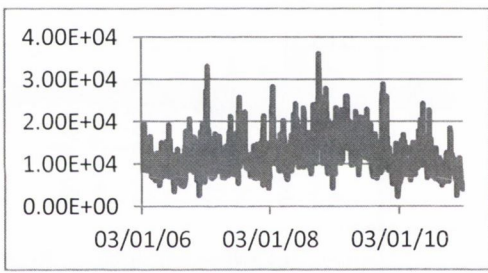
(c) Boeing Volume



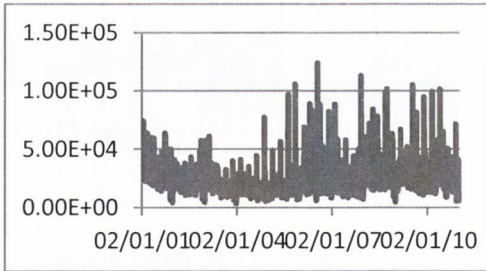
(d) Chevron Volume



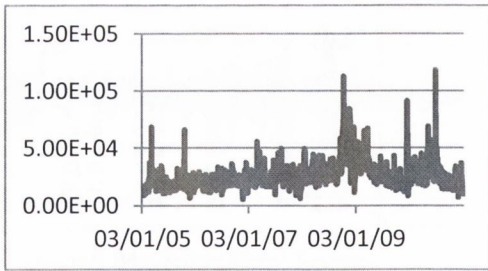
(e) Cisco Volume



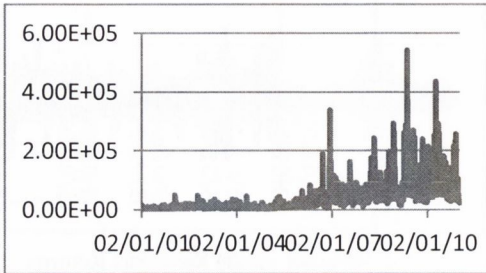
(f) ConocoPhillips Volume



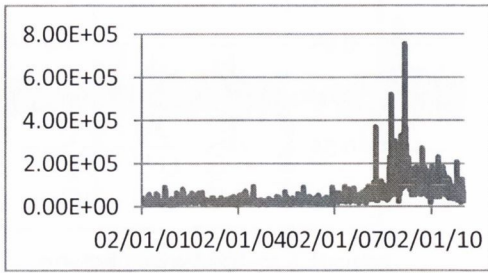
(g) Dell Volume



(h) ExxonMobil Volume

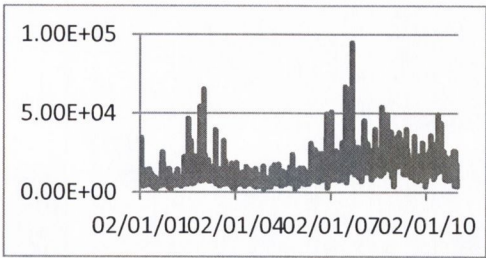


(i) Ford Motor Volume

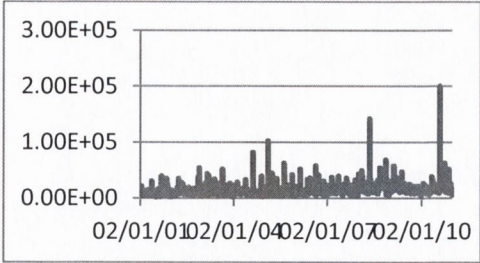


(j) General Electric Volume

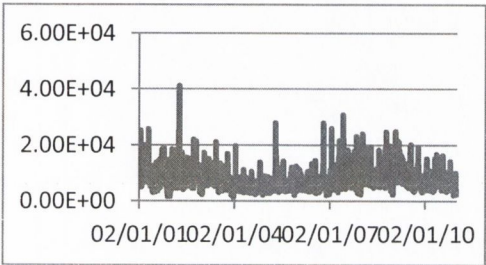
Figure 4.4: Trading Volumes (continued)



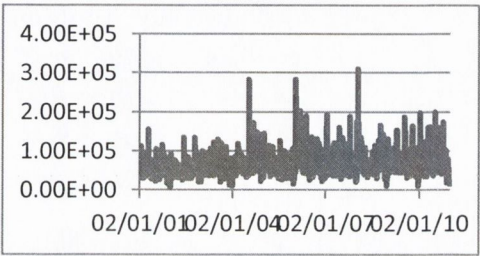
(k) Home Depot Volume



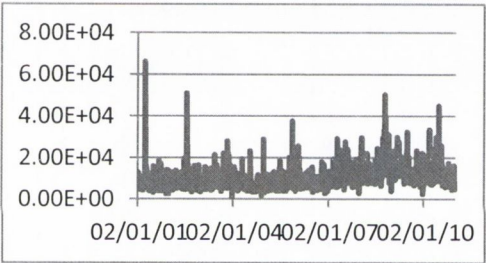
(l) HP Volume



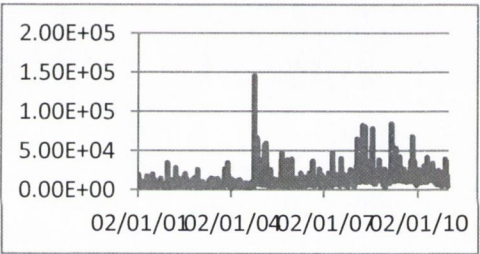
(m) IBM Volume



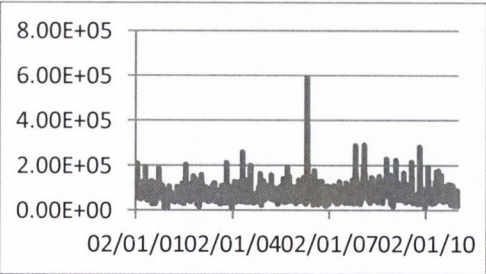
(n) Intel Volume



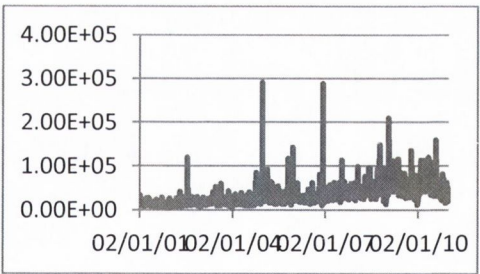
(o) Johnson & Johnson Volume



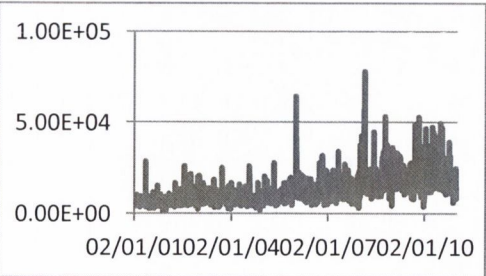
(p) Merck Volume



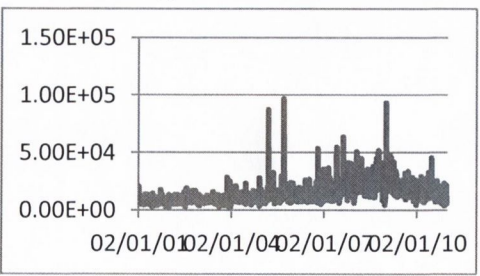
(q) Microsoft Returns



(r) Pfizer Volume

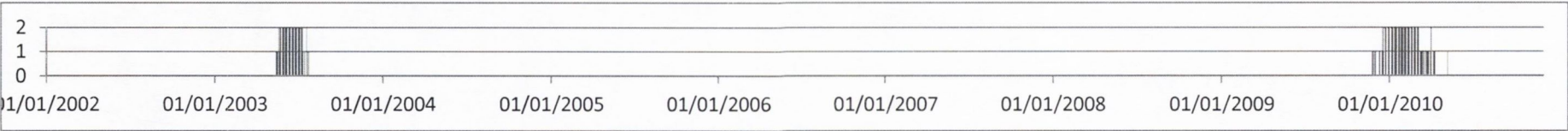


(s) Verizon Volume

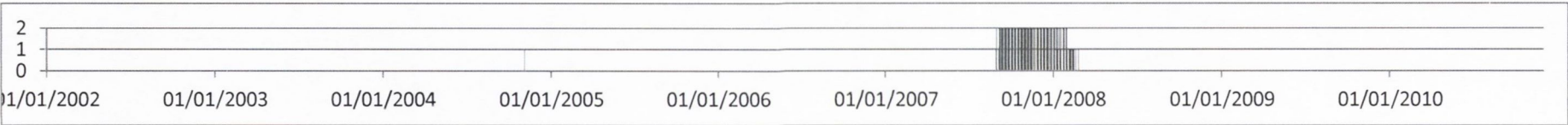


(t) Wal-mart Volume

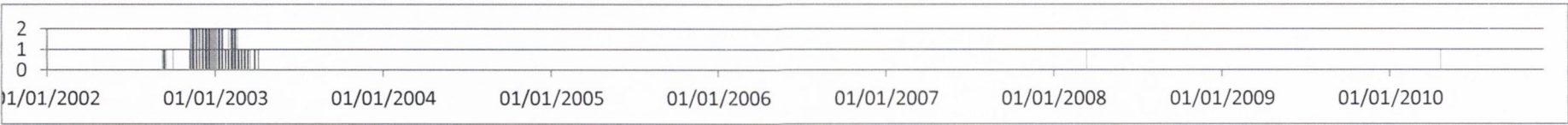
Figure 4.5: Predictability of sentiment on return (1-year sliding windows)



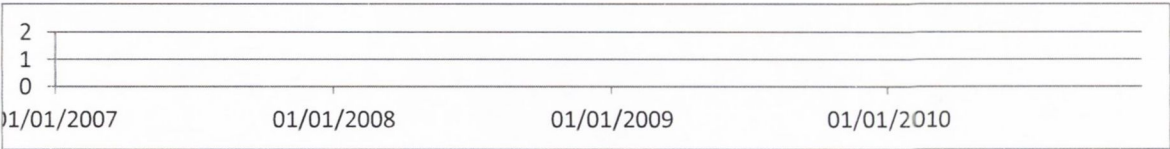
(a) Apple



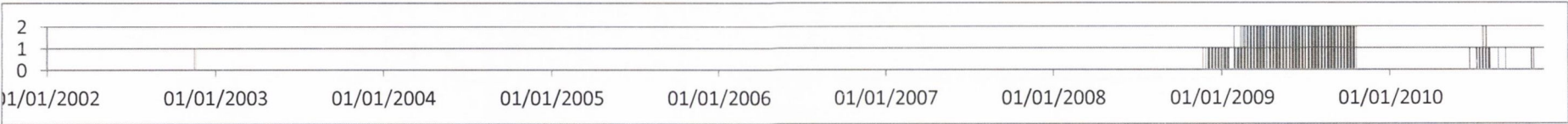
(b) AT&T



(c) Boeing

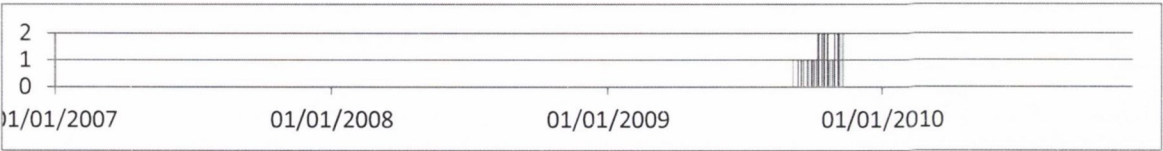


(d) Chevron

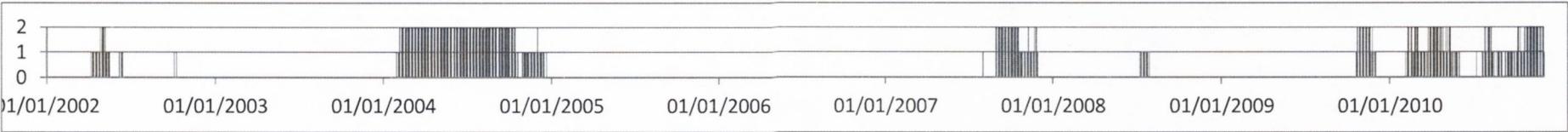


(e) Cisco

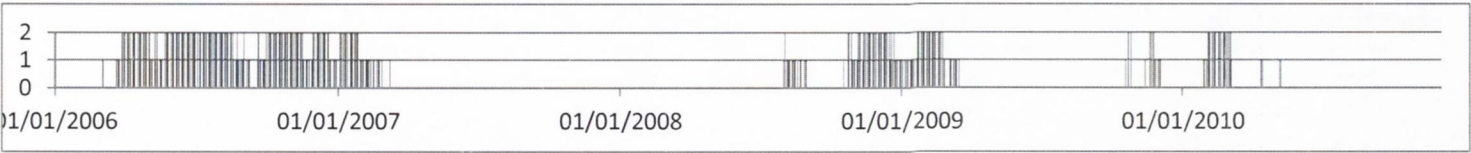
(f) Figure 4.5 (continued from previous page)



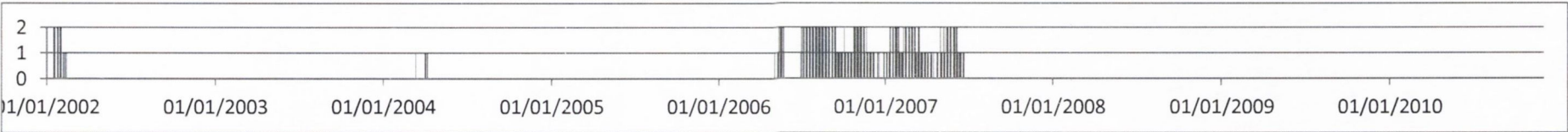
(g) ConocoPhillips



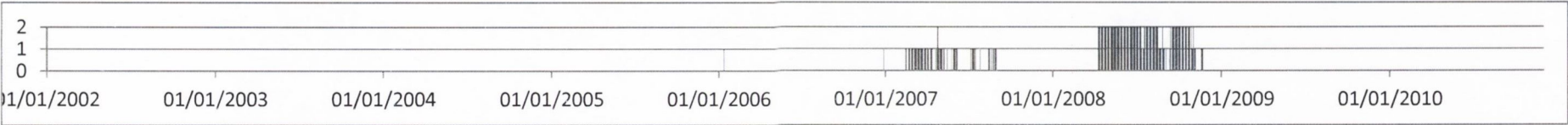
(h) Dell



(i) Exxon Mobil



(j) Ford Motor



(k) General Electric

Figure 4.5 (continued from previous page)

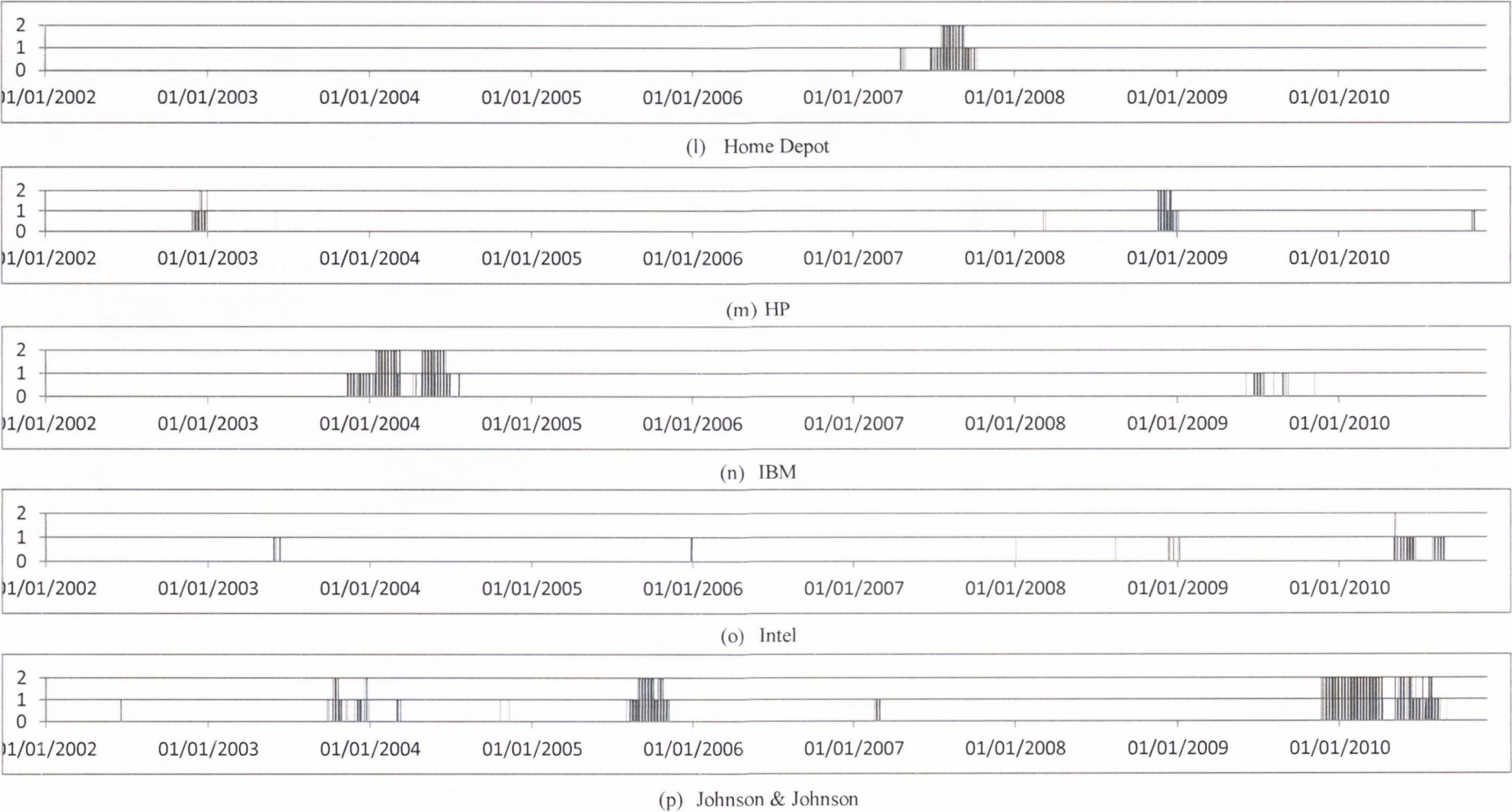


Figure 4.5 (continued from previous page)

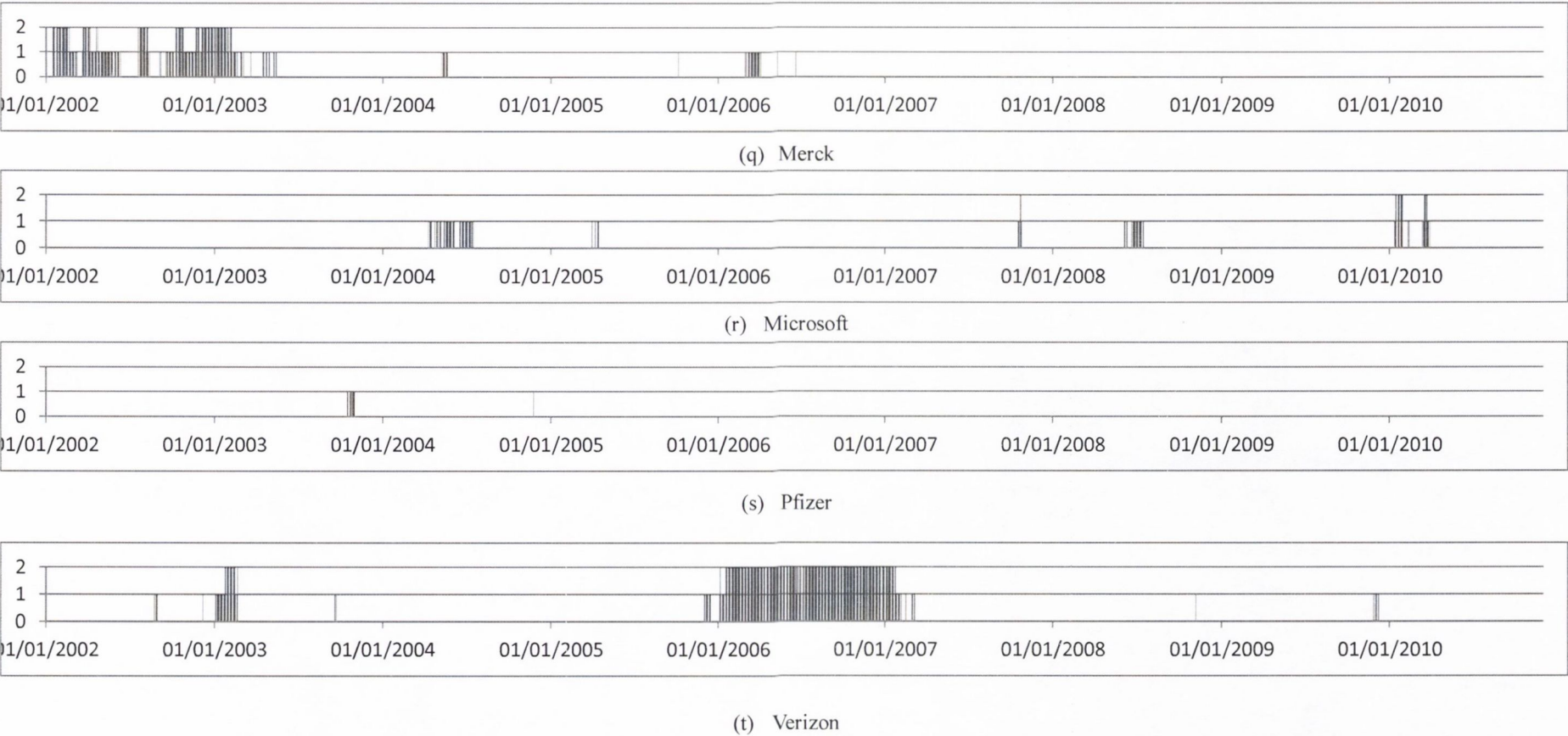
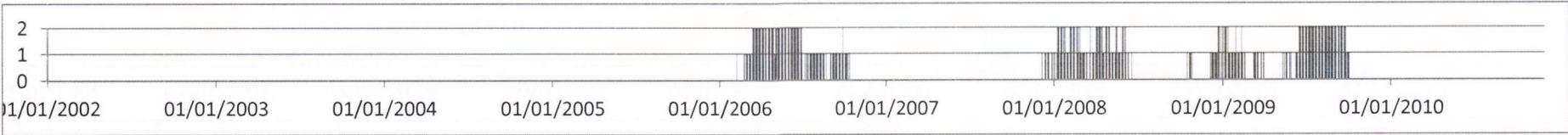


Figure 4.5 (continued from previous page)



(u) Wal-mart

Notes: Graph (a) to (t) display the results of rolling regressions for all firms, specifically whether the five lags of FN significantly forecast stock return. In each graph, the horizontal axis in each graph illustrates the date at which the individual 252-day sliding window ends. The vertical axis indicates whether the five lags of FN significantly forecast returns in the individual regression. A value of 1 on the vertical axis indicates significance at the 10 percent level, while a value of 2 indicates significance at the 5 percent level.

Table 4.1: Average number of articles and number of words per day (firm-specific news)

Company	Number of observations	Average Number of Articles/ day	Average Number of Words/day
Apple	2515	20.01	10714.35
AT&T	2515	16.29	9309.99
Boeing	2515	42.68	21030.80
Chevron	1259	13.64	6730.83
Cisco	2515	20.38	12368.96
ConocoPhillips	1259	10.37	6173.93
Dell	2515	22.19	11279.08
ExxonMobil	1511	10.04	5701.69
Ford Motor	2515	11.26	7079.89
General Electric	2515	19.81	9734.83
Home Depot	2515	11.00	6570.31
HP	2515	12.67	7791.52
IBM	2515	37.41	19678.86
Intel	2515	17.76	9694.44
Johnson & Johnson	2515	10.72	5515.98
Merck	2515	13.22	9035.28
Microsoft	2515	46.25	24590.28
Pfizer	2515	17.46	9948.85
Verizon	2515	28.05	16166.98
Walmart	2515	27.85	14352.54

Notes: This table presents the average number of articles and number of words per day for each firm.

Table 4.2: Descriptive Statistics of *FN* series

Company	Mean	Variance	Skewness	Kurtosis	J-B P-val	Q10 P-val	DF t-stat
Apple	1.05	0.30	1.70	6.03	0.00	0.00	-34.37
AT&T	1.04	0.37	1.98	8.60	0.00	0.00	-35.89
Boeing	1.35	0.22	1.11	2.52	0.00	0.00	-30.37
Chevron	1.40	0.48	1.16	2.93	0.00	0.00	-25.38
Cisco	0.85	0.23	2.22	9.65	0.00	0.00	-35.99
ConocoPhillips	1.13	0.29	2.39	18.93	0.00	0.00	-40.00
Dell	1.04	0.22	1.45	3.89	0.00	0.00	-35.91
ExxonMobil	1.49	0.77	1.47	4.20	0.00	0.00	-28.46
Ford Motor	1.27	0.53	1.30	3.03	0.00	0.00	-37.92
GE	1.02	0.22	1.19	2.87	0.00	0.00	-40.25
Home Depot	1.19	0.37	1.29	3.41	0.00	0.00	-42.03
HP	0.93	0.40	2.22	7.88	0.00	0.00	-31.77
IBM	0.85	0.10	1.49	3.96	0.00	0.00	-35.07
Intel	0.92	0.37	2.19	7.61	0.00	0.00	-33.91
Johnson & Johnson	1.41	0.59	1.29	4.21	0.00	0.00	-38.57
Merck	1.50	0.62	1.03	2.94	0.00	0.00	-35.24
Microsoft	1.21	0.32	1.41	2.12	0.00	0.00	-27.41
Pfizer	1.52	0.39	1.43	4.88	0.00	0.00	-39.25
Verizon	1.26	0.21	1.66	7.10	0.00	0.00	-37.58
Walmart	1.51	0.22	1.12	3.62	0.00	0.00	-40.27

Notes: This table summarizes the descriptive statistics of *FN* series, including mean, variance, skewness, excess kurtosis, the Jarque-Bera (J-B) test of normality, the Ljung-Box Q test for autocorrelation and the Dickey-Fuller unit-root test.

Table 4.3: Summary statistics of firm-level returns and trading volumes

Company	Obs	Variable	Skewness	Kurtosis	J-B P-val	Q10 P-val	DF t-stat
Apple	2515	Return	-0.13	4.10	0.00	0.07	-51.06
		Volume	1.74	4.52	0.00	0.00	-18.12
AT&T	2515	Return	0.20	5.33	0.00	0.00	-51.45
		Volume	1.44	4.28	0.00	0.00	-15.94
Boeing	2515	Return	-0.31	6.10	0.00	0.03	-49.95
		Volume	2.15	7.76	0.00	0.00	-23.40
Chevron	1259	Return	0.20	13.25	0.00	0.00	-40.88
		Volume	1.78	6.36	0.00	0.00	-14.11
Cisco	2515	Return	0.18	7.40	0.00	0.03	-52.21
		Volume	3.92	41.93	0.00	0.00	-26.97
ConocoPhillips	1259	Return	-0.38	6.66	0.00	0.00	-38.61
		Volume	1.09	2.09	0.00	0.00	-15.49
Dell	2515	Return	0.00	4.73	0.00	0.00	-51.28
		Volume	2.44	11.16	0.00	0.00	-27.46
ExxonMobil	1511	Return	0.07	12.72	0.00	0.00	-46.09
		Volume	2.30	9.12	0.00	0.00	-15.43
Ford Motor	2515	Return	0.01	12.33	0.00	0.00	-48.42
		Volume	3.02	16.00	0.00	0.00	-16.84
GE	2515	Return	0.08	8.13	0.00	0.01	-50.80
		Volume	4.52	38.03	0.00	0.00	-14.97
Home Depot	2515	Return	0.13	5.02	0.00	0.08	-50.31
		Volume	2.28	10.61	0.00	0.00	-19.44
HP	2515	Return	0.01	7.69	0.00	0.04	-51.18
		Volume	6.62	100.35	0.00	0.00	-27.55
IBM	2515	Return	0.31	6.13	0.00	0.08	-51.14
		Volume	2.11	8.78	0.00	0.00	-23.85
Intel	2515	Return	-0.22	6.09	0.00	0.05	-52.81
		Volume	2.49	14.04	0.00	0.00	-28.14
Johnson & Johnson	2515	Return	-0.63	19.87	0.00	0.00	-52.27
		Volume	2.50	13.01	0.00	0.00	-22.15
Merck	2515	Return	-1.92	30.46	0.00	0.49	-50.11
		Volume	4.20	37.80	0.00	0.00	-23.38
Microsoft	2515	Return	0.21	6.50	0.00	0.00	-54.87
		Volume	3.80	42.50	0.00	0.00	-27.80
Pfizer	2515	Return	-0.33	5.43	0.00	0.00	-51.28
		Volume	2.58	17.32	0.00	0.00	-18.47
Verizon	2515	Return	0.13	5.56	0.00	0.00	-52.51
		Volume	1.83	6.40	0.00	0.00	-19.40
Walmart	2515	Return	0.26	4.25	0.00	0.00	-54.37
		Volume	2.64	15.78	0.00	0.00	-22.08

Notes: This table provides descriptive statistics of daily stock returns and volume, including skewness, excess kurtosis, the Jarque-Bera (J-B) test of normality, the Ljung-Box Q test for autocorrelation and the Dickey-Fuller unit-root test.

Table 4.4: Equity data and summary statistics

Company	Size (billion)	Revenue (million)	Return (%)	Volume (thousand)	Avg <i>FN</i>
Pfizer	152.3	67,809	-5.1	33,254	1.52
Walmart	197.7	421,849	1.8	13,566	1.51
Merck	108.0	45,987	-0.6	11,546	1.50
ExxonMobil	376.9	354,674	10.4	25,541	1.49
Johnson & Johnson	176.4	61,587	4.9	10,046	1.41
Chevron	211.0	196,337	14.3	11,387	1.40
Boeing	49.9	64,306	7.8	4,845	1.35
Ford Motor	41.3	128,954	28.4	36,995	1.27
Verizon	104.5	106,565	3.5	12,288	1.26
Microsoft	222.9	62,484	9.4	68,595	1.21
Home Depot	59.3	67,997	3.9	12,885	1.19
ConocoPhillips	95.0	184,966	9.8	12,065	1.13
Apple	357.3	65,225	68.3	22,572	1.05
AT&T	172.2	124,629	2.1	17,552	1.04
Dell	28.2	61,494	3.5	23,699	1.04
General Electric	169.8	151,628	-1.3	44,335	1.02
HP	54.6	126,033	8.9	14,027	0.93
Intel	127.5	43,623	6.4	62,793	0.92
Cisco	101.6	40,040	3.2	60,549	0.85
IBM	220.0	99,870	10.7	7,461	0.85

Notes: This table presents the firms' market capitalization on 15/11/2011, annual revenue (provided by the 2011 Fortune 500 list), average annual return (simple return), average daily trading volume, and average daily *FN* scores of each firm. The firms are sorted by the average *FN* scores, from high to low, for rough examination of the relationship between negative sentiment and returns.

Table 4.5: Panel unit root tests

	LLC	Breitung	HT
FN	-120.00***	-82.15***	0.32***
Turnover	-65.48***	-64.07***	0.74***
SUE	1.24	-5.64***	0.99***
B/M	-8.59***	-16.92***	0.95***
MV	-10.12***	-2.78***	0.96***

Notes: This table reports the statistics and the significance level (1 percent ***) of panel unit root tests on each variables. Levin–Lin–Chu (LLC) test, Breitung test, and Harris–Tzavalis

Table 4.6: Panel estimation: do sentiment and volumes predict returns?
(Sentiment measure: First lag of *FN*)

	Dependent Variable: Returns (<i>R</i>)			
	(1)	(2)	(3)	(4)
<i>FN</i> (-1)	-3.567E-04 (-2.38**)	-3.101E-04 (-2.01**)	-3.567E-04 (-2.12**)	-3.101E-04 (-1.75*)
<i>R</i> (-1)	-0.013 (-1.10)		-0.013 (-1.81*)	
<i>R</i> (-2)	-0.031 (-1.86*)		-0.031 (-2.34**)	
<i>AR</i> (-1)		-0.017 (-1.26)		-0.017 (-4.34***)
<i>AR</i> (-2)		0.010 (0.84)		0.010 (2.46***)
<i>V</i> (-1)	0.040 (1.64)	0.042 (1.75*)	0.040 (2.17**)	0.042 (2.50***)
<i>SUE</i>	4.380E-05 (0.35)	4.360E-05 (0.36)	4.380E-05 (0.93)	4.360E-05 (0.96)
<i>B/M</i>	-0.004 (-3.32***)	-0.003 (-3.45***)	-0.004 (-12.79***)	-0.003 (-14.95***)
<i>MV</i>	1.360E-05 (0.42)	1.280E-05 (0.40)	1.360E-05 (0.54)	1.280E-05 (0.53)
<i>VIX</i> (-1)	1.270E-05 (6.74***)	1.350E-05 (7.42***)	1.270E-05 (12.54***)	1.350E-05 (15.44***)
<i>Monday</i>	5.718E-04 (0.76)	5.186E-04 (0.68)	5.718E-04 (2.54***)	5.186E-04 (2.34**)
<i>January</i>	-6.613E-04 (-0.79)	-5.991E-04 (-0.74)	-6.613E-04 (-2.68***)	-5.991E-04 (-2.43**)
<i>Constant</i>	5.690E-05 (0.11)	6.810E-06 (0.01)	5.690E-05 (0.25)	6.810E-06 (0.03)
Estimation method	Pooled OLS	Pooled OLS	Random effects	Random effects
R-squared	0.01	0.01	0.01	0.01
Observations	42704	42704	42704	42704

Notes: This table reports the estimated coefficients and the corresponding t-statistics (in brackets) for each of the regressions in the four columns. The symbols ***, ** and * indicate t-stats significance at 1 percent, 5 percent and 10 percent level, respectively. The dependent variable is returns, and the first lag of *FN* (*FN*(-1)) is the sentiment measure. Regressions (1) and (2) are estimated by pooled OLS with standard errors clustered by calendar quarter. Regression (3) and (4) are estimated using random-effects model with Huber–White robust standard errors. Besides *FN*(-1), other independent variables in column (1) and (3) include the first and second lag of returns (*R*), the first lag of trading volumes (*V*), standardized unexpected earnings (*SUE*) (use its value at the end of the preceding quarter), book-to-market ratio (*B/M*), market value (*MV*), past volatility proxied by the first lag of VIX index, dummies for the Monday and January effects. In columns (2) and (4), the lags of return are replaced by lags of abnormal return (*AR*), which is the return adjusted for Fama-French (1993) three factors.

Table 4.7: Panel estimation: do sentiment and volumes predict returns?
(Sentiment measure: Change in *FN*)

	Dependent Variable: Returns (R)			
	(1)	(2)	(3)	(4)
<i>D(FN)</i>	-6.465E-04 (-3.57***)	-6.573E-04 (-3.57***)	-6.465E-04 (-2.89***)	-6.573E-04 (-2.90***)
<i>R(-1)</i>	-0.013 (-1.09)		-0.013 (-1.78*)	
<i>R(-2)</i>	-0.030 (-1.81*)		-0.030 (-2.26**)	
<i>AR(-1)</i>		-0.017 (-1.25)		-0.017 (-4.27***)
<i>AR(-2)</i>		0.010 (0.82)		0.010 (2.42**)
<i>V(-1)</i>	0.041 (1.68*)	0.043 (1.79*)	0.041 (2.21**)	0.043 (2.53***)
<i>SUE</i>	4.870E-05 (0.39)	4.780E-05 (0.40)	4.870E-05 (0.99)	4.780E-05 (1.01)
<i>B/M</i>	-0.004 (-3.33***)	-0.003 (-3.45***)	-0.004 (-12.78***)	-0.003 (-14.95***)
<i>MV</i>	1.340E-05 (0.41)	1.260E-05 (0.40)	1.340E-05 (0.53)	1.260E-05 (0.52)
<i>VIX(-1)</i>	1.270E-05 (6.73***)	1.350E-05 (7.38***)	1.27E-05 (12.53***)	1.35E-05 (15.48***)
<i>Monday</i>	4.550E-04 (0.60)	4.059E-04 (0.53)	4.550E-04 (2.08**)	4.060E-04 (1.90*)
<i>January</i>	-6.716E-04 (-0.80)	-6.096E-04 (-0.76)	-6.716E-04 (-2.70***)	-6.096E-04 (-2.45**)
<i>Constant</i>	-3.443E-04 (-0.88)	-3.396E-04 (-0.82)	-3.443E-04 (-1.79*)	-3.396E-04 (-1.92*)
Estimation method	Pooled OLS	Pooled OLS	Random effects	Random effects
R-squared	0.01	0.01	0.01	0.01
Observations	42704	42704	42704	42704

Notes: This table reports the estimated coefficients and the corresponding t-statistics (in brackets) for each of the regressions in the four columns. The symbols ***, ** and * indicate t-stats significance at 1 percent, 5 percent and 10 percent level, respectively. The dependent variable is returns, and the change in *FN* (*D(FN)*) is the sentiment measure. Regressions (1) and (2) are estimated by pooled OLS with standard errors clustered by calendar quarter. Regression (3) and (4) are estimated using random-effects model with Huber–White robust standard errors. Besides *D(FN)*, other independent variables in column (1) and (3) include the first and second lag of returns (*R*), the first lag of trading volumes (*V*), standardized unexpected earnings (*SUE*) (use its value at the end of the preceding quarter), book-to-market ratio (*B/M*), market value (*MV*), past volatility proxied by the first lag of VIX index, dummies for the Monday and January effects. In columns (2) and (4), the lags of return are replaced by lags of abnormal return (*AR*), which is the return adjusted for Fama-French (1993) three factors.

Table 4.8: Panel estimation: do returns and volumes cause sentiment?

Dependent Variable: Sentiment (<i>FN</i>)		
	(1)	(2)
<i>FN</i> (-1)	0.344	0.284
	34.59***	13.67***
<i>FN</i> (-2)	0.157	0.099
	17.88***	8.92***
<i>R</i> (-1)	-0.611	-0.641
	-4.04***	-4.10***
<i>V</i> (-1)	-0.567	2.618
	-1.06	1.27
<i>SUE</i>	-0.008	-1.248E-02
	-1.92*	-2.13**
<i>B/M</i>	0.008	0.009
	3.41***	4.94***
<i>MV</i>	6.620E-05	-1.407E-04
	0.34	-0.53
<i>VIX</i> (-1)	1.340E-05	9.250E-06
	0.83	0.59
<i>Monday</i>	-0.050	-0.043
	-5.42***	-2.95***
<i>January</i>	0.002	0.003
	0.18	0.23
<i>Constant</i>	0.598	0.712
	32.98***	24.56***
Estimation method	Pooled OLS	Fixed effects
R-squared	0.19	0.11
Observations	42721	42721

Notes: This table reports the estimated coefficients and the corresponding t-statistics (in brackets) for each of the regressions in the two columns. The symbols ***, ** and * indicate t-stats significance at 1 percent, 5 percent and 10 percent level, respectively. The dependent variable is the sentiment measure, *FN*. Regressions (1) is estimated by pooled OLS with standard errors clustered by calendar quarter. Regression (2) is estimated using fixed-effects model with Huber–White robust standard errors. Independent variables include the first two lags of *FN*, the first lag of returns (*R*), the first lag of trading volumes (*V*), standardized unexpected earnings (*SUE*) (use its value at the end of the preceding quarter), book-to-market ratio (*B/M*), market value (*MV*), past volatility proxied by the first lag of VIX index, dummies for the Monday and January effects.

Table 4.9: Panel estimation: do sentiment and returns cause volumes?

Dependent Variable: Volumes (<i>V</i>)		
	(1)	(2)
<i>V</i> (-1)	0.634 23.74***	0.603 35.30***
<i>V</i> (-2)	0.210 6.94***	0.179 44.48***
<i>R</i> (-1)	-0.008 -2.00**	-0.008 -3.03***
<i>FN</i> (-1)	-2.462E-04 -6.60***	-1.333E-04 -1.80*
<i>SUE</i>	2.210E-05 1.22	7.640E-06 0.22
<i>B/M</i>	2.454E-04 1.01	2.524E-04 16.87***
<i>MV</i>	5.870E-06 1.92*	6.110E-06 1.90*
<i>VIX</i> (-1)	2.810E-07 1.38	4.220E-07 1.62
<i>Monday</i>	-0.001 -7.40***	-0.001 -5.92***
<i>January</i>	2.657E-04 1.55	2.883E-04 5.67***
<i>Constant</i>	0.002 11.24***	0.002 12.38***
Estimation method	Pooled OLS	Fixed effects
R-squared	0.67	0.56
Observations	42721	42721

Notes: This table reports the estimated coefficients and the corresponding t-statistics (in brackets) for each of the regressions in the two columns. The symbols ***, ** and * indicate t-stats significance at 1 percent, 5 percent and 10 percent level, respectively. The dependent variable is trading volumes (*V*). Regressions (1) is estimated by pooled OLS with standard errors clustered by calendar quarter. Regression (2) is estimated using fixed-effects model with Huber–White robust standard errors. Independent variables include the first two lags of *V*, the first lag of returns (*R*), the first lag of sentiment measure(*FN*), standardized unexpected earnings (*SUE*) (use its value at the end of the preceding quarter), book-to-market ratio (*B/M*), market value (*MV*), past volatility proxied by the first lag of VIX index, dummies for the Monday and January effects.

Table 4.10: VAR modeling over full sample period

Company	Usable Obs	Return			Volume			Fin Neg		
		Return	Volume	Fin Neg	Return	Volume	Fin Neg	Return	Volume	Fin Neg
Apple	2509	2.85 (0.01)	1.32 (0.25)	0.67 (0.65)	3.62 (0.00)	797.40 (0.00)	2.69 (0.02)	1.45 (0.17)	3.85 (0.00)	79.35 (0.00)
AT&T	2509	2.71 (0.02)	0.35 (0.88)	1.02 (0.40)	5.12 (0.00)	894.65 (0.00)	5.80 (0.00)	0.22 (0.95)	14.73 (0.00)	44.01 (0.00)
Boeing	2509	2.66 (0.02)	1.08 (0.37)	1.72 (0.13)	8.50 (0.00)	366.84 (0.00)	2.65 (0.02)	0.62 (0.69)	7.48 (0.00)	141.60 (0.00)
Cisco	2509	2.42 (0.03)	0.74 (0.59)	0.37 (0.87)	2.13 (0.06)	230.40 (0.00)	1.07 (0.37)	1.22 (0.30)	18.66 (0.00)	47.95 (0.00)
Dell	2509	4.25 (0.00)	1.07 (0.37)	2.55 (0.03)	0.72 (0.61)	213.67 (0.00)	1.12 (0.35)	2.82 (0.02)	6.29 (0.00)	56.17 (0.00)
Ford Motor	2509	5.64 (0.00)	2.36 (0.04)	1.10 (0.36)	5.47 (0.00)	884.16 (0.00)	4.99 (0.00)	1.45 (0.20)	8.06 (0.00)	44.55 (0.00)
GE	2509	1.21 (0.30)	5.69 (0.00)	1.94 (0.08)	15.35 (0.00)	1354.63 (0.00)	2.13 (0.06)	3.16 (0.01)	4.42 (0.00)	37.81 (0.00)
Home Depot	2509	2.32 (0.06)	1.79 (0.08)	0.44 (0.82)	3.82 (0.00)	581.86 (0.00)	2.49 (0.03)	1.51 (0.18)	13.36 (0.00)	12.40 (0.00)
HP	2509	1.04 (0.39)	0.87 (0.50)	1.56 (0.17)	1.63 (0.15)	221.59 (0.00)	0.57 (0.73)	0.45 (0.81)	4.62 (0.00)	142.83 (0.00)
IBM	2509	1.50 (0.19)	2.85 (0.01)	0.52 (0.76)	5.47 (0.00)	368.84 (0.00)	0.53 (0.76)	0.71 (0.61)	4.69 (0.00)	69.62 (0.00)
Intel	2509	3.36 (0.01)	0.74 (0.60)	0.45 (0.82)	0.82 (0.54)	207.85 (0.00)	3.46 (0.00)	0.35 (0.88)	1.48 (0.19)	79.89 (0.00)
Johnson & Johnson	2509	5.08 (0.00)	1.48 (0.19)	1.09 (0.31)	6.96 (0.00)	478.26 (0.00)	3.07 (0.01)	1.02 (0.41)	2.44 (0.03)	46.22 (0.00)
Merck	2509	0.62 (0.68)	0.61 (0.69)	0.51 (0.77)	2.39 (0.04)	382.85 (0.00)	0.75 (0.59)	2.53 (0.03)	1.34 (0.24)	105.25 (0.00)

Table 4.10 (continued from previous page)

Company	Usable Obs	Return			Volume			Fin Neg		
		Return	Volume	Fin Neg	Return	Volume	Fin Neg	Return	Volume	Fin Neg
Microsoft	2509	6.74 (0.00)	0.88 (0.50)	1.24 (0.29)	3.16 (0.01)	207.52 (0.00)	3.05 (0.01)	1.05 (0.39)	1.11 (0.35)	236.44 (0.00)
Pfizer	2509	6.49 (0.00)	1.55 (0.17)	0.62 (0.68)	1.26 (0.28)	810.47 (0.00)	0.33 (0.89)	2.07 (0.06)	0.18 (0.97)	48.29 (0.00)
Verizon	2509	3.50 (0.00)	0.34 (0.89)	2.11 (0.06)	3.29 (0.01)	653.52 (0.00)	1.92 (0.09)	1.66 (0.14)	5.14 (0.00)	46.05 (0.00)
Walmart	2509	5.99 (0.00)	0.92 (0.47)	0.98 (0.43)	2.22 (0.05)	538.44 (0.00)	1.44 (0.21)	1.02 (0.41)	2.31 (0.04)	28.62 (0.00)
Exxon Mobil	1505	16.90 (0.00)	1.07 (0.37)	2.06 (0.07)	10.76 (0.00)	377.14 (0.00)	1.59 (0.16)	0.38 (0.86)	3.52 (0.00)	35.24 (0.00)
Chevron	1253	9.72 (0.00)	1.78 (0.11)	0.57 (0.72)	11.54 (0.00)	291.91 (0.00)	0.45 (0.81)	1.10 (0.36)	1.95 (0.08)	30.95 (0.00)
ConocoPhillips	1253	5.18 (0.00)	1.33 (0.24)	1.08 (0.37)	14.26 (0.00)	217.67 (0.00)	0.39 (0.85)	0.75 (0.59)	3.31 (0.01)	6.75 (0.00)

Notes: This table presents the results of estimating VAR(5) model (VAR(4) for Chevron) over the full sample period, focusing on the F-statistics (with their p-values in brackets). Except the own lags of the dependent variable in each equation, p-values for the other two variables that are significant at the 5 percent level are marked in bold and italic, and 10 percent significance levels are marked in bold only. The column headed 'Usable obs' provides the number of usable observations for respective companies. The three columns headed 'Return' 'Volume' and 'Fin Neg' denote the dependent variable in the vector of endogenous variables, and the three columns beneath these headings are the explanatory variables in these equations. The F-statistics with their marginal significance levels indicate the joint significance of all lags of each explanatory variable in each equation.

Table 4.11: Summary of causality tests

	P-values ≤ 0.05	$0.05 < \text{P-values} \leq 0.10$
Sentiment causes returns	1	3
Returns cause sentiment	3	1
Sentiment causes volumes	8	2
Volumes cause sentiment	15	1
Volumes cause returns	3	1
Returns cause volumes	16	1

Notes: This table summarizes the results of testing H_1 to H_6 , employing VAR models on individual firm data (whole period). The column headed ‘P-values ≤ 0.05 ’ presents the number of significant cases at the 5 percent level, and the column headed ‘ $0.05 < \text{P-values} \leq 0.10$ ’ presents the number of marginally significant cases.

Table 4.12: The direct and indirect effects of *FN* on returns

Apple	None
AT&T	None
Boeing	None
Chevron	None
Cisco	None
ConocoPhillips	None
Dell	$S \rightarrow R$
Exxon Mobil	$S \rightarrow R$
Ford Motor	$S \rightarrow V \rightarrow R$
GE	$S \rightarrow R, S \rightarrow V \rightarrow R$
Home Depot	$S \rightarrow V \rightarrow R$
HP	None
IBM	None
Intel	None
Johnson & Johnson	None
Merck	None
Microsoft	None
Pfizer	None
Verizon	$S \rightarrow R$
Walmart	None

Notes: This table summarizes the direct and indirect effects of *FN* on returns, at the 10 percent level of significance.

Table 4.13: The direct and indirect effects of firm-specific sentiment on firm-level equity returns (rolling regressions)

(a)								
Company		Direct (%)	Indirect (%)	Either (%)	Total number of regressions			
Apple		5.80	0.71	6.46	1255			
AT&T		5.36	0.00	5.36	1255			
Boeing		4.69	0.00	4.69	1255			
Chevron		0.00	0.00	0.00	751			
Cisco		11.47	0.84	12.31	1255			
ConocoPhillips		4.19	0.00	4.19	751			
Dell		24.52	0.04	24.57	1255			
ExxonMobil		31.08	1.99	31.24	1003			
Ford Motor		11.91	0.58	12.48	1255			
GE		10.36	3.94	12.57	1255			
Home Depot		3.45	4.96	5.93	1255			
HP		3.32	7.35	10.67	1255			
IBM		6.91	6.86	13.77	1255			
Intel		3.19	3.05	6.24	1255			
Johnson & Johnson		12.88	4.65	17.31	1255			
Merck		12.04	1.90	13.94	1255			
Microsoft		5.00	0.44	5.36	1255			
Pfizer		0.89	0.13	1.02	1255			
Verizon		15.10	4.47	19.08	1255			
Walmart		19.12	0.18	19.21	1255			

(b)								
	Mean	Standard Deviation	Minimum	First Quartile	Median	Third Quartile	Maximum	T-stats (mean=0)
Direct	9.56	8.05	0.00	3.82	6.35	12.46	31.08	5.31
Indirect	2.10	2.44	0.00	0.09	0.77	4.21	7.35	3.86
Either	11.32	8.03	0.00	5.36	11.49	15.36	31.24	6.30

Notes: Panel (a) displays the proportion of significant cases among all the 252-day rolling-window regressions. It shows at the 10 percent significance level, the percentage of the time that the five lags of sentiment measure (*FN*) have direct, indirect or either effect on stock returns. The ‘indirect effect’ indicates that sentiment cause volumes, whilst volumes cause returns. Panel (b) displays the descriptive statistics.

Table 4.14: Robustness analysis: the direct and indirect effects of firm-specific sentiment on firm-level equity returns (rolling regressions)

(a)

Company	Direct (%)	Indirect (%)	Either (%)	Total number of regressions
Apple	7.30	13.55	20.81	1255
AT&T	15.18	0.00	15.18	1255
Boeing	1.95	0.13	2.08	1255
Chevron	13.46	0.00	13.46	751
Cisco	9.47	0.00	9.47	1255
ConocoPhillips	13.56	0.00	13.56	751
Dell	11.55	0.80	11.69	1255
ExxonMobil	5.02	0.00	5.02	1003
Ford Motor	6.64	0.27	6.91	1255
GE	9.52	12.13	19.61	1255
Home Depot	6.37	1.28	7.26	1255
HP	7.75	1.73	9.47	1255
IBM	9.12	1.28	10.40	1255
Intel	4.12	0.00	4.12	1255
Johnson & Johnson	3.63	4.43	8.01	1255
Merck	10.71	2.39	13.10	1255
Microsoft	5.40	0.00	5.40	1255
Pfizer	9.07	0.00	9.07	1255
Verizon	9.83	0.00	9.83	1255
Walmart	13.37	4.03	13.46	1255

(b)

	Mean	Standard Deviation	Minimum	First Quartile	Median	Third Quartile	Maximum	T-stats (mean=0)
Direct	8.65	3.66	1.95	5.89	9.10	11.13	15.18	10.57
Indirect	2.10	3.91	0.00	0.00	0.20	2.06	13.55	2.40
Either	10.40	4.85	2.08	7.08	9.65	13.46	20.81	9.58

Notes: Panel (a) displays the percentage of significant cases in all the 252-day rolling-window regressions. It shows at the 10 percent significance level, the percentage of the time that the five lags of sentiment measure (*FN*) have direct, indirect or either effect on stock abnormal returns. Abnormal returns are raw returns minus the returns of a value-weighted portfolio with similar size/book-to-market-characteristics. The ‘indirect effect’ indicates that sentiment cause volumes, whilst volume causes abnormal returns. Panel (b) displays the descriptive statistics.

Table 4.15: Robustness analysis: winsorizing returns, volumes and sentiment data

Company	Original Data (%)	Winsorized Data (%)
Apple	5.80	6.11
AT&T	5.36	5.49
Boeing	4.69	4.96
Chevron	0.00	0.20
Cisco	11.47	12.62
ConocoPhillips	4.19	2.79
Dell	24.52	23.42
ExxonMobil	31.08	33.55
Ford Motor	11.91	12.97
GE	10.36	7.70
Home Depot	3.45	3.59
HP	3.32	4.29
IBM	6.91	11.51
Intel	3.19	4.34
Johnson & Johnson	12.88	13.72
Merck	12.04	14.34
Microsoft	5.00	5.31
Pfizer	0.89	1.28
Verizon	15.10	15.23
Walmart	19.12	16.11
Average	9.56	9.98

Notes: This table compares the results of the modified model to the original model. The time series of return, volume and sentiment are all winsorized at the 1 percent level. Three exogenous variables (i.e. volatility measure proxied by VIX index and dummy variables for Monday effect and January effect) are still included in the model to control for volatility and potential return anomalies. Both columns present the percentage of significant cases in all the rolling-window regressions.

CHAPTER 5

RESULTS – CASE STUDY OF SOVEREIGN DEBT

5.1 Introduction

Since the introduction of the euro, many studies have investigated the determinants of sovereign bond yield spreads in the euro zone. Special attention has been paid to the period of the global financial crisis beginning from late 2008, since when the long-term sovereign bond yield spreads relative to Germany have increased greatly for most euro zone countries. Most recently from late 2009, the European sovereign debt crisis has drawn much attention worldwide. Countries such as Greece, Ireland, Portugal and Spain have experienced danger of a possible default, and their sovereign yield spreads widened dramatically as investors continually reassessed the default and other risks. During the debt crisis period, a large volume of news stories reported credit rating adjustments, analyzed the causes of the crisis, appropriate fiscal policies, the possibility of rescue packages, and investors' confidence.

News and information instantly affects the market and is reflected in the credit default swap (CDS) premium, liquidity of trading, general risk aversion, and the sovereign bond yield spread. In fact, credit risk, liquidity risk and general risk aversion are the three determinants of sovereign yield spreads identified by the extant literature. However, it is reasonable to argue that the content and complexity of news may not be completely

absorbed into the three factors. Although textual sentiment has been analyzed in equity markets, it has not been analyzed in debt markets. Textual sentiment is likely to be one facet of news that may contribute to yield spreads. An interesting research question then arises: is textual sentiment another determinant of sovereign yield spreads?

This chapter examines the role of sentiment during the ongoing European sovereign debt crisis from January 2009 to June 2012. Greece, Ireland, Portugal and Spain are the selected sample as they were the most affected countries and had the largest volume of news. Relevant news stories from each country are collected, and country-specific sentiment is extracted based on Loughran and McDonald's (2011) finance negative word list. Panel regressions are performed at daily frequencies to examine whether sentiment has significant impact on yield spreads in addition to the traditionally recognized variables.

The research in this chapter has implications for two areas of study. Firstly, it provides evidence of whether textual sentiment has additional explanatory power on sovereign yield spreads, and if it is true, the direction and magnitude of the effects. Meanwhile, the effects of sovereign yield spreads on textual sentiment are also examined. Secondly, it contributes to the literature on textual sentiment analysis by being the first to show whether the well-documented role of sentiment in equity markets also applies to bond markets. No assumption has been made in relation to investors' cognitive ability or emotional biases. In fact, textual sentiment may be a facet of news or information that has been overlooked in the literature but should actually be priced, and it does not necessarily reflect behavioral characteristics. Overall, this research provides evidence for asset pricing and suggests another angle for studying market efficiency.

The rest of the chapter is organized as follows. Section 5.2 presents a review of the literature on the determination of sovereign yield spreads in the Eurozone. The literature includes research on both the pre-crisis and crisis period. Section 5.3 describes the data in detail and provides preliminary analysis of the data. Hypotheses regarding the relationship between sentiment and yield spreads are also described. Section 5.4 reports the results of

testing the hypotheses through panel estimations and correspondent robustness analysis. Section 5.5 summarizes the main findings and draws together the conclusions.

5.2 Review of sovereign yield spread determinants literature

There is a relatively large body of literature exploring the determinants of sovereign bond yield spreads in the EU. The yield spreads studied are usually the country's 10-year, 5-year, or 2-year government bond yield differences relative to Germany. Researchers have identified three main factors that have a critical impact on the interest rates the governments pay relative to the benchmark government bond. First, the yield spreads are determined by the creditworthiness of the government. The 'credit risk' of a government has three types: default risk, credit spread risk and downgrade risk. Default risk is the probability that the bond issuer fails to make the required payments on the debt obligation. Credit spread risk is the risk of an increase in the yield spread on an asset against the risk-free benchmark. Downgrade risk is the possibility that the credit rating of an issuer is downgraded by a major credit-rating organization. It is anticipated that the higher the credit risk, the higher the yield spreads required by investors. The so-called 'sovereign credit risk' could be reflected by the country's fiscal conditions or certain macroeconomic variables, credit ratings or financial market information, such as the CDS premium paid for the debt. Liquidity risk is defined as the risk that a bond cannot be traded quickly enough in the market to prevent a loss or make the required profit. From time to time it has been found to have an important impact on yield spreads. The literature generally uses transaction costs (bid/ask spreads), trading volumes or the market value of bonds to proxy for liquidity risk. Sovereign bond spreads also reflect international risk aversion or general risk aversion, which is the risk premium investors would like to pay over the riskless asset. This factor can be considered to be a reflection of investor sentiment on risk. It is usually proxied by the yield spread between the US corporate and government bonds.

Codogno et al (2003) are one of the first researchers to study the determinants of sovereign bond yield spreads in the EU. Their sample includes 11 Eurozone countries, and the monthly data ranges from 1995 to 2002. They employ seemingly unrelated regressions analysis (SUR) of government bond spreads, and use public debt as the fraction of GDP as the proxy for a country's fiscal position. The results in relation to the role of indebtedness are mixed. No evidence is found for 10 Eurozone countries, with the exception of Austria, Spain and Italy. The spreads for these countries can be explained by the interaction effect of general risk aversion and the debt ratio. Codogno et al (2003) use three different liquidity measures: bid/ask spreads, trading volumes and the turnover ratio. Trading volumes are found to be the best performing liquidity indicator. They indicate that the role of liquidity is limited, since this effect is usually dominated by general risk aversion (with the exception of France). For some countries such as Belgium, France and Portugal, the effect of general risk aversion (global risk factor) seems to dominate. Meanwhile, global risk is found to have a greater impact in countries with high public debt level.

Geyer et al (2004) study five Eurozone countries, using weekly data from 1999 to 2002. A multi-issuer version of a pricing model is employed which was originally proposed by Duffie and Singleton (1999). They adopt a state-space approach to implement the model, whereby risk factors are extracted and then regressed on macroeconomic and financial variables. The authors find no significant effects from credit related macroeconomic variables (current account balance, industrial production growth and economic sentiment) or liquidity related variables. However, the role of global risk factors (corporate bond spreads and swap spreads) is found to be significant in determining the Euro government bond spreads.

For a pooled sample of 13 EU countries, Bernoth et al (2006) reveal that fiscal variables which are proxied by the budget balance or government debt have a significant effect on government bond yield spreads. In particular, since the start of the EMU in 1999, a measure of debt service (debt payments to current revenue) is found to be a better-performing indicator of fiscal imbalances than either the debt-to-GDP ratio or the deficit-to GDP ratio. Moreover, they find evidence that the effect of debt service on

spreads is non-linear. The authors find that, while liquidity (measured as the ratio of a country's bonds outstanding to total euro area bonds outstanding) has historically played a role in explaining spreads, this effect is no longer apparent following EMU. The international risk factor (proxied as the US corporate bond spread) is also found to play a significant role.

The role of sovereign credit risk and international risk aversion for either the pre-crisis or crisis period (after Sept 2008) has been confirmed by several studies. Studies that find significant effects of credit risk include Heppke-Falk and Huefner (2004), Gomez-Puig (2006), Faini (2006), Beber et al (2009), Schwarz (2009), and Barbosa and Costa (2010). Particularly, Heppke-Falk and Huefner (2004) provide evidence which suggests that following the EMU, expected budget deficits may have a more important role in explaining the risk premiums incorporated in government bond yield spreads. Many papers, such as Sgherri and Zoli (2009), Attinasi et al (2009), Barrios et al (2009), Haugh et al (2009), Manganelli and Wolswijk (2009), and Gerlach et al (2010) conclude that international risk aversion is the main driver of sovereign bond yield spreads.

In relation to the role of liquidity, the results are more complicated and controversial. Similar to Codogno et al (2003), Geyer et al (2004), and Bernoth et al (2006), some later studies, such as Gomez-Puig (2006), Jankowitsch et al (2006), Favero et al (2008), Beber et al (2009), Schwarz (2009), Ejsing and Sihvonen (2009), Attinasi et al (2009), Barrios et al (2009), Haugh et al (2009) and Gerlach et al (2010) find that liquidity effects are quantitatively limited and only relevant for some countries. Pagano and Von Thadden (2004), Jankowitsch et al (2006), and Gomez-Puig (2006) find that liquidity matters and is related to the bond market's size. Gomez-Puig (2006) indicates that after the introduction of the euro, liquidity is a highly relevant factor in determining sovereign yield spreads. Moreover, Beber et al (2009) conclude that credit risks often explain the largest part of bond spreads. In the case of heightened market uncertainty, however, liquidity considerations gain in importance.

Since the global economic and financial crisis in 2008, the most recent literature focuses on the impact of the crisis. Empirical findings from several studies (e.g. Barrios et al (2009), Ejlsing and Sihvonen (2009), Mody (2009) and Barbosa and Costa (2010)) suggest that the importance of credit risk and liquidity risk has increased in the crisis period. Moreover, the importance of credit risk has increased to a greater extent than liquidity. Barbosa and Costa (2010) conclude that prior to the collapse of Lehman Brothers, international risk aversion accounted, on average for approximately 70 percent of euro area sovereign bond yield spreads. Since September 2008, sovereign credit quality and liquidity have played a more important role in determining the yield spreads. Mody (2009), Sgherri and Zoli (2009), and Schuknecht et al (2010) also find that sovereign credit risk appears to be the major driver of the changes in sovereign yield spreads in the crisis period. Mody (2009) highlights that at the beginning of the financial market turmoil in the second half of 2007 and early 2008, spreads were largely determined by common risk factors. After March 2008, the different conditions of the countries' national financial sectors became determinants of the sovereign bond spreads. Furthermore, the effect of financial sector risk magnified in the period after the collapse of Lehman Brothers. Attinasi et al (2009) conclude from their results that the poor condition of the financial sector and the governments' announcement of bank rescue packages led to the transfer of credit risk from the financial sector to the public sector. This occurred when investors re-assessed the creditworthiness of the country yet the size of the rescue package does not statistically affect the yield spreads.

There are very few studies examining the relationship between sentiment and sovereign bond yield spreads. Only one study by Spyrou (2011) investigates whether investor sentiment is a determinant of sovereign bond yield spreads. Spyrou (2011) uses two sentiment measures, the University of Michigan Consumer Sentiment Index and the Greek Economic Sentiment Indicator, as the proxy for investor sentiment. Investor sentiment is tested to see if it has a significant impact on Greek yield spreads in the light of the sovereign debt crisis since 2009, controlling for fiscal conditions, liquidity, and general market conditions. Results reveal that local and international investor sentiment is indeed a

significant determinant for both the level and changes in the spread. However no studies have been found on the effects of textual sentiment on sovereign yield spreads and only a few relevant studies on the effect of credit-related announcements. Afonso et al (2011) employ an event study analysis to investigate the reaction of government yield spreads before and after the announcements of major credit rating agencies for EU countries. They find that changes in both the credit rating notations and the outlook (with some differences across rating agency) lead to a significant response of government rating bond yield spreads, and the effect is particularly important for the case of negative announcements.

5.3 Description and analysis of the data

The first part of this section describes the data, variable construction and the potential problems associated with them. It also presents a preliminary analysis of the data. The second part presents hypotheses regarding the relationship between textual sentiment and the sovereign yield spreads.

5.3.1 Data description and variable construction

The relevant variables include sentiment measures, government bond yield spreads, measures of credit risk, liquidity, and international risk aversion. The description of news data is displayed in the first place.

News

Compared with general economic and finance news, news that focuses on sovereign debt, credit risk and rescue actions should be more relevant to sovereign bond yield spreads, thus sentiment is to be extracted from this category of news. Because of the scarcity of such news in normal periods, the European sovereign-debt crisis period January 2009 to June 2012 is chosen as the sample period, during which there is much more qualified news. Greece, Ireland, Portugal and Spain are chosen as the countries being studied because they

have had the most difficult situations and have the greatest volume of news during this period.

As in Chapter 4, country-specific news stories are also downloaded from *LexisNexis*. For each country the search criteria is that the headline contains the country name, and there are at least five mentions of ‘sovereign’ or ‘debt’ in the body of the news. The sources chosen are ‘All English News’, which contains newspapers, magazines, newswires, etc. Index terms regarding ‘Public Finance’ subject as well as the option ‘Strong References Only’ are selected to make sure that the news stories retrieved are highly relevant to public debt and government solvency. I select ‘Duplication options – On – High similarity’ to get rid of the articles highly similar to a story that was published earlier.

Table 5.1 presents the number of news stories collected for each country, grouped by year. Ireland had the largest volume of news in 2010, when there was a marked increase in Irish 2-year bond yields in April and at the year end when the EU-IMF rescue deal was agreed. Greece had an increasing amount of news year by year, as its condition worsened. Both Spain and Portugal had increasing news as well, but they had only less than 100 news stories in 2009, when there were fewer problematic issues reported. Spain saw the greatest amount of news in the first half of 2012, as it gradually became the main concern for the Euro-zone and a bailout package was reported in June to be under consideration.

Sentiment Measures

After news stories were downloaded, the same method as described in Section 4.2.1 was used to split articles, group them by date and generate sentiment. For each country, the daily value of the basic sentiment score (BSC) is the fraction of negative words relative to the total words in a given day. We move all articles on weekends to Friday and omit the very few articles on holidays. There are many missing values because of no news on those days. Therefore BSC is not good for time series modeling at daily frequency because there would be too many losses of data points. Tetlock (2007) finds that textual sentiment has an immediate significant impact on the next day’s stock returns, and there is a large degree of return reversal between the next 2-5 days. Given this result, we assume that examining the

effect of sentiment at lower frequency (e.g. weekly, monthly) would be meaningless because after a few days the impact of sentiment may have already reversed and the immediate effects could not be detected. Therefore, two measures are constructed to minimize the impact of missing values and to maintain the effect of the most recent news. The first measure, denoted as $NSI_{[-2,0]}$, is the fraction of negative words relative to the total words in the latest three trading days (from day $t-2$ to day t). If there is no article in the three days, the value is set to be 0. The second measure, denoted as $NS2_{[-2,0]}$, is constructed through two steps. Firstly, let the value of BSC be 0 if there is no article on that day. Second, average the BSC value of the latest three days. The arbitrary choice of the period (i.e. from day $t-2$ to day t) to combine sentiment is to make sure that there are less days with no sentiment score and at the same time to maintain the effectiveness of news and information.

The plots of the two sentiment series for each country are presented in Figure 5.1. The descriptive statistics of the sentiment series are reported in Table 5.2. On average, Greece has the most negative words each day throughout the sample period, regardless of the measures. Portugal has the fewest negative words. The variance in sentiment is similar for Greece, Spain and Portugal, whilst Ireland has the smallest variance. The J-B column denotes the p-value of tests against the null hypothesis that the sentiment series is normally distributed. Results clearly show that none of the sentiment series is normally distributed with the exception of $NSI_{[-2,0]}$ for Spain. The Q10 column displays the p-value of tests for autocorrelation up to the 10th lag. All results indicate that every sentiment series is serially correlated.

The rest of this section presents the description of the dependent variable and independent variables other than sentiment. This data is obtained from Datastream, unless otherwise stated.

Yield Spread (Dependent Variable)

The measure of sovereign bond yield spread for each country is as follows:

$$YS_{i,t} = yield_{i,t} - yield_{BD,t}, \text{ where } i = \text{GR, IE, PT or ES} \quad (5.1)$$

Here, $YS_{i,t}$ is the yield spread. It is a country's 10+ - year sovereign bond index¹ ($yield_{i,t}$) minus the Germany (BD) counterpart ($yield_{BD,t}$). i denotes the sample countries: Greece (GR), Ireland (IE), Portugal (PT) and Spain (ES). It is a daily time series for each country. Figure 5.2 plots the yield spread for each country during the sample period. Ireland reached the highest point at 9.332 percent in July 2011, and then gradually lowered to around 4-5 percent in the first half of 2012. Greece began a sharp rise since 2010, and it reached an astonishing 27.975 percent in June 2012. Both Spain and Portugal began a steady rise from early 2010 to the highest point at 5.233 percent and 11.346 percent respectively.

Measures of Sovereign Credit

Sovereign credit is measured by two groups of variables. The first group contains only one variable, the CDS premium. CDS is a financial swap agreement in which the seller of the CDS compensates the buyer in the event of a debt default or other credit event. The buyer of the CDS makes a series of payments (the CDS 'premium') to the seller in exchange for a payoff if the debt defaults. The more likely the buyer is to default, the higher the premium that the seller asks for. Therefore, the CDS premium can be considered as a proxy for credit risk. The measure used in this research, denoted as *CDS*, is each country's sovereign 10-year bond CDS premium minus the Germany counterpart.

The second group of measures includes credit-related macroeconomic and fiscal variables that have appeared in the literature. These variables are usually observed at lower frequencies. To coordinate with other daily variables, only variables that are available in the *Economists Intelligence Unit*² at the highest frequency (quarterly) during the sample period are selected. These variables are the current account balance (percent of GDP), industrial production (percent change per annum) and the budget balance (percent of GDP),

¹ The indexes used are FTSE global government bond sub-indices, and the index value is the average yield of all bonds in the sub-index portfolio.

² www.eiu.com

denoted as *CAB*, *IND* and *BBL* respectively. The three quarterly variables are transformed to daily variables by assuming that the values are the same within each quarter, and they are also stated as the difference relative to Germany. Another frequently used variable in the literature, public debt (percent of GDP) or debt-to-GDP ratio, is only available at annual frequency, thus it is not appropriate for this research.

Measure of Liquidity

Similar to Sgherri and Zoli (2009), the proxy for liquidity (*LIQ*) used in this research is the market value of long-term sovereign bond. Specifically, it is the market value (in US dollars) of the Bank of American Merrill Lynch (BOFA ML) Government 10+Y Bond Index (sub-indices for each country), stated as the difference against Germany.

International Risk Aversion

In line with the literature, the common risk factor, i.e. international risk aversion, is measured by the difference between the US corporate bond yield and Treasury bond yield. In particular it is the difference between the yield of BOFA ML US High-Yield 1-10Y corporate bond index and the yield of BOFA ML US 1-10Y Treasury bond index, denoted as *RISK*.

5.3.2 Hypotheses

The hypotheses (null and the alternative) regarding the role of negative sentiment on sovereign yield spreads are straightforward :

H_1^0 : Negative textual sentiment does not significantly affect sovereign bond yield spreads.

H_1^a : Negative textual sentiment is a determinant of sovereign bond yield spreads.

The hypotheses regarding the reversal causal effect of sovereign yield spreads on negative sentiment:

H_2^0 : Sovereign bond yield spreads do not significantly affect negative textual sentiment.

H_2^a : Sovereign bond yield spreads significantly affect negative textual sentiment.

All the tests of these hypotheses are 2-tailed.

5.4 Results

This section presents the results of testing whether textual sentiment is a determinant of sovereign yield spreads, as well as whether sovereign yield spreads have explanatory power on textual sentiment.

Before running panel regressions, a variety of panel tests were obtained, in order to detect if the data is stationary, if heteroskedasticity, serial correlation and cross-sectional dependence exist in data. These tests, together with the panel estimations, are implemented by using the corresponding commands in Stata 12³.

Pre-regression Tests

Stationarity Test

Panel unit-root tests were taken on each data series (as panel data) to examine if they are stationary. Four tests were employed: the Levin-Lin-Chu (2002) (LLC) test, the Breitung (2000) test, the Harris-Tsavalis (1999) (HT) test and the Im-Pesaran-Shin (2003) test. As described in Chapter 3 Section 3.3.2, each test has its benefits and drawbacks, and no one particular test is dominant. Table 5.3 panel (a) reports the statistics and the significance level (1 percent ***, 5 percent **, 10 percent *). *YS*, *CDS* and *LIQ* are found to be non-stationary for any of the tests. Both sentiment measures ($NS1_{[-2,0]}$ and $NS2_{[-2,0]}$) are stationary at the 1 percent level of significance. Results are mixed for international risk aversion (*RISK*) and the three macroeconomic variables, *CAB*, *IND* and *BBL*. Given the results, *YS*, *CDS*, *LIQ* and *RISK* are then transformed to the first difference: *ys*, *cds*, *liq* and *risk*. Panel (b) of Table 5.3 reports the statistics of the four tests on *ys*, *cds*, *liq* and *risk*.

³ <http://www.stata.com>

They are all stationary at 1 percent level, and ready for panel estimation. The three macroeconomic variables, *CAB*, *IND* and *BBL* are left unchanged, because if taking first difference, the majority of their daily value would become 0, rendering these controlling variables useless.

Heteroskedasticity, Serial Correlation and Cross-Sectional Dependence

The heteroskedasticity test calculates a modified Wald statistics for groupwise heteroskedasticity in the residuals of a fixed effect regression model, following Greene (2000, p. 598).

The serial correlation test implements a test for serial correlation in the idiosyncratic errors of a linear panel-data model, as discussed by Wooldridge (2002). Under the null of no serial correlation, the residuals from the regression of the first-differenced variables should have an autocorrelation of -0.5. This implies that the coefficient on the lagged residuals in a regression of the lagged residuals on the current residuals should be -0.5.

The cross-sectional dependence test uses the methods shown in Pesaran (2004). Pesaran's statistic follows the standard normal distribution and it is able to handle balanced and unbalanced panels.

The three tests were obtained following the estimation of the proposed models. The results show that heteroskedasticity, serial correlation and cross-sectional dependence all exist. A sample of the results of each test is presented in Table A.3 in the Appendix. These results indicate that a standard error which is robust to heteroskedasticity, serial correlation and cross-sectional dependence should be used in the panel estimation.

5.4.1 Does sentiment determine yield spreads?

Panel estimations

Firstly, the Hausman's specification test, as described in Chapter 3 Section 3.3.2, is employed to test which model, fixed effects or random effects, should be used. The random effects model assumes that the individual effects are uncorrelated with the other regressors. The test was employed on several proposed models, and the majority of results show that the fixed-effect model is preferred (rejects the null that no correlation between individual effects and other regressors). A sample of the test results is presented in Table A.2 in the Appendix. Given the results, the fixed effect model is chosen to be used in the following estimation.

Two groups of regressions are formed to test if sentiment has significant impacts on yield spreads. The first group uses the CDS premium as a control for credit risk, and the second uses one of the three macroeconomic variables to proxy for credit risk. The general equation is written as follows:

$$ys_{i,t} = \alpha_{10} + \alpha_{11} \cdot NEG_{i,t} + \alpha_{12} \cdot CR_{i,t} + \alpha_{13} \cdot liq_{i,t} + \alpha_{14} \cdot risk_t + \mu_i + v_{i,t} \quad (5.2)$$

where $i = \text{GR, IE, PT or ES}$. The dependent variable is ys , the first difference of YS . α_{10} is the constant, NEG denotes the sentiment measure (NSI or $NS2$)⁴, and CR denotes one of the credit risk measures: cds , CAB , IND or BBL . The other variables have been described in Section 5.3.1. All equations are estimated using a panel data approach with fixed-effects, so μ_i denotes the unobservable individual-specific effects, and v_{it} denotes the remaining errors.

Table 5.4 reports the estimated coefficients and the t-statistics of the variables. Columns (1) to (8) each report the results of one model. The variables presented in the first column are independent variables. The exact sentiment measure (NSI or $NS2$) used is indicated at the end of the table. The variance and covariance matrix of the residuals is obtained on the

⁴ Short for $NSI_{[-2,0]}$ and $NS2_{[-2,0]}$

basis of the Driscoll and Kraay (1998) approach. This approach corrects heteroskedasticity, serial correlation and cross-sectional dependence in the residuals, making the estimation results more reliable. The respective results for the two sentiment measures may serve as an initial robustness check. The results show that in all regressions, negative sentiment has a significant or marginally significant positive impact on the change in the yield spread, after controlling for sovereign credit risk, liquidity, and international risk aversion, the three well-documented determinants of sovereign bond yield spread. The results are more robust for *NSI*, with 3 out of 4 cases significant at the 5 percent level. Only 1 out of 4 cases is significant at the 5 percent level, for *NS2*. This could indicate that *NSI* is a better measure to combine sentiment within the closest 3 trading days⁵. The magnitude of sentiment's impact is sizeable in comparison with the other variables. For a 10 percent increase in negative sentiment (i.e. the percentage of negative words), the change in the yield spreads rises by 3.8 to 7.1 basis points. Consistent with the literature, liquidity and international risk aversion are found to have strongly significant impacts on yield spreads, especially international risk aversion which has a large magnitude of impact. The results for credit risk are mixed. The CDS premium is marginally significant, but none of the three credit-related macroeconomic variables are significant. According to results not reported here, the R-square increases by 8 to 9 basis points if sentiment is included in the regressions compared to when it is excluded.

Robustness Analysis

Pooled OLS Estimation

Hausman's test was applied in order to decide whether a fixed-effects or random-effects model should be used. However, pooled OLS is another option for estimation. In this part, pooled OLS is applied on the same models to examine the robustness of results. Sentiments' effects are even stronger in this set of regressions. *NSI* is significant at the 1 percent level in 3 out of 4 cases, and *NS2* is significant at the 5 percent level in 3 cases (Table 5.5). Sentiment shows a marginally significant effect in the 2 other cases. The

⁵ Measures that combine sentiment within the latest 5 trading days were also employed in the regressions, but there were no significant results.

performances of credit risk measures, liquidity measures and international risk aversion are very similar to the former regressions.

Altering Model Specifications

In the previous regressions, most of the independent variables are defined as the difference against Germany, with the exception of international risk aversion (same for all countries) and sentiment measures. Initially consideration was given to collecting news stories of Germany in order to construct a ‘sentiment spread’ measure. However, Germany was often involved in the news during the European sovereign debt crisis concerning other countries’ situations, hence it is difficult to get a ‘pure’ Germany sentiment that reflects Germany’s sovereign credit. To mitigate the ‘mixed sentiment’ problem, news stories in which the headline contains the name of other countries should be excluded from the Germany text corpus, but this could not completely separate ‘Germany’s sentiment’ from other countries’ sentiment. In fact, although the ‘pureness’ of individual sentiment for countries deeply affected by the crisis such as Greece, Ireland, Portugal or Spain are expected to be higher, there could still be ‘mixed sentiment’ that mutually affects each country. As the crisis deepened, it has led to far-reaching consequences that extend beyond the borders of the ‘question countries’. Due to the growing interconnection of the global financial system and a general loss of investor confidence, ‘contagion’ exists both in and out of the region. For these reasons, it is not only impossible but it may also be unnecessary to entirely purify sentiment.

In this section, data and model specifications are revised in order to examine if the impact of sentiment is robust. The idea is to construct each variable (except international risk aversion) as the difference from its average value in each of the four countries. In this way, each variable including sentiment, becomes a relative measure against its average, and is no longer solely the difference against Germany. This makes the method of variable construction more consistent. For example, the average yield spread is computed as:

$$Yield_t^{AV} = \frac{1}{4} (Yield_t^{GR} + Yield_t^{IE} + Yield_t^{PT} + Yield_t^{ES}) \quad (5.3)$$

Then the new yield spread measure, denoted as *CSYS* (country-specific yield spread), is defined as:

$$CSYS_t^i = Yield_t^i - Yield_t^{AV} \quad i = GR, IE, PT \text{ or } ES \quad (5.4)$$

The average of the first sentiment measure is computed as:

$$NS1_t^{AV} = \frac{1}{4} (NS1_t^{GR} + NS1_t^{IE} + NS1_t^{PT} + NS1_t^{ES}) \quad (5.5)$$

The first new sentiment measure is defined as:

$$CSNS1_t^i = NS1_t^i - NS1_t^{AV} \quad i = GR, IE, PT \text{ or } ES \quad (5.6)$$

Using this method, the other variables, *NS2*, *CDS*, *LIQ*, *CAB*, *IND*, and *BBL* are transformed to *CSNS2*, *CSCDS*, *CSLIQ*, *CSCAB*, *CSIND*, and *CSBBL*, respectively. After taking panel unit root tests, *CSYS*, *CSCDS* and *CSLIQ* are transformed to the stationary first difference *csys*, *cscds* and *csliq*, respectively. Unlike the aforementioned variables, *risk* (international risk aversion) has the same values for all countries, thus it could not be transformed. Initial examination detects that *risk* is not a significant determinant of *csys*. Therefore in the following panel regressions *risk* is excluded from independent variables.

$$csys_{i,t} = \alpha_{20} + \alpha_{21} \cdot CSNS_{i,t} + \alpha_{22} \cdot CSCR_{i,t} + \alpha_{23} \cdot csliq_{i,t} + \mu_i + v_{i,t} \quad (5.7)$$

In Equation 5.7, *CSNS* is either *CSNS1* or *CSNS2*, and *CSCR* denotes *cscds*, *CSCAB*, *CSIND*, or *CSBBL*. The same estimation method and robust standard errors are used.

Overall, negative sentiment shows a more significant role on yield spread (Table 5.6). This time *CSNS2* is a better performer, with all four cases significant at the 1 percent level. *CSNS1* is significant at the 1 percent level once, 5 percent twice, and 10 percent once. Liquidity continues to be a strongly significant factor, but none of the credit risk factors are significant. Due to the unusual way in which yield spread measure and other variables are

constructed, these results should be only treated as complementary analysis and robustness check.

Based on the results, hypothesis H_1^0 should be rejected. As negative sentiment increases, the change in sovereign yield spreads increases. Negative sentiment is an additional determinant of sovereign yield spread besides credit risk, liquidity and international risk aversion.

5.4.2 Determinants of country-specific sentiment

On the other hand, do sovereign debt yield spreads have explanatory power on the country-specific negative sentiment? How about credit risk measures, liquidity measure and international risk aversion? The following model is considered:

$$NEG_{i,t} = \alpha_{30} + \alpha_{31} \cdot ys_{i,t} + \alpha_{32} \cdot CR_{i,t} + \alpha_{33} \cdot liq_{i,t} + \alpha_{34} \cdot risk_t + \varepsilon_t \quad (5.8)$$

As in Equation 5.2, NEG denotes either $NS1$ or $NS2$, and CR denotes one of the credit risk measures: cds , CAB , IND or BBL . The models are estimated using the pooled OLS method, and the variance and covariance matrix of the residuals is obtained on the basis of the Driscoll and Kraay (1998) approach.

The results show that the change in the yield spreads is positively and significantly associated with negative sentiment at the 1 percent or 5 percent level in 6 out of 8 cases, and the other 2 cases are marginally significant (Table 5.7). If the change in yield spreads increases 10 percent, negative sentiment would rise by 2.4 to 4.5 percent approximately. The credit risk variables, liquidity measure and international risk aversion are also shown to have impacts on negative sentiment at different significance levels. Among these variables, CDS premium, industrial production, the budget balance and international risk aversion have stronger effects. The direction of budget balance's effects seems counterintuitive (greater budget balance is associated with more negative words), which

may be due to the limitation of inconsistent data frequency in the model. It is interesting to note that greater industrial production is associated with lower negative sentiment, and that in the cases where industrial production (*IND*) is incorporated in the regression, the R-squared increases significantly. The current account balance does not show any significant effects.

Robustness analysis

As in the previous section, each variable (except the international risk aversion) is transformed to be a relative measure against its average. The model becomes:

$$CSNS_{i,t} = \alpha_{40} + \alpha_{41} \cdot csys_{i,t} + \alpha_{42} \cdot CSCR_{i,t} + \alpha_{43} \cdot csliq_{i,t} + \varepsilon_t \quad (5.9)$$

In Equation 5.9, *CSNS* is either *CSNS1* or *CSNS2*, and *CSCR* denotes *cscds*, *CSCAB*, *CSIND*, or *CSBBL*.

The measure of yield spreads (*csys*) has consistent and significant impacts on both sentiment measures at the 1 percent level in 7 out of 8 cases; the other case is significant at the 5 percent level (Table 5.8). Among the rest of the independent variables, only the CDS premium measure (*cscds*) and the budget balance measure (*CSBBL*) show significant explanatory power on negative sentiment.

Given the results, hypothesis H_2^0 is rejected. The other potential determinants of the country-specific negative sentiment include the CDS premium, international risk aversion and some macroeconomic or fiscal variables. This should be an interesting area for future research.

5.5 Summary and Conclusions

Although there is an increasing body of literature studying the role of textual sentiment in the stock market, little research has been conducted on fixed-income securities. Although there is a substantial body of studies investigating the determinants of sovereign yield spreads, the association between textual sentiment and yield spreads has yet to be linked. This is probably because textual sentiment analysis is still relatively new and not widely known. This chapter examines if the well-documented effect of textual sentiment on equity, also applies to sovereign debt market. In comparison to the literature exploring the determinants of sovereign yield spread, this research focuses on the examination of sentiment rather than testing different measures of credit risk, liquidity and general risk aversion.

Greece, Ireland, Portugal and Spain were selected as the sample countries. The European sovereign debt crisis period Jan 2009 to June 2012 was selected as the data range, due to the surge in news stories regarding a country's credit situations during this period. Sentiment was generated using the dictionary-based approach based on Loughran and McDonald (2011)'s finance negative word list, and two measures of sentiment were constructed. The results show that textual negative sentiment in the most recent three trading days has significant effects on sovereign yield spreads, after controlling for credit risk, liquidity and international risk aversion. For a 10 percent increase in the percentage of negative words, the change of sovereign yield spread would rise by 3.8 to 7.1 basis points. Consistent with the literature, this research finds that liquidity and general risk aversion each has a strongly significant effect on sovereign spreads. However, the performance of credit risk measures is not very good, possibly due to the inconsistent data frequencies compared with other variables.

Traditionally sovereign debt yield spreads are found to be determined by fundamentals such as fiscal conditions or market assessment of a variety of risk. Textual sentiment could represent the hard-to-quantify aspects of fundamentals and communicate additional pricing-relevant information, or it could reflect investor confidence and the market's assessment of the risk-return relationship. In addition, it could, reflect the behavioral or irrational aspects of the market, representing the hidden pricing factors that may not be

completely justified by the available facts. Results from this research confirm that textual sentiment indeed has additional explanatory power over traditional factors. The findings are consistent with the literature of textual sentiment analysis, which has studied equity market only.

On the other hand, sovereign bond yield spreads are found to be significantly directly related to negative textual sentiment. The increases of the change in yield spreads are associated with higher negative sentiment. The CDS premium, international risk aversion and some macroeconomic or fiscal variables also have potential effects on negative sentiment.

There are a few limitations in this research. *First*, the data period selected is during the European sovereign debt crisis period. It would be more difficult to examine the role of sentiment in non-crisis periods when there are much fewer relevant news stories, hence too many missing values of sentiment. The contemporaneous effect of sentiment could still be assessed, but time series analysis would be quite difficult. Therefore the evolution of sentiment's effect, e.g. whether it was permanently priced in or reversed, is difficult to assess. The question of whether the role of sentiment in the non-crisis periods is different from the crisis periods has not been revealed in this research. Future research could address this issue. *Second*, fiscal variables are usually low frequency data, making it hard to reconcile the stationarity problem with high-frequency sentiment measures. For example, some well-documented fiscal variables such as public debt and debt service ratio could not be employed in this research.

Figure 5.1: Textual sentiment series

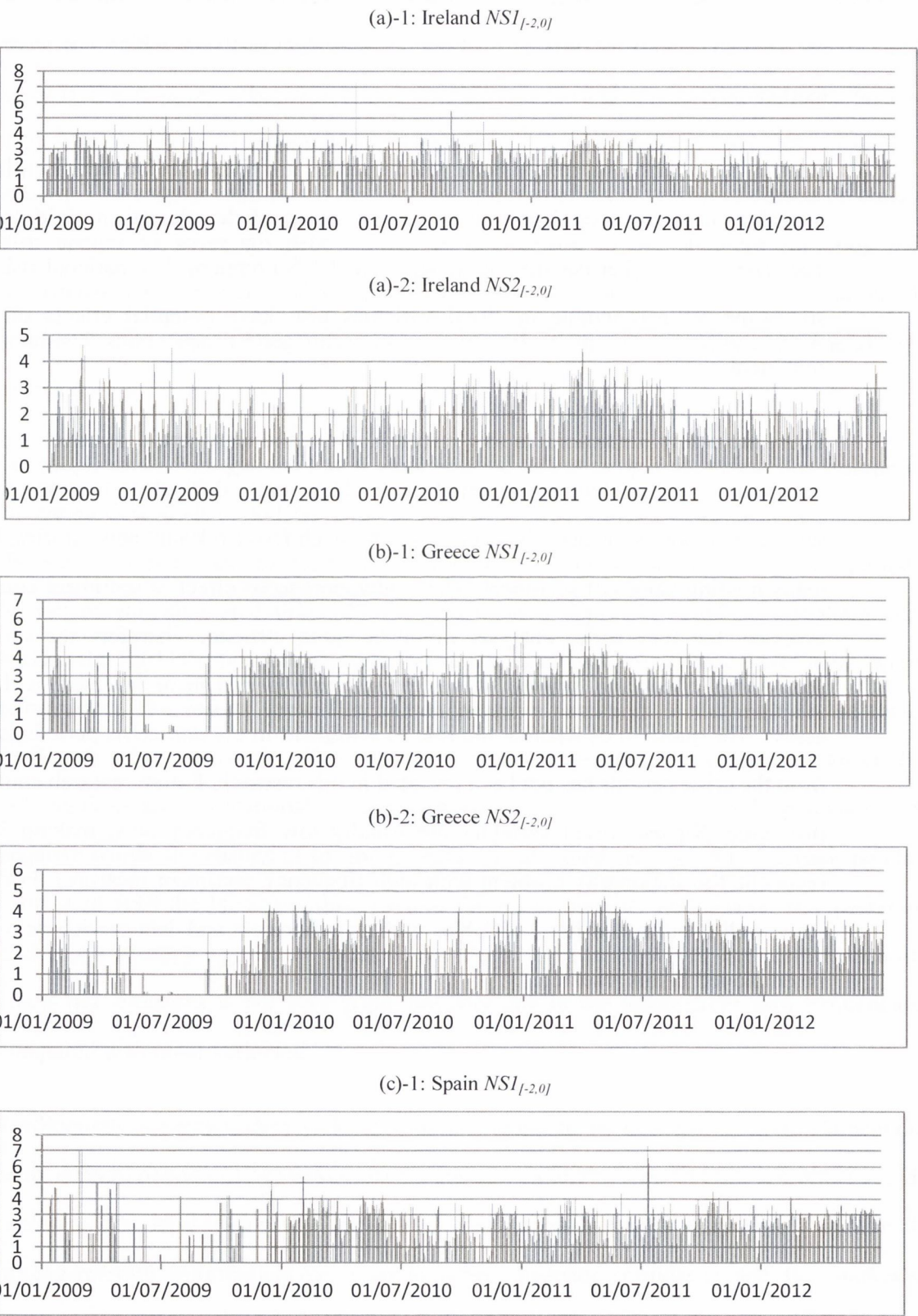
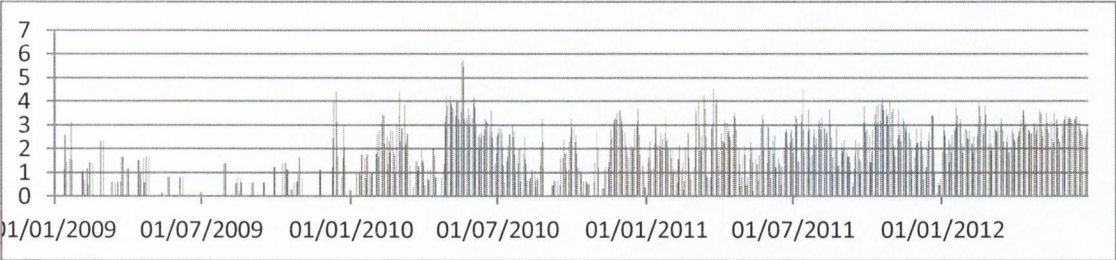
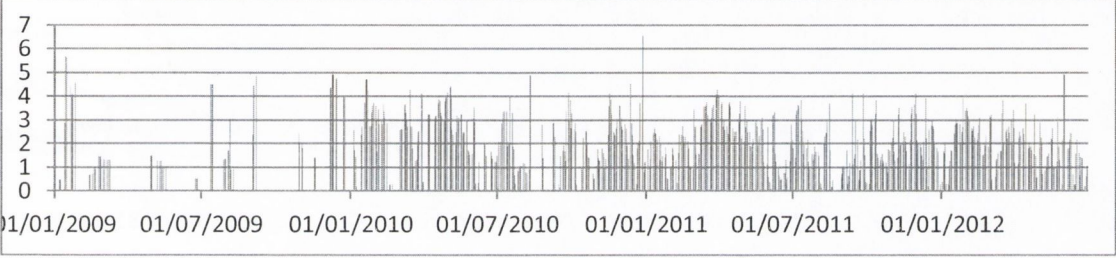


Figure 5.1: Textual Sentiment Series (continued)

(c)-2:Spain $NS2_{[-2,0]}$



(d)-1: Portugal $NS1_{[-2,0]}$



(d)-2: Portugal $NS2_{[-2,0]}$

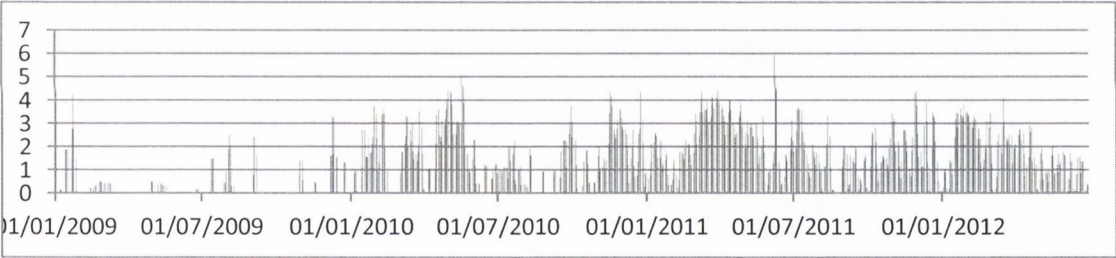
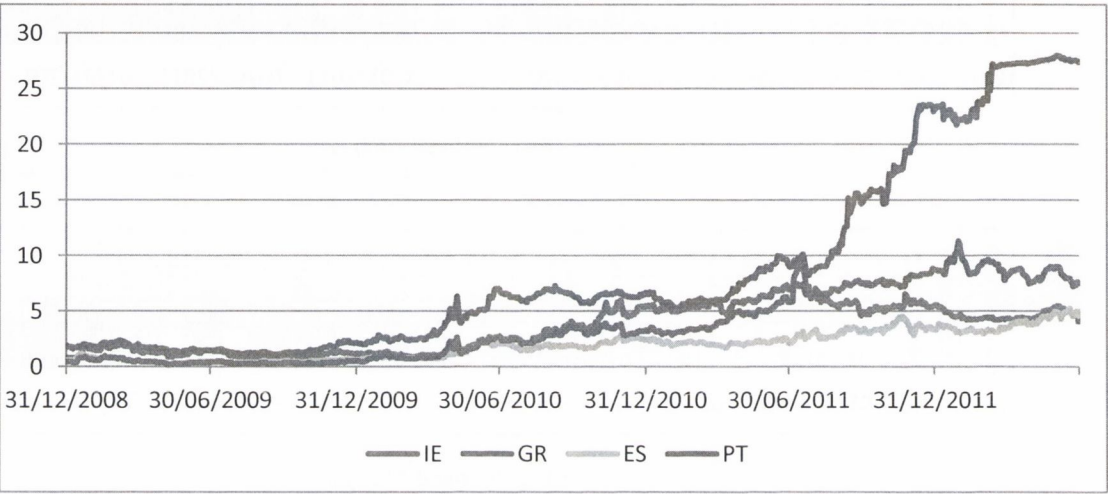


Figure 5.2: Sovereign 10+-year bond yield spreads against Germany



Data source: Datastream and author's calculation

Table 5.1: Number of country-specific news stories

	2009	2010	2011	2012 (First half)
Ireland	450	1058	753	321
Greece	352	1382	1815	1235
Spain	81	897	1198	1763
Portugal	48	714	1169	275

Notes: This table presents the number of collected news stories regarding each country’s sovereign creditworthiness, fiscal conditions, debt crisis, etc.

Table 5.2: Descriptive statistics of sentiment

Country	Measure	Mean	Variance	Skewness	Kurtosis	J-B P-val	Q10 P-val
Ireland	$NSI_{[-2,0]}$	2.37	0.94	-0.12	0.77	0.00	0.00
	$NS2_{[-2,0]}$	1.86	0.93	0.10	-0.65	0.00	0.00
Greece	$NSI_{[-2,0]}$	2.63	1.95	-0.77	-0.27	0.00	0.00
	$NS2_{[-2,0]}$	2.17	1.80	-0.38	-1.08	0.00	0.00
Spain	$NSI_{[-2,0]}$	2.11	1.93	-0.03	-0.23	0.36	0.00
	$NS2_{[-2,0]}$	1.72	1.58	0.11	-0.99	0.00	0.00
Portugal	$NSI_{[-2,0]}$	1.59	1.92	0.46	-0.70	0.00	0.00
	$NS2_{[-2,0]}$	1.33	1.60	0.66	-0.54	0.00	0.00

Notes: This table provides descriptive statistics of two sentiment measures, including mean, variance, skewness, excess kurtosis, the Jarque-Bera (J-B) test of normality, and the Ljung-Box Q test for autocorrelation. The number of observations is 910.

Table 5.3: Panel data unit-root tests

(a)				
Variable	LLC	Breitung	IPS	HT
<i>YS</i>	2.58	3.26	3.58	1.00
<i>NSI</i>	-23.53***	-13.30***	-25.02***	0.70***
<i>NS2</i>	-18.07***	-9.16***	-15.89***	0.86***
<i>CDS</i>	3.14	2.07	3.02	1.00
<i>CAB</i>	-0.28	-2.74***	-1.93**	0.99***
<i>IND</i>	-2.24**	0.08	-0.49	1.00
<i>BBL</i>	-0.10	-2.06**	-0.98	0.98***
<i>LIQ</i>	1.10	-0.08	2.01	1.00
<i>RISK</i>	-8.44***	3.47	-9.38***	0.99*

(b)				
Variable	LLC	Breitung	IPS	HT
<i>ys</i>	-42.22***	-37.49***	-44.04***	0.06***
<i>cds</i>	-35.77***	-38.39***	-42.27***	0.15***
<i>liq</i>	-44.66***	-39.65***	-46.18***	0.04***
<i>risk</i>	-32.04***	-13.30***	-36.82***	0.37***

Notes: This table reports the statistics and the significance level (1 percent ***, 5 percent **, 10 percent *) of panel unit root tests on each variables. The Levin–Lin–Chu (LLC) test, the Breitung test, the Im–Pesaran–Shin (IPS) test and Harris–Tzavalis (HT) test are undertaken. IPS stat listed is the Z-t-tilde-bar stat. Variables shown in panel (b) are the first-differenced variables.

Table 5.4: Fixed-effects regression with Driscoll-Kraay (1998) standard errors

Dependent Variable: Δy								
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NEG	0.004 (2.52**)	0.004 (1.78*)	0.004 (2.37**)	0.004 (2.48**)	0.005 (1.79*)	0.005 (1.82*)	0.005 (2.14**)	0.007 (1.82*)
<i>cds</i>	1.322E-04 (1.86*)	1.317E-04 (1.84*)						
<i>CAB</i>			0.070 (0.49)			0.056 (0.40)		
<i>IND</i>				-7.790E-05 (-0.20)			3.220E-05 (0.09)	
<i>BBL</i>					3.691E-04 (0.76)			2.690E-04 (0.55)
<i>liq</i>	-9.260E-06 (-4.78***)	-9.200E-06 (-4.76***)	-8.120E-06 (-3.91***)	-9.150E-06 (-4.35***)	-8.330E-06 (-2.85***)	-8.060E-06 (-3.89***)	-9.090E-06 (-4.33***)	-8.260E-06 (-2.82***)
<i>risk</i>	0.169 (3.84***)	0.168 (3.78***)	0.176 (3.61***)	0.172 (3.51***)	0.235 (3.99***)	0.174 (3.54***)	0.172 (3.49***)	0.231 (3.88***)
CONSTANT	0.003 (0.81)	0.004 (1.12)	0.014 (0.76)	0.005 (1.27)	0.013 (1.19)	0.012 (0.73)	0.005 (1.38)	0.008 (0.83)
NEG measure	$NSI_{[-2,0]}$	$NS2_{[-2,0]}$	$NSI_{[-2,0]}$	$NSI_{[-2,0]}$	$NSI_{[-2,0]}$	$NS2_{[-2,0]}$	$NS2_{[-2,0]}$	$NS2_{[-2,0]}$
R-squared (Within)	0.07	0.07	0.06	0.05	0.05	0.06	0.05	0.05
Observations	3640	3640	3120	3380	1560	3120	3380	1560

Notes: 1) The table presents the estimated coefficients and their respective T-statistics (in the brackets). Panel regression method is fixed-effects model. The variance and covariance matrix of the residuals is obtained on the basis of the Driscoll and Kraay (1998) approach. The dependent variable is the 1st difference of yield spread. *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively. 2) NEG is the negative sentiment measure. The two specific measures used are $NSI_{[-2,0]}$ and $NS2_{[-2,0]}$. The former is the fraction of negative words in all news stories released from day t-2 to t. The latter is the average of the daily sentiment score from day t-2 to t. 3) Variables in lower case is the first difference of original variables. *cds* represents the CDS premium, *liq* is the proxy for liquidity risk, and *risk* represents the general risk aversion. *CAB*, *IND* and *BBL* correspond to current account balance, industrial production and fiscal budget balance, respectively. *cds*, *liq*, *CAB*, *IND* and *BBL* are defined in differences against Germany.

Table 5.5: Robustness analysis using pooled OLS estimation with Driscoll-Kraay (1998) standard errors

Dependent Variable: ys								
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NEG	0.004 (2.93***)	0.005 (2.06**)	0.004 (2.68***)	0.004 (2.83***)	0.005 (1.79*)	0.006 (2.04**)	0.005 (2.19**)	0.007 (1.83*)
<i>cds</i>	1.352E-04 (1.91*)	1.347E-04 (1.89*)						
<i>CAB</i>			-0.077 (-1.12)			-0.081 (-1.18)		
<i>IND</i>				-3.562E-04 (-1.00)			-2.572E-04 (-0.79)	
<i>BBL</i>					7.254E-04 (1.42)			6.208E-04 (1.22)
<i>liq</i>	-9.280E-06 (-4.79***)	-9.220E-06 (-4.77***)	-8.210E-06 (-3.94***)	-9.180E-06 (-4.35***)	-8.330E-06 (-2.85***)	-8.130E-06 (-3.91***)	-9.120E-06 (-4.34***)	-8.260E-06 (-2.82***)
<i>risk</i>	0.168 (3.83***)	0.166 (3.76***)	0.179 (3.70***)	0.168 (3.46***)	0.234 (3.99***)	0.176 (3.61***)	0.167 (3.43***)	0.230 (3.88***)
CONSTANT	0.002 (0.49)	0.002 (0.74)	-0.005 (-0.50)	0.003 (0.77)	0.018 (1.67)	-0.006 (-0.61)	0.003 (0.88)	0.013 (1.34)
NEG measure	$NSI_{[-2,0]}$	$NS2_{[-2,0]}$	$NSI_{[-2,0]}$	$NSI_{[-2,0]}$	$NSI_{[-2,0]}$	$NS2_{[-2,0]}$	$NS2_{[-2,0]}$	$NS2_{[-2,0]}$
R-squared	0.07	0.07	0.06	0.05	0.05	0.06	0.05	0.05
Observations	3640	3640	3120	3380	1560	3120	3380	1560

Notes: 1) The table presents the estimated coefficients and their respective T-statistics (in the brackets). The estimation method is pooled OLS. The variance and covariance matrix of the residuals is obtained on the basis of the Driscoll and Kraay (1998) approach. The dependent variable is the 1st difference of yield spread. *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively. 2) NEG is the negative sentiment measure. The two specific measures used are $NSI_{[-2,0]}$ and $NS2_{[-2,0]}$. The former is the fraction of negative words in all news stories released from day t-2 to t. The latter is the average of the daily sentiment score from day t-2 to t. 3) Variables in lower case is the first difference of original variables. *cds* represents the CDS premium, *liq* is the proxy for liquidity risk, and *risk* represents the general risk aversion. *CAB*, *IND* and *BBL* correspond to current account balance, industrial production and fiscal budget balance, respectively. *cds*, *liq*, *CAB*, *IND* and *BBL* are defined in differences against Germany.

Table 5.6: Robustness analysis using the fixed effects models with Driscoll-Kraay (1998) standard errors

Dependent Variable: <i>csys</i>								
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NEG	0.004 (2.20**)	0.008 (3.00***)	0.003 (1.89*)	0.004 (2.10**)	0.005 (1.87*)	0.007 (2.77***)	0.008 (2.95***)	0.011 (3.06***)
<i>cscds</i>	8.970E-05 (1.53)	8.910E-05 (1.51)						
<i>CSCAB</i>			0.082 (0.57)			0.087 (0.61)		
<i>CSIND</i>				-5.980E-06 (-0.01)			-1.985E-04 (-0.41)	
<i>CSBBL</i>					-4.333E-04 (-0.81)			-6.428E-04 (-1.20)
<i>csliq</i>	-1.580E-05 (-4.46***)	-1.570E-05 (-4.46***)	-1.290E-05 (-3.70***)	-1.390E-05 (-3.99***)	-1.710E-05 (-2.74***)	-1.280E-05 (-3.71***)	-1.380E-05 (-3.99***)	-1.670E-05 (-2.74***)
CONSTANT	-2.420E-10 -0.26	-2.460E-10 -0.26	6.850E-11 0.13	2.740E-10 0.36	2.388E-03 1.04	6.500E-11 0.13	2.750E-10 0.36	9.987E-04 0.47
NEG measure	<i>CSNS1</i>	<i>CSNS2</i>	<i>CSNS1</i>	<i>CSNS1</i>	<i>CSNS1</i>	<i>CSNS2</i>	<i>CSNS2</i>	<i>CSNS2</i>
R-squared (Within)	0.03	0.04	0.02	0.02	0.02	0.02	0.02	0.03
Observations	3640	3640	3120	3380	1560	3120	3380	1560

Notes: 1) The table presents the estimated coefficients and their respective T-statistics (in the brackets). Estimation method is fixed-effects. The variance and covariance matrix of the residuals is obtained on the basis of the Driscoll and Kraay (1998) approach. The dependent variable is *csys*, a yield spread measure specifically constructed for robustness analysis. *, **, and *** indicate significance level of 1 percent, 5 percent and 10 percent, respectively. 2) NEG is the negative sentiment measure. The two specific measures used are $CSNS1_{[-2,0]}$ and $CSNS2_{[-2,0]}$. 3) Variables in lower case are the first difference of the original variables. All the variables are constructed using the same method, which is described in Section 5.4.1 (robustness analysis).

Table 5.7: Determinants of the country-specific negative sentiment

Dependent Variable: NEG								
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>ys</i>	0.326 (3.06***)	0.316 (2.08**)	0.432 (2.71***)	0.284 (3.04***)	0.239 (1.89*)	0.453 (1.97**)	0.281 (2.23**)	0.349 (1.77*)
<i>cds</i>	1.642E-04 (2.96***)	2.437E-04 (2.66***)						
<i>CAB</i>			0.565 (0.60)			1.104 (1.16)		
<i>IND</i>				-0.043 (-7.25***)			-0.052 (-8.67***)	
<i>BBL</i>					0.022 (4.04***)			0.028 (4.02***)
<i>liq</i>	1.530E-05 (1.80*)	7.930E-07 (0.10)	5.645E-01 (1.71*)	1.630E-05 (1.82*)	2.220E-06 (0.21)	2.760E-07 (0.03)	1.210E-06 (0.16)	-6.390E-06 (-0.60)
<i>risk</i>	0.959 (2.60***)	1.158 (2.82***)	0.962 (2.49***)	0.413 (1.40)	0.510 (1.30)	1.185 (2.84***)	0.476 (1.86*)	0.843 (2.15**)
CONSTANT	2.185 (38.60***)	1.779 (26.34***)	2.219 (21.64***)	1.955 (31.66***)	2.865 (29.66***)	1.835 (16.02***)	1.485 (26.40***)	2.415 (19.10***)
NEG measure	<i>NS1</i>	<i>NS2</i>	<i>NS1</i>	<i>NS1</i>	<i>NS1</i>	<i>NS2</i>	<i>NS2</i>	<i>NS2</i>
R-squared	0.01	0.02	0.01	0.08	0.03	0.02	0.14	0.05
Observations	3640	3640	3120	3380	1560	3120	3380	1560

Notes: 1) This table reports the estimated coefficients and their respective T-statistics (in the brackets). The estimation method is pooled OLS. The variance and covariance matrix of the residuals is obtained on the basis of the Driscoll and Kraay (1998) approach. The dependent variable is the negative sentiment measures. The symbols *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively. 2) NEG is the negative sentiment measure. The two specific measures used are $NS1_{[-2,0]}$ and $NS2_{[-2,0]}$. The former is the fraction of negative words in all news stories released from day t-2 to t. The latter is the average of the daily sentiment score from day t-2 to t. 3) Variables in lower case is the first difference of original variables. *ys* denotes the sovereign yield spreads, *cds* denotes the CDS premium, *liq* is the proxy for liquidity risk, and *risk* represents the general risk aversion. *CAB*, *IND* and *BBL* correspond to current account balance, industrial production and fiscal budget balance, respectively. *cds*, *liq*, *CAB*, *IND* and *BBL* are defined in differences against Germany.

Table 5.8: Determinants of the country-specific negative sentiment (robustness analysis)

Dependent Variable: NEG								
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>csys</i>	0.413 (3.55***)	0.527 (3.93***)	0.546 (3.13***)	0.417 (3.55***)	0.383 (2.19**)	0.705 (3.84***)	0.550 (4.04***)	0.673 (3.82***)
<i>cscds</i>	2.004E-04 (2.25**)	2.129E-04 (1.78*)						
<i>CSCAB</i>			0.789 (0.82)			0.056 (0.06)		
<i>CSIND</i>				-0.009 (-0.95)			0.001 (0.14)	
<i>CSBBL</i>					0.029 (4.22***)			3.197E-02 (3.83***)
<i>cslig</i>	3.010E-05 (1.39)	8.700E-07 (0.05)	3.680E-05 (1.56)	3.150E-05 (1.42)	2.850E-06 (0.08)	7.710E-06 (0.43)	3.080E-06 (0.18)	-2.480E-05 (-0.74)
CONSTANT	-6.500E-10 (-0.89)	3.000E-10 (0.44)	-1.290E-09 (-1.61)	-7.940E-10 (-1.13)	0.376 (10.63***)	-2.330E-10 (-0.37)	-2.770E-10 (-0.45)	0.289 (9.22***)
NEG measure	<i>CSNS1</i>	<i>CSNS2</i>	<i>CSNS1</i>	<i>CSNS1</i>	<i>CSNS1</i>	<i>CSNS2</i>	<i>CSNS2</i>	<i>CSNS2</i>
R-squared	0.00	0.01	0.00	0.04	0.04	0.01	0.01	0.06
Observations	3640	3640	3120	3380	1560	3120	3380	1560

Notes: 1) The table presents the estimated coefficients and their respective T-statistics (in the brackets). The estimation method is pooled OLS. The variance and covariance matrix of the residuals is obtained on the basis of the Driscoll and Kraay (1998) approach. Symbols *, **, and *** indicate significance level of 1 percent, 5 percent and 10 percent, respectively. 2) The dependent variable NEG is the negative sentiment measure. The two specific measures used are $CSNS1_{[-2,0]}$ and $CSNS2_{[-2,0]}$. 3) Variables in lower case are the first difference of the original variables. All the variables are specifically constructed for robustness analysis. The method is described in Section 5.4.1 (robustness analysis).

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Introduction

The role of textual sentiment in financial markets has been of increasing interest to researchers in finance in the most recent decade, particularly since 2007. Researchers have focused on the relationship between textual sentiment and firm fundamentals or market variables. The majority of existing studies have used corporate disclosures, such as annual reports and earnings press releases, as the information source. The literature which studies news stories (or news commentaries) sentiment remains relatively small. This thesis extracted sentiment from firm-specific news stories and examined its impact on asset prices and market activities. Textual sentiment in the news stories could be a relatively objective reflection of the conditions of a firm or other objects. It can also possibly convey new information which has additional explanatory power on asset prices over fundamentals. Compared with corporate disclosures, news stories have the advantage of being available at a higher frequency and covering numerous kinds of events. Compared with internet postings, news stories are generally more objective, less ‘noisier’, and require less pre-processing time.

This thesis has provided a comprehensive analysis of the application of content analysis methods in finance and the role of textual sentiment in different markets. It mainly

addressed the following issues. *First*, by employing panel regressions and VAR models on individual stocks, the interrelations between firm-specific sentiment, firm-level equity returns and trading volumes were examined. The indirect effects of sentiment, namely the extent to which sentiment drives firm-level returns by first working through trading volumes, were also tested. *Second*, this thesis incorporated a first-time research on the relationship between textual sentiment and sovereign debt yield spreads, using data from Greece, Ireland, Portugal and Spain during the European sovereign debt crisis. Novel findings were obtained for both equity and sovereign debt analysis.

This chapter provides concluding comments on each topic and suggests possible areas for consideration in future research. The rest of the chapter is organized as follows. Section 6.2 outlines the primary results and main findings from the thesis. Section 6.3 discusses the limitations of the thesis and provides possible areas for future research. Section 6.4 briefly summarizes this chapter.

6.2 Main findings and contributions

Equity

Chapter 4 examined the interrelations between firm-specific textual sentiment, firm-level equity returns and trading volumes. Panel data analysis discovered that ‘finance negative’ words negatively and significantly predict firm-level equity returns on the following day, mostly at the 5 percent level. The impact of a 10 percent increase in the percentage of negative words today on next day’s stock return is approximately -0.31 to -0.36 basis points. There is also strong evidence of the reverse causal effect of returns on next day’s negative sentiment. The difference between today’s and yesterday’s negative sentiment scores has an even stronger effect (significant at the 1 percent level) on today’s return. Higher negative sentiment causes lower trading volumes on the following day. In addition, there is strong evidence of the ‘indirect’ effects of sentiment on equity returns: sentiment causes trading volumes, whilst trading volumes cause returns.

At the individual firm level, the most evident relations are that trading volumes cause sentiment, and that returns cause trading volumes. The results of rolling-window regressions show that on average, firm-specific sentiment directly causes price returns 9.56 percent of the time. The predictability tends to be concentrated during discrete periods that most likely align with significant newsworthy episodes for each firm. The indirect effects of sentiment on returns are less frequent, and the two effects do not coexist often.

These results are consistent with the literature (e.g. Tetlock (2007), Tetlock et al. (2008), Garcia (2012)): negative words are found to have an immediate negative impact on equity returns, while returns forecast negative words; higher returns predict lower next-day trading volumes. The novel contributions to the literature are as follows. *First*, this research is the first to generate consecutive firm-specific sentiment from news stories for 20 large and liquid MNCs, most of whom have 10-year daily observations. This permits time-series analysis that can be performed to evaluate the interrelations between several lags of sentiment, returns and trading volumes. Previous studies either examined the relations between low-frequency firm-specific sentiment and firm fundamentals, or investigated the effects of daily textual sentiment at the *market* level (i.e. sentiment is extracted from general economics and finance news). In other words, this thesis fills a gap in the literature as nobody has hitherto examined the time-series pattern of sentiment's role for individual stocks. *Second*, this research evaluated the 'indirect' effects of sentiment on stock returns by assuming that textual sentiment acts as a previously unconsidered source of information that drives firm-level returns by first working through trading volumes. *Third*, we have shown that the effects of textual sentiment on equity returns are time-varying and tend to be concentrated during discrete periods. This has not been examined in the existing literature.

Sovereign Debt

Chapter 5 investigated whether textual sentiment is a determinant of sovereign bond yield spread in the context of the European sovereign-debt crisis. The results show that negative sentiment in country-specific news stories within the latest three trading days has a significant positive effect on sovereign debt yield spreads, after controlling for credit risk,

liquidity and international risk aversion measures. Two sentiment measures are used and the results are qualitatively similar. For a 10 percent increase in the percentage of negative words, the change of sovereign yield spreads would rise by 3.8 to 7.1 basis points. Sentiment also contributes to the R-square of regressions. Consistent with the literature, the traditionally identified determinants of sovereign yield spreads, liquidity and international risk aversion measures, are found to have strongly significant effects on sovereign spreads. Meanwhile, sovereign bond yield spreads are found to be significantly directly related to negative textual sentiment. The increases of the change in yield spreads are associated with higher negative sentiment.

This research makes a novel contribution by considering textual sentiment as an additional explanatory variable on sovereign debt yield spreads. Textual sentiment may capture the hard-to-quantify aspects of fundamentals and communicate additional pricing-relevant information. It may also reflect investor confidence and the market's assessment of the risk-return relationship. In addition, it may reflect the behavioral aspects of the market, representing the hidden pricing factors that may not be completely justified by the available facts at hand. This research contributes to both the textual sentiment analysis and sovereign yield spreads literature. It is the first research that discovers that the well-documented role of textual sentiment on equity also applies to debt, and it suggests another perspective for studying the determinants of bond yield spreads or interest rates.

6.3 Limitations in thesis and future work

This section outlines the major limitations of this thesis and suggests some areas for future research.

6.3.1 Limitations

Sample construction issue

In Chapter 4, the sample must be confined to large and liquid MNCs, because the time series analysis requires us to minimize missing values in order to ensure a stable daily news flow. This resulted in only 20 firms qualifying for this selection. The industries represented are quite few, concentrating on energy and technology firms. The sample was not chosen by random selection, so inferences only apply to these firms. The role of textual sentiment may be different for smaller, less liquid firms. It is important to note, however, that this does not bias our tests or results, but maximizes the integrity of our data sample.

Similarly, in Chapter 5, Greece, Ireland, Portugal and Spain were selected to form the sample because they were the most affected countries in the European sovereign debt crisis and had the largest volume of news in this regard. This sample was not chosen by random selection. In addition, the effects of sentiment were only examined in the crisis period. It would be more difficult to examine the role of sentiment in non-crisis periods, when there are much fewer relevant news stories, hence too many missing values of sentiment. So the question that whether the role of sentiment in non-crisis period is different to the crisis period has not been answered in this thesis. Even in the European sovereign debt crisis period, there were not sufficient news stories to generate consecutive daily country-specific sentiment scores. This made the construction of sentiment measures very inflexible.

‘Mixed sentiment’ issue

There is one potential limitation in sentiment extraction that the text corpora may not be absolutely firm-specific or country-specific, as there may be ‘mixed’ sentiment for other relevant firms or countries. Examinations on sub-samples discovered that in the firm-specific news stories corpora (Chapter 4), the names of other Fortune 500 and Global 500 firms¹ (869 firms in total) were mentioned only around 0.3 to 0.6 percent in the total number of words. The firm-specific news stories that mention other firms are either related to their partnerships and business cooperation, joint events, and lawsuits, or are stock market briefings of several related firms. Sentences of isolated descriptions and comments on other firms are few. Obviously the mentioning of other firms’ names is integral to the

¹ Based on the 2011 Fortune 500 and Global 500 lists. The frequencies of mentioning of these firms are examined because besides the sample firm itself (e.g. Boeing), other Fortune 500 and Global 500 firms are assumed to be more likely to appear in its news stories than smaller firms.

content of the firm-specific news. Considering that our text corpora are large and that strict criteria were followed to filter the articles, we believe that any minimal ‘mixed’ sentiment could not materially affect the results we obtained.

Country-specific news stories (Chapter 5) in the context of the European sovereign-debt crisis are inevitably related to other Eurozone countries. The crisis has led to far-reaching consequences that have extended beyond the borders of the ‘question countries’. Due to the growing interconnection of the global financial system and a general loss of investor confidence, it is not only impossible but it may also be unnecessary to completely purify sentiment. Moreover, in the robustness analysis, data and model specifications are revised to mitigate any potential ‘mixed’ sentiment problem.

To summarize, this limitation in sentiment extraction has been fully assessed and accounted for, and does not bias our tests or results.

Data frequency issue

In Chapter 5, sentiment measure employed is at daily frequencies. However, some fundamental variables are only available at low frequencies (e.g. quarterly, annual). This makes it hard to reconcile the stationarity problem with sentiment measures. Therefore, several well-documented determinants of sovereign bond yield spreads, such as the public debt and debt service ratio, could not be used as controlling variables in the regressions. Hence it is not known whether sentiment still contributes to explaining sovereign yield spreads after taking these variables into consideration.

6.3.2 Future work

In light of the contributions and limitations of the thesis, future work could address the following areas.

This thesis has demonstrated that the direct effects of textual sentiment are time-varying and concentrated during discrete periods. Further research could explore the reasons why the effects of sentiment are time-varying and why the predictability on returns is

concentrated during certain periods. It may be that some episodes are particularly newsworthy for a firm, and investors have heightened their sensitivity towards news articles during these periods. It is of great importance to enhance our understanding of the complex relationship between investor behaviors and textual sentiment.

The Mixture of Distributions Hypothesis (MDH) proposed by Clark (1973) suggests that volume and volatility should be positively correlated as they jointly depend on a common underlying latent mixing variable. Textual sentiment could act as a previously unconsidered source of information (i.e. the latent variable) that drives stock returns by first working through trading volumes. Future research could use the firm-specific textual sentiment data to test the MDH.

Extending textual sentiment analysis, researchers are now beginning to explore the effects of complexity (or ‘readability’) of text in the financial markets. For example, Lehavy, Li, and Merkley (2011) and Loughran and McDonald (2013) have examined the relationship between the readability of 10-Ks and analysts’ information processing mechanism and their forecasts. Future research could investigate the complexity of firm-specific news stories and its effects on market activities.

This thesis has affirmed that the well-documented role of textual sentiment in stock market also applies to sovereign debt markets. Similar effects could exist in corporate bond markets as well as other financial markets, markets for commodities, services and derivative asset markets. Future research could examine the relationship between firm-specific textual sentiment and corporate bond yield spreads, and could also investigate the role of sentiment in the other markets mentioned here as well as in other contexts.

6.4 Summary and Conclusions

This chapter has drawn together the main findings and novel contributions of this thesis as well as suggesting future areas for research. Overall, this thesis provides innovative ways to examine the role of firm-level sentiment on individual stocks, and the connection

between textual sentiment and sovereign debt yield spreads. The primary results, which are consistent with the literature, are that negative words have a significant impact on asset prices, both equity and debt. Overall, the analysis constitutes strong evidence for the consideration of textual sentiment as a potentially important factor in asset pricing models. This thesis has potential limitations in the aspects of sample construction, sentiment extraction and nature of data. Future paths of research regarding limitations in both this thesis and the existing literature have been discussed.

APPENDIX

1. Excerpt from the introductory session on GI's official website

'The General Inquirer is basically a mapping tool. It maps each text file with counts on dictionary-supplied categories. The currently distributed version combines the "Harvard IV-4" dictionary content-analysis categories, the "Lasswell" dictionary content-analysis categories, and five categories based on the social cognition work of Semin and Fiedler, making for 182 categories in all...The output is a matrix of "tag counts" for each category, with separate rows of counts for each file processed...The main output from the Inquirer can be displayed as an Excel spreadsheet, with both raw frequency counts and indexes scaled for document length. Statistical tests then evaluate whether there are statistically reliable differences between the texts or groupings of texts being studied.'

2. Examples of the Harvard-IV-4 categories: the word categories are neither mutually exclusive nor exhaustive—one word may fall into multiple categories and some words are not categorized at all.

<u>Pstv</u> : 1045 positive words
<u>Ngvtv</u> : 1160 negative words
<u>Strong</u> : 1902 words implying strength
<u>Weak</u> : 755 words implying weakness
<u>Active</u> : 2045 words implying an active orientation.
<u>Passive</u> : 911 words indicating a passive orientation
<u>Pleasur</u> : 168 words indicating the enjoyment of a feeling, including words indicating confidence, interest and commitment.
<u>Pain</u> : 254 words indicating suffering, lack of confidence, or commitment.
<u>Feel</u> : 49 words describing particular feelings, including gratitude, apathy, and optimism, not those of pain or pleasure.
<u>Ovrst</u> : "Overstated", 696 words indicating emphasis in realms of speed, frequency, causality, inclusiveness, quantity or quasi-quantity, accuracy, validity, scope, size, clarity, exceptionality, intensity, likelihood, certainty and extremity.
<u>Undrst</u> : "Understated", 319 words indicating de-emphasis and caution in these realms.
.....

3. Description of DICTION 5.0 software on <http://rhetorica.net/diction.htm>

DICTION 5.0 by Dr. Roderick P. Hart

DICTION 5.0 counts words based on 33 separate dictionaries (such as familiarity, human interest, tenacity, and self-reference) and two sets of variables and creates numerical frequencies and standard scores for these lists of words. Included with the dictionary scores are scores for five master variables and four calculated variables. The calculated variables, which are based on word ratios, include: (1) insistence, a measure of “code-restriction” that indicates a “preference for a limited, ordered world”; (2) embellishment, a measure of the ratio of adjectives to verbs; (3) variety, a measure of conformity to, or avoidance of, a limited set of expressions (different words/total words); and (4) complexity, a measure of word size based on the Flesch method. The master variables, which the program assumes best capture the major tonal features of a text, are derived from calculations using the scores from the 33 dictionaries and include: (1) certainty, a measure of language “indicating resoluteness, inflexibility, and completeness and a tendency to speak ex cathedra”; (2) activity, a measure of “movement, change, [and] the implementation of ideas and the avoidance of inertia”; (3) optimism, a measure of “language endorsing some

person, group, concept or event or highlighting their positive entailments”; (4) realism, a measure of language “describing tangible, immediate, recognizable matters that affect people’s everyday lives”; and (5) commonality, a measure of language “highlighting the agreed-upon values of a group and rejecting idiosyncratic modes of engagement.” One of the strengths of DICTION 5.0 is that it can compare texts to several sets of norms based on Hart’s research over the past 17 years. The norms are based on the analysis of 22,027 texts of various sorts written between 1948 and 1998.

The data produced by DICTION 5.0 includes scores compared to a "normal range" and standard scores that indicate standard deviations from the norm. Hart’s use of the term "normal range" is somewhat idiosyncratic. It does not refer to the range of plus or minus two standard deviations from the mean, encompassing 95 percent of the data, as the term is typically used in statistical analysis. Instead, Hart uses the term to indicate the range of plus or minus one standard deviation. This normal range encompasses 68 percent of the data. Any score outside the normal range is considered statistically significant in DICTION 5.0. The standard score of an observation is the number of standard deviation units it is above or below the mean; the larger the standard score, the farther it is from the mean.

Formulas for the Master Variables

Certainty = [Tenac. + Level. + Collec. + Insist.] - [Numer. + Ambiv. + Self + Variety]

Optimism = [Praise + Satis. + Inspir.] - [Blame + Hard. + Denial]

Realism = [Famil. + Spat. + Temp. + Prsnt. + Human. + Concr.] - [Past + Complx.]

Activity = [Aggres. + Accomp. + Commun. + Motion] - [Cog. + Passv. + Embell.]

Commonality = [Centr. + Coop. + Rapport] - [Divers. + Exclu. + Liber.]

4. A complete list of Loughran and McDonald's (2011) finance negative words

abandon	accidental	alienates	assaulting	bridge
abandoned	accidentally	alienating	assaults	broken
abandoning	accidents	alienation	assertions	burden
abandonment	accusation	alienations	attrition	burdened
abandonments	accusations	allegation	aversely	burdening
abandons	accuse	allegations	backdating	burdens
abdicated	accused	allege	bad	burdensome
abdicates	accuses	alleged	bail	burned
abdication	accusing	allegedly	bailout	calamities
abdications	acquiesce	alleges	balk	calamitous
aberrant	acquiesced	alleging	balked	calamity
aberration	acquiesces	annoy	bankrupt	cancel
aberrational	acquiescing	annoyance	bankruptcies	canceled
aberrations	acquit	annoyances	bankruptcy	canceled
abetting	acquits	annoyed	bankrupted	cancellation
abnormal	acquittal	annoying	bankrupting	cancellations
abnormalities	acquittals	annoys	bankrupts	cancelled
abnormality	acquitted	annul	bans	cancelling
abnormally	acquitting	annulled	barred	cancels
abolish	adulterate	annulling	barrier	careless
abolished	adulterated	annulment	barriers	carelessly
abolishes	adulterating	annulments	bottleneck	carelessness
abolishing	adulteration	annuls	bottlenecks	casualties
abrogate	adulterations	anomalies	boycott	catastrophe
abrogated	adversarial	anomalous	boycotted	catastrophes
abrogates	adversaries	anomalously	boycotting	catastrophic
abrogating	adversary	anomaly	boycotts	catastrophically
abrogation	adverse	anticompetitive	breach	caution
abrogations	adversely	antitrust	breached	cautionary
abrupt	adversities	argue	breaches	cautioned
abruptly	adversity	argued	breaching	cautioning
abruptness	aftermath	arguing	break	cautions
absence	aftermaths	argument	breakage	cease
absences	against	argumentative	breakages	ceased
absenteeism	aggravate	arguments	breakdown	ceases
abuse	aggravated	arrearage	breakdowns	ceasing
abused	aggravates	arrears	breaking	censure
abuses	aggravating	arrest	breaks	censured
abusing	aggravation	arrested	bribe	censures
abusive	aggravations	arrests	bribed	censuring
abusively	alerted	artificially	briberies	challenge
abusiveness	alerting	assault	bribery	challenged
accident	alienate	assaulted	bribes	challenges
	alienated		bribing	challenging

chargeoffs	complicating	confront	contradicts	criticisms
circumvent	complication	confrontation	contrary	criticize
circumvented	complications	confrontational	controversial	criticized
circumventing	compulsion	confrontations	controversies	criticizes
circumvention	concealed	confronted	controversy	criticizing
circumventions	concealing	confronting	convict	crucial
circumvents	concede	confronts	convicted	crucially
claiming	conceded	confuse	convicting	culpability
claims	concedes	confused	conviction	culpable
clawback	conceding	confuses	convictions	culpably
closed	concern	confusing	corrected	cumbersome
closeout	concerned	confusingly	correcting	curtail
closeouts	concerns	confusion	correction	curtailed
closing	conciliating	conspiracies	corrections	curtailing
closings	conciliation	conspiracy	corrects	curtailment
closure	conciliations	conspirator	corrupt	curtailments
closures	condemn	conspiratorial	corrupted	curtails
coerce	condemnation	conspirators	corrupting	cut
coerced	condemnations	conspire	corruption	cutback
coerces	condemned	conspired	corruptions	cutbacks
coercing	condemning	conspires	corruptly	damage
coercion	condemns	conspiring	corruptness	damaged
coercive	condone	construe	costly	damages
collapse	condoned	construed	counterclaim	damaging
collapsed	confess	construes	counterclaimed	dampen
collapses	confessed	construing	counterclaiming	dampened
collapsing	confesses	contempt	counterclaims	danger
collision	confessing	contend	counterfeit	dangerous
collisions	confession	contended	counterfeited	dangerously
collude	confine	contending	counterfeiter	dangers
colluded	confined	contends	counterfeiters	deadlock
colludes	confinement	contention	counterfeiting	deadlocked
colluding	confinements	contentions	counterfeits	deadlocking
collusion	confines	contentious	countermeasure	deadlocks
collusions	confining	contentiously	countermeasures	deadweight
collusive	confiscate	contested	crime	deadweights
complain	confiscated	contesting	crimes	debarment
complained	confiscates	contraction	criminal	debarments
complaining	confiscating	contractions	criminally	debarred
complains	confiscation	contradict	criminals	deceased
complaint	confiscations	contradicted	crises	deceit
complaints	conflict	contradicting	crisis	deceitful
complicate	conflicted	contradiction	critical	deceitfulness
complicated	conflicting	contradictions	critically	deceive
complicates	conflicts	contradictory	criticism	deceived

Appendix

deceives	deficiency	demoting	detention	diminish
deceiving	deficient	demotion	detentions	diminished
deception	deficit	demotions	deter	diminishes
deceptions	deficits	denial	deteriorate	diminishing
deceptive	defraud	denials	deteriorated	diminution
deceptively	defrauded	denied	deteriorates	disadvantage
decline	defrauding	denies	deteriorating	disadvantaged
declined	defrauds	denigrate	deterioration	disadvantageous
declines	defunct	denigrated	deteriorations	disadvantages
declining	degradation	denigrates	deterred	disaffiliation
deepened	degradations	denigrating	deterrence	disagree
deepening	degrade	denigration	deterrences	disagreeable
deepens	degraded	deny	deterrent	disagreed
deeper	degrades	denying	deterrents	disagreeing
deepest	degrading	deplete	detering	disagreement
deface	delay	depleted	deters	disagreements
defaced	delayed	depletes	detract	disagrees
defacement	delaying	depleting	detracted	disallow
defamation	delays	depletion	detracting	disallowance
defamations	deleterious	depletions	detriment	disallowances
defamatory	deliberate	deprecation	detrimental	disallowed
defame	deliberated	depress	detrimentally	disallowing
defamed	deliberately	depressed	detriments	disallows
defames	delinquencies	depresses	devalue	disappear
defaming	delinquency	depressing	devalued	disappearance
default	delinquent	deprivation	devalues	disappearances
defaulted	delinquently	deprive	devaluing	disappeared
defaulting	delinquents	deprived	devastate	disappearing
defaults	delist	deprives	devastated	disappears
defeat	delisted	depriving	devastating	disappoint
defeated	delisting	derelict	devastation	disappointed
defeating	delists	dereliction	deviate	disappointing
defeats	demise	derogatory	deviated	disappointingly
defect	demised	destabilization	deviates	disappointment
defective	demises	destabilize	deviating	disappointments
defects	demising	destabilized	deviation	disappoints
defend	demolish	destabilizing	deviations	disapproval
defendant	demolished	destroy	devolve	disapprovals
defendants	demolishes	destroyed	devolved	disapprove
defended	demolishing	destroying	devolves	disapproved
defending	demolition	destroys	devolving	disapproves
defends	demolitions	destruction	difficult	disapproving
defensive	demote	destructive	difficulties	disassociates
defer	demoted	detain	difficultly	disassociating
deficiencies	demotes	detained	difficulty	disassociation

disassociations	disgorgement	displaces	dissolutions	downgraded
disaster	disgorgements	displacing	distort	downgrades
disasters	disgorges	dispose	distorted	downgrading
disastrous	disgorging	dispossess	distorting	downsize
disastrously	disgrace	dispossessed	distortion	downsized
disavow	disgraceful	dispossesses	distortions	downsizes
disavowal	disgracefully	dispossessing	distorts	downsizing
disavowed	dishonest	disproportion	distract	downsizings
disavowing	dishonestly	disproportional	distracted	downtime
disavows	dishonesty	disproportionate	distracting	downtimes
disciplinary	dishonor	disproportionately	distraction	downturn
disclaim	dishonorable	dispute	distractions	downturns
disclaimed	dishonorably	disputed	distracts	downward
disclaimer	dishonored	disputes	distress	downwards
disclaimers	dishonoring	disputing	distressed	drag
disclaiming	dishonors	disqualification	disturb	drastic
disclaims	disincentives	disqualifications	disturbance	drastically
disclose	disinterested	disqualified	disturbances	drawback
disclosed	disinterestedly	disqualifies	disturbed	drawbacks
discloses	disinterestedness	disqualify	disturbing	dropped
disclosing	disloyal	disqualifying	disturbs	drought
discontinuance	disloyally	disregard	diversion	droughts
discontinuances	disloyalty	disregarded	divert	duress
discontinuation	dismal	disregarding	diverted	dysfunction
discontinuations	dismally	disregards	diverting	dysfunctional
discontinue	dismiss	disreputable	diverts	dysfunctions
discontinued	dismissal	disrepute	divest	easing
discontinues	dismissals	disrupt	divested	egregious
discontinuing	dismissed	disrupted	divesting	egregiously
discourage	dismisses	disrupting	divestiture	embargo
discouraged	dismissing	disruption	divestitures	embargoed
discourages	disorderly	disruptions	divestment	embargoes
discouraging	disparage	disruptive	divestments	embargoing
discredit	disparaged	disrupts	divests	embarrass
discredited	disparagement	dissatisfaction	divorce	embarrassed
discrediting	disparagements	dissatisfied	divorced	embarrasses
discredits	disparages	dissent	divulge	embarrassing
discrepancies	disparaging	dissented	divulged	embarrassment
discrepancy	disparagingly	dissenter	divulges	embarrassments
disfavor	disparities	dissenters	divulging	embezzle
disfavored	disparity	dissenting	doubt	embezzled
disfavoring	displace	dissents	doubted	embezzlement
disfavors	displaced	dissident	doubtful	embezzlements
disgorge	displacement	dissidents	doubts	embezzler
disgorged	displacements	dissolution	downgrade	embezzles

Appendix

embezzling	evasions	expropriate	flaws	fugitives
encroach	evasive	expropriated	forbid	gratuitous
encroached	evict	expropriates	forbidden	gratuitously
encroaches	evicted	expropriating	forbidding	grievance
encroaching	evicting	expropriation	forbids	grievances
encroachment	eviction	expropriations	force	grossly
encroachments	evictions	expulsion	forced	groundless
encumber	evicts	expulsions	forcing	guilty
encumbered	exacerbate	extenuating	foreclose	halt
encumbering	exacerbated	fail	foreclosed	halted
encumbers	exacerbates	failed	forecloses	hamper
encumbrance	exacerbating	failing	foreclosing	hampered
encumbrances	exacerbation	failings	foreclosure	hampering
endanger	exacerbations	fails	foreclosures	hampers
endangered	exaggerate	failure	forego	harass
endangering	exaggerated	failures	foregoes	harassed
endangerment	exaggerates	fallout	foregoing	harassing
endangers	exaggerating	false	foregone	harassment
enjoin	exaggeration	falsely	forestall	hardship
enjoined	excessive	falsification	forestalled	hardships
enjoining	excessively	falsifications	forestalling	harm
enjoins	exculpate	falsified	forestalls	harmed
erode	exculpated	falsifies	forfeit	harmful
eroded	exculpates	falsify	forfeited	harmfully
erodes	exculpating	falsifying	forfeiting	harming
eroding	exculpation	falsity	forfeits	harms
erosion	exculpations	fatalities	forfeiture	harsh
erratic	exculpatory	fatality	forfeitures	harsher
erratically	exonerate	fatally	forgers	harshes
erred	exonerated	fault	forgery	harshly
erring	exonerates	faulted	fraud	harshness
erroneous	exonerating	faults	frauds	hazard
erroneously	exoneration	faulty	fraudulence	hazardous
error	exonerations	fear	fraudulent	hazards
errors	exploit	fears	fraudulently	hinder
errs	exploitation	felonies	frivolous	hindered
escalate	exploitations	felonious	frivolously	hindering
escalated	exploitative	felony	frustrate	hinders
escalates	exploited	fictitious	frustrated	hindrance
escalating	exploiting	fin	frustrates	hindrances
evade	exploits	fines	frustrating	hostile
evaded	expose	fired	frustratingly	hostility
evades	exposed	firing	frustration	hurt
evading	exposes	flaw	frustrations	hurting
evasion	exposing	flawed	fugitive	idle

idled	impossible	incapacity	ineffectiveness	insufficiency
idling	impound	incarcerate	inefficiencies	insufficient
ignore	impounded	incarcerated	inefficiency	insufficiently
ignored	impounding	incarcerates	inefficient	insurrection
ignores	impounds	incarcerating	inefficiently	insurrections
ignoring	impracticable	incarceration	ineligibility	intentional
ill	impractical	incarcerations	ineligible	interfere
illegal	impracticalities	incidence	inequitable	interfered
illegalities	impracticality	incidences	inequitably	interference
illegality	imprisonment	incident	inequities	interferences
illegally	improper	incidents	inequity	interferes
illegible	improperly	incompatibilities	inevitable	interfering
illicit	improprieties	incompatibility	inexperience	intermittent
illicitly	impropriety	incompatible	inexperienced	intermittently
illiquid	imprudent	incompetence	inferior	interrupt
illiquidity	imprudently	incompetency	inflicted	interrupted
imbalance	inability	incompetent	infraction	interrupting
imbalances	inaccessible	incompetently	infractions	interruption
immature	inaccuracies	incompetents	infringe	interruptions
immoral	inaccuracy	incomplete	infringed	interrupts
impair	inaccurate	incompletely	infringement	intimidation
impaired	inaccurately	incompleteness	infringements	intrusion
impairing	inaction	inconclusive	infringes	invalid
impairment	inactions	inconsistencies	infringing	invalidate
impairments	inactivate	inconsistency	inhibited	invalidated
impairs	inactivated	inconsistent	inimical	invalidates
impasse	inactivates	inconsistently	injunction	invalidating
impasses	inactivating	inconvenience	injunctions	invalidation
impede	inactivation	inconveniences	injure	invalidity
impeded	inactivations	inconvenient	injured	investigate
impedes	inactivity	incorrect	injures	investigated
impediment	inadequacies	incorrectly	injuries	investigates
impediments	inadequacy	incorrectness	injuring	investigating
impeding	inadequate	indecent	injurious	investigation
impending	inadequately	indecent	injury	investigations
imperative	inadvertent	indefeasible	inordinate	involuntarily
imperfection	inadvertently	indefeasibly	inordinately	involuntary
imperfections	inadvisability	indict	inquiry	irreconcilable
imperil	inadvisable	indictable	insecure	irreconcilably
impermissible	inapplicable	indicted	insensitive	irrecoverable
implicate	inappropriate	indicting	insolvencies	irrecoverably
implicated	inappropriately	indictment	insolvency	irregular
implicates	inattention	indictments	insolvent	irregularities
implicating	incapable	ineffective	instability	irregularity
impossibility	incapacitated	ineffectively	insubordination	irregularly

Appendix

irreparable	litigations	miscalculation	mismanaging	monopolizes
irreparably	lockout	miscalculations	mismatch	monopolizing
irreversible	lockouts	mischief	mismatched	monopoly
jeopardize	lose	misclassification	mismatches	moratoria
jeopardized	loses	misclassified	mismatching	moratorium
justifiable	losing	misconduct	misplaced	moratoriums
kickback	loss	misdated	misrepresent	mothballed
kickbacks	losses	misdeemeanor	misrepresentatio	mothballing
knowingly	lost	misdeemeanors	n	negative
lack	lying	misdirected	misrepresentatio	negatively
lacked	malfeasance	mishandle	ns	negatives
lacking	malfunction	mishandled	misrepresented	neglect
lackluster	malfunctioned	mishandles	misrepresenting	neglected
lacks	malfunctioning	mishandling	misrepresents	neglectful
lag	malfunctions	misinform	miss	neglecting
lagged	malice	misinformation	missed	neglects
lagging	malicious	misinformed	misses	negligence
lags	maliciously	misinforming	misstate	negligences
lapse	malpractice	misinforms	misstated	negligent
lapsed	manipulate	misinterpret	misstatement	negligently
lapses	manipulated	misinterpretation	misstatements	nonattainment
lapsing	manipulates	misinterpretations	misstates	noncompetitive
late	manipulating	misinterpreted	misstating	noncompliance
laundering	manipulation	misinterpreting	misstep	noncompliances
layoff	manipulations	misinterprets	missteps	noncompliant
layoffs	manipulative	misjudge	mistake	noncomplying
lie	markdown	misjudged	mistaken	nonconforming
limitation	markdowns	misjudges	mistakenly	nonconformities
limitations	misapplication	misjudging	mistakes	nonconformity
lingering	misapplications	misjudgment	mistaking	nondisclosure
liquidate	misapplied	misjudgments	mistrial	nonfunctional
liquidated	misapplies	mislabel	mistrials	nonpayment
liquidates	misapply	mislabeled	misunderstand	nonpayments
liquidating	misapplying	mislabeling	misunderstanding	nonperformance
liquidation	misappropriate	mislabelled	misunderstandings	nonperformances
liquidations	misappropriated	mislabels	misunderstood	nonperforming
liquidator	misappropriates	mislead	misuse	nonproducing
liquidators	misappropriating	misleading	misused	nonproductive
litigant	misappropriation	misleadingly	misuses	nonrecoverable
litigants	misappropriations	misleads	misusing	nonrenewal
litigate	misbranded	misled	monopolistic	nuisance
litigated	miscalculate	mismanage	monopolists	nuisances
litigates	miscalculated	mismanaged	monopolization	nullification
litigating	miscalculates	mismanagement	monopolize	nullifications
litigation	miscalculating	mismanages	monopolized	nullified

nullified	outages	overruns	persists	preventing
nullifies	outdated	overshadow	pervasive	prevention
nullify	outmoded	overshadowed	pervasively	prevents
nullifying	overage	overshadowing	pervasiveness	problem
objected	overages	overshadows	petty	problematic
objecting	overbuild	overstate	picket	problematical
objection	overbuilding	overstated	picketed	problems
objectionable	overbuilds	overstatement	picketing	prolong
objectionably	overbuilt	overstatements	plaintiff	prolongation
objections	overburden	overstates	plaintiffs	prolongations
obscene	overburdened	overstating	plea	prolonged
obscenity	overburdening	oversupplied	plead	prolonging
obsolescence	overcapacities	oversupplies	pleaded	prolongs
obsolete	overcapacity	oversupply	pleading	prone
obstacle	overcharge	oversupplying	pleadings	prosecute
obstacles	overcharged	overtly	pleads	prosecuted
obstruct	overcharges	overturn	pleas	prosecutes
obstructed	overcharging	overturned	pled	prosecuting
obstructing	overcome	overturning	poor	prosecution
obstruction	overcomes	overturns	poorly	prosecutions
obstructions	overcoming	overvalue	poses	protest
offence	overdue	overvalued	posing	protested
offences	overestimate	overvaluing	postpone	protester
offend	overestimated	panic	postponed	protesters
offended	overestimates	panics	postponement	protesting
offender	overestimating	penalize	postponements	protestor
offenders	overestimation	penalized	postpones	protestors
offending	overestimations	penalizes	postponing	protests
offends	overload	penalizing	precipitated	protracted
omission	overloaded	penalties	precipitous	protraction
omissions	overloading	penalty	precipitously	provoke
omit	overloads	peril	preclude	provoked
omits	overlook	perils	precluded	provokes
omitted	overlooked	perjury	precludes	provoking
omitting	overlooking	perpetrate	precluding	punished
onerous	overlooks	perpetrated	predatory	punishes
opportunistic	overpaid	perpetrates	prejudice	punishing
opportunistically	overpayment	perpetrating	prejudiced	punishment
oppose	overpayments	perpetration	prejudices	punishments
opposed	overproduced	persist	prejudicial	punitive
opposes	overproduces	persisted	prejudicing	purport
opposing	overproducing	persistence	premature	purported
opposition	overproduction	persistent	prematurely	purportedly
oppositions	overrun	persistently	pressing	purporting
outage	overrunning	persisting	pretrial	purports

Appendix

question	refinances	repudiating	sacrifices	slippage
questionable	refinancing	repudiation	sacrificial	slippages
questionably	refinancings	repudiations	sacrificing	slow
questioned	refusal	resign	scandalous	slowdown
questioning	refusals	resignation	scandals	slowdowns
questions	refuse	resignations	scrutinize	slowed
quit	refused	resigned	scrutinized	slower
quitting	refuses	resigning	scrutinizes	slowest
racketeer	refusing	resigns	scrutinizing	slowing
racketeering	reject	restate	scrutiny	slowly
rationalization	rejected	restated	secrecy	slowness
rationalizations	rejecting	restatement	seize	sluggish
rationalize	rejection	restatements	seized	sluggishly
rationalized	rejections	restates	seizes	sluggishness
rationalizes	rejects	restating	seizing	solvencies
rationalizing	relinquish	restructure	sentenced	solvency
reassessment	relinquished	restructured	sentencing	staggering
reassessments	relinquishes	restructures	serious	stagnant
reassign	relinquishing	restructuring	seriously	stagnate
reassigned	relinquishment	restructurings	seriousness	stagnated
reassigning	relinquishments	retaliate	setback	stagnates
reassignment	reluctance	retaliated	setbacks	stagnating
reassignments	reluctant	retaliates	sever	stagnation
reassigns	renegotiate	retaliating	severe	standstill
recall	renegotiated	retaliation	severed	standstills
recalled	renegotiates	retaliations	severely	stolen
recalling	renegotiating	retaliatory	severities	stoppage
recalls	renegotiation	retribution	severity	stoppages
recession	renegotiations	retributions	sharply	stopped
recessionary	renounce	revocation	shocked	stopping
recessions	renounced	revocations	shortage	stops
reckless	renouncement	revoke	shortages	strain
recklessly	renouncements	revoked	shortfall	strained
recklessness	renounces	revokes	shortfalls	straining
redact	renouncing	revoking	shrinkage	strains
redacted	reparation	ridicule	shrinkages	stress
redacting	reparations	ridiculed	shut	stressed
redaction	repossessed	ridicules	shutdown	stresses
redactions	repossesses	ridiculing	shutdowns	stressful
redress	repossessing	riskier	shuts	stressing
redressed	repossession	riskiest	shutting	stringent
redresses	repossessions	risky	slander	subjected
redressing	repudiate	sabotage	slandered	subjecting
refinance	repudiated	sacrifice	slandorous	subjection
refinanced	repudiates	sacrificed	slanders	subpoena

subpoenaed	terminations	uncollectibles	understates	unjustifiably
subpoenas	testify	uncompetitive	understating	unjustified
substandard	testifying	uncompleted	underutilization	unjustly
sue	threat	unconscionable	underutilized	unknowing
sued	threaten	unconscionably	undesirable	unknowingly
sues	threatened	uncontrollable	undesired	unlawful
suffer	threatening	uncontrollably	undetected	unlawfully
suffered	threatens	uncontrolled	undetermined	unlicensed
suffering	threats	uncorrected	undisclosed	unliquidated
suffers	tightening	uncover	undocumented	unmarketable
suing	tolerate	uncovered	undue	unmerchantable
summoned	tolerated	uncovering	unduly	unnecessarily
summoning	tolerates	uncovers	uneconomic	unnecessary
summons	tolerating	undeliverable	uneconomical	unneded
summonses	toleration	undelivered	uneconomically	unobtainable
surrender	tortuous	undercapitalized	unemployed	unoccupied
surrendered	tortuously	undercut	unemployment	unpaid
surrendering	tragedies	undercuts	unethical	unperformed
surrenders	tragedy	undercutting	unethically	unplanned
susceptibility	tragic	underestimate	unexcused	unpopular
susceptible	tragically	underestimated	unexpected	unpredictability
suspect	traumatic	underestimates	unexpectedly	unpredictable
suspected	trouble	underestimating	unfair	unpredictably
suspects	troubled	underestimation	unfairly	unpredicted
suspend	troubles	underfunded	unfavorable	unproductive
suspended	turbulence	underinsured	unfavorably	unprofitability
suspending	turmoil	undermine	unfavourable	unprofitable
suspends	unable	undermined	unfeasible	unqualified
suspension	unacceptable	undermines	unfit	unrealistic
suspensions	unacceptably	undermining	unfitness	unreasonable
suspicion	unaccounted	underpaid	unforeseeable	unreasonableness
suspicious	unannounced	underpayment	unforeseen	unreasonably
suspiciously	unanticipated	underpayments	unforseen	unrecoverable
taint	unapproved	underpays	unfortunate	unrecovered
tainted	unattractive	underperform	unfortunately	unreimbursed
tainting	unauthorized	underperformance	unfounded	unreliable
taints	unavailability	underperformed	unfriendly	unremedied
tampered	unavailable	underperforming	unfulfilled	unreported
tense	unavoidable	underproduced	unfunded	unresolved
terminate	unavoidably	underproduction	uninsured	unrest
terminated	unaware	underreporting	unintended	unsafe
terminates	uncollectable	understate	unintentional	unsalable
terminating	uncollected	understated	unintentionally	unsaleable
termination	uncollectibility	understatement	unjust	unsatisfactory
	uncollectible	understatements	unjustifiable	unsatisfied

Appendix

unsavory	untruths	violated	vulnerable	worry
unscheduled	unusable	violates	vulnerably	worrying
unsold	unwanted	violating	warn	worse
unsound	unwarranted	violation	warned	worsen
unstable	unwelcome	violations	warning	worsened
unsubstantiated	unwilling	violative	warnings	worsening
unsuccessful	unwillingness	violator	warns	worsens
unsuccessfully	upset	violators	wasted	worst
unsuitability	urgency	violence	wasteful	worthless
unsuitable	urgent	violent	wasting	writedown
unsuitably	usurious	violently	weak	writedowns
unsuited	usurp	vitiate	weaken	writeoff
unsure	usurped	vitiated	weakened	writeoffs
unsuspected	usurping	vitiates	weakening	wrong
unsuspecting	usurps	vitiating	weakens	wrongdoing
unsustainable	usury	vitiating	weaker	wrongdoings
untenable	vandalism	voided	weakest	wrongful
untimely	verdict	voiding	weakly	wrongfully
untruth	verdicts	volatile	weakness	wrongly
untruthful	vetoed	volatility	weaknesses	
untruthfully	victims	vulnerabilities	willfully	
untruthfulness	violate	vulnerability	worries	

Table A.1: VAR Lag Length Tests

Company	Test	LR statistics	Significance
Apple	5 v 4	46.25	0.00
	5 v 3	130.13	0.00
	5 v 2	206.56	0.00
	5 v 1	315.77	0.00
Chevron	5 v 4	13.91	0.13
	4 v 3	24.40	0.00
	4 v 2	75.57	0.00
	4 v 1	223.00	0.00

Notes: This table presents the sample results of VAR lag length tests, for individual firms on their whole sample period.

Table A.2: Hausman's specification test (sample results)

	Equation 4.4	Equation 4.5	Equation 4.6	Equation 5.2
Test Statistics	11.13	2898.88	1209.45	12.16
P-value	0.19	0.00	0.00	0.01
FE or RE?	RE	FE	FE	FE

Notes: This table displays the sample results of Hausman's specification test (to choose between the fixed effects and the random effects model) for Equation 4.4, 4.5, 4.6 and 5.2.

Table A.3: Testing heteroskedasticity, serial correlation and cross-sectional dependence
in the panel data (Chapter 5)

(a) The fixed-effects regression

Dependent variable: <i>ys</i>	
Independent Variables	(1)
NEG	0.004 (1.91*)
<i>cds</i>	1.322E-04 (8.66***)
<i>liq</i>	-9.260E-06 (-4.78***)
<i>risk</i>	0.169 (6.92***)
CONSTANT	0.003 (0.57)
NEG measure	$NSI_{[-2,0]}$
R-squared (Within)	0.07
Observations	3640

(b) Test results

	Statistics	P-values	Indication
Heteroskedasticity	847.42	0.00	Heteroskedasticity exists
Serial correlation	27.16	0.01	Serial correlation exists
Cross-sectional dependence	23.26	0.00	Cross-sectional dependence exists

Notes: This table displays the results of testing if heteroskedasticity, serial correlation and cross-sectional dependence exist in data (Chapter 5). Panel (a) presents the estimated coefficients and T-statistics (in the bracket, and the symbols *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively) of a proposed fixed-effects regression. *ys* denotes the sovereign yield spreads, *cds* denotes the CDS premium, *liq* is the proxy for liquidity risk, and *risk* represents the general risk aversion. Panel (b) displays the test statistics and p-values of each test. The results of other proposed models (not reported) are qualitatively the same.

BIBLIOGRAPHY

Afonso A. D. Furceri and P. Gomes (2011). Sovereign credit ratings and financial markets linkage: application to European data. ECB Working Paper No. 1347. Available at SSRN: <http://ssrn.com/abstract=1847505>.

Antweiler, W. and M. Z. Frank (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance*, **59**, 1259-1294.

Attinasi, M-G., Checherita, C. and C. Nickel (2009). What explains the surge in euro area sovereign spreads during the financial crisis of 2007-09?. ECB Working Paper, No.1131. Available at SSRN: <http://ssrn.com/abstract=1520351>.

Baker, M. and J. Wurgler (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, **21**, 129-151.

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data*, 3rd edition. Chichester: John Wiley & Sons.

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*, 4th edition. Chichester: John Wiley & Sons.

Barbosa, L. and S. Costa (2010). Determinants of sovereign bond yield spreads in the euro area in the context of the economic and financial crisis. Banco de Portugal, Working Paper, No.22.

Barrios S., P. Iversen, M. Lewandowska and R. Setzer (2009), Determinants of intra-euro area government bond spreads during the financial crisis, European Economy. Economic Papers No. 388.

Beber, A., M. W. Brandt, and K. A. Kavajecz (2009). Flight-to-Quality or Flight-to-Liquidity? Evidence from the Euro-Area Bond Market. *Review of Financial Studies*, **22**, 925-957.

- Bera, A. K., M. L. Higgins, and S. Lee (1992). Interaction between autocorrelation and conditional heteroscedasticity: A random-coefficient approach. *Journal of Business & Economic Statistics*, **10**, 133-142.
- Bernoth, K., J. von Hagen and L. Schuknecht (2006), Sovereign Risk Premiums in the European Government Bond Market, Governance and the Efficiency of Economic Systems (GESY), Discussion Paper, No. 151.
- Black, F. (1986). Noise. *Journal of Finance*, **41**, 529-543.
- Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.
- Breitung, J., (2000). The local power of some unit root tests for panel data. *Advances in Econometrics*, **15**, 161–177.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, **52**, 57-82.
- Chen, H., P. De, Y. Hu and B. Hwang (2012). Customers as advisors: the role of social media in financial markets. Available at SSRN: <http://ssrn.com/abstract=1807265>.
- Clark, P. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, **41**(1), 135-155.
- Codogno, L., C. Favero and A. Missale (2003), Yield spreads on EMU government bonds, *Economic Policy*, **18**(37), 503–532.
- Das, S.R. and M.Y. Chen (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science*. **53**, 1375–1388.
- Davis, A.K., J.M. Piger, and L.M. Sedor (2008). Beyond the numbers: managers' use of optimistic and pessimistic tone in earnings press releases. Available at SSRN: <http://ssrn.com/abstract=875399>.
- De Bondt, W.F.M., and R. Thaler (1985). Does the stock market overreact? *Journal of Finance*, **40**, 793-905.
- DeGennaro, R., and R. Shrieves (1997). Public information releases, private information arrival and volatility in the foreign exchange market. *Journal of Empirical Finance*, **4**, 295–315.
- Demers, E., and C. Vega (2010). Soft information in earnings announcement: news or noise?, INSEAD Working Paper No. 2010/33/AC. Available at SSRN: <http://ssrn.com/abstract=1153450>.
- Dickey D. A., Fuller W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, **74**, 427–

431.

Doran, J. S., D. R. Peterson and S. M. Price (2010). Earnings conference call content and stock price: the case of REITs. *Journal of Real Estate Finance and Economics*, **45**(2), 402-434.

Duffie, D. and K. Singleton (1999). Modeling the term structure of defaultable bonds. *Review of Financial Studies*, **12**, 687-720.

Ejsing, J. W. and J. Sihvonen (2009). Liquidity premia in German government bonds. ECB Working Paper, No.1081. Available at SSRN: <http://ssrn.com/abstract=1456858>.

Elton, E., M. Gruber, and J. Busse (1998). Do investors care about sentiment?. *Journal of Business*, **71**, 477-500.

Enders, W. (1994). *Applied Econometric Time Series*. New York: Wiley & Sons.

Engelberg, J. (2008). Costly information processing: evidence from earnings announcements. AFA 2009 San Francisco Meetings Paper. Available at SSRN: <http://ssrn.com/abstract=1107998>.

Engle, R., D. Lilien, and R. Robins (1987). Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica*, **55**, 391-407.

Faini, R. (2006). Fiscal Policy and Interest Rates in Europe. *Economic Policy*, **21**(47), 443-489.

Fama, E. F. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, **25**, 383-417.

Fama, E. F. (1991). Efficient Capital Markets: II. *Journal of Finance*, **46**, 1575-1617.

Fama, E. F. and K. French, (1992). The cross-section of expected stock returns. *Journal of Finance*, **47**, 427-65.

Fama, E. F. and K. French, (1993). Common risk factors in the returns on bonds and stocks. *Journal of Financial Economics*, **33**, 3-56.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, **49**, 283-306.

Favero, C., M. Pagano, and E.-L. vonThadden (2008). How does liquidity affect government bond yields?. Centre for Economic Policy Research, Discussion Paper, No.6649.

Feldman, R., S. Govindaraj, J. Livnat and B. Segal (2008). The incremental information content of tone change in Management Discussion and Analysis. Available at SSRN: <http://ssrn.com/abstract=1126962>.

- Garcia, D. (2012). Sentiment during recessions. *Journal of Finance*, Forthcoming. Available at SSRN: <http://ssrn.com/abstract=1571101>.
- Gerlach, S. S. Alexander, and G. B. Guntram (2010). Banking and sovereign risk in the euro area. Deutsche Bank, Discussion Paper, No.09/2010.
- Geyer, A., Kossmeier, S. and S. Pichler (2004), Measuring systematic risk in EMU government yield spreads. *Review of Finance*, **8**, 171–197.
- Glosten, L., R. Jagannathan, and D. Runkle (1993): On the Relationship between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, **48**, 1779–1801.
- Gomez-Puig, M. (2006). The Impact of Monetary Union on EU-15 Sovereign Debt Yield Spreads. University of Barcelona, Working Paper in Economics 147.
- Greene, W. (2000). *Econometric Analysis*. New York: Prentice-Hall.
- Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review*, **70**, 393–408.
- Hales, J. W., J. Kuang, and S. Venkataraman (2010). Who believes the hype? an experimental examination of how language affects investor judgments. Available at SSRN: <http://ssrn.com/abstract=1667244>.
- Harris, L. (1986). Cross-security tests of the mixture of distributions hypothesis. *Journal of Financial and Quantitative Analysis*, **21(1)**, 39–46.
- Harris, R. D. F. and E. Tzavalis (1999). Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics*, **91**, 201–226.
- Hart, R. P. (2000). *DICTION 5.0: The text-analysis program*. Thousand Oaks, CA: Scolari/Sage Publications.
- Haugh, D., P. Ollivaud, D. Turner (2009). What drives sovereign risk premiums? An analysis of recent evidence from the euro area. OECD Economics Department, Working Paper No. 718.
- Hausman, J.A., (1978). Specification tests in econometrics, *Econometrica*, **46**, 1251–1271.
- Henry, E. (2008). Are investors influenced by the way earnings press releases are written?, *The Journal of Business Communication*, **45(4)**, 363–407.
- Henry, E. and A.J., Leone (2009). Measuring qualitative information in capital markets research. Available at SSRN: <http://ssrn.com/abstract=1470807>.

- Heppke-Falk, K. and F. Hüfner (2004). Expected budget deficits and interest rate swap spreads - evidence for France, Germany and Italy, Deutsche Bundesbank Discussion Paper No. 40/2004.
- Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance*, **56**, 1533-97.
- Hsiao C., Mountain DC, Ho-Ilman K. (1995). Bayesian integration of end-use metering and conditional demand analysis. *Journal of Business & Economic Statistics*, **13**, 315–326.
- Hsiao, C., (2003). Analysis of Panel Data. Cambridge University Press, Cambridge.
- Hsiao, C., (2007). Panel data analysis – advantages and challenges. Invited paper, University of Southern California.
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In Vol.1 of Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 221–233. Berkeley: University of California Press.
- Hutson, E., C. Kearney and M. Lynch (2008), Volume and skewness in international equity markets, *Journal of Banking and Finance*, **32**, 1255-1268.
- Im, K. S., M. H. Pesaran and Y. Shin (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, **115**, 53–74.
- Jankowitsch, R., H. Mosenbacher and S. Pichler (2006). Measuring the liquidity impact on EMU government bond prices. *European Journal of Finance*, **12**(2), 153–169.
- Jegadeesh, N. and A. D. Wu (2012). Word Power: A new approach for content analysis. AFA 2012 Chicago Meetings Paper. Available at SSRN: <http://ssrn.com/abstract=1787273>.
- Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, **48**, 65-91.
- Kahneman, D., and A. Tversky (1979). Prospect theory: an analysis of decisions under risk. *Econometrica*, **47**, 263-291.
- Keim, D. R. (1983). Size related anomalies and stock return seasonality: further empirical evidence. *Journal of Financial Economics*, **12**, 13-32.
- Klevmarken, N.A., 1989. Panel studies: What can we learn from them? Introduction. *European Economic Review*, **33**, 523–529.
- Kothari, S.P., X. Li, and J. E. Short (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a

- study using content analysis. *The Accounting Review*, **84**(5), 1639–1670.
- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. 2nd edition, Thousand Oaks, CA: Sage.
- Lee, C., A. Shleifer, and T. Thaler (1991). Investor sentiment and the closed-end fund puzzle, *Journal of Finance*, **46**, 75-109.
- Lehavy, R., F. Li, and K. Merkley (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *Accounting Review*, **86**, 1087-1115.
- Levin, A., C. F. Lin and C. Chu, (2002). Unit root test in panel data: Asymptotic and finite sample properties. *Journal of Econometrics*, **108**, 1–25.
- Li, F., (2006). Do stock market investors understand the risk sentiment of corporate annual reports? Available at SSRN: <http://ssrn.com/abstract=898181>.
- Li, F., (2010). The information content of forward-looking statements in corporate filings - A Naive Bayesian machine learning algorithm approach. *Journal of Accounting Research*, **48**, 1049-1102.
- Liang, Kung-Yee, and Scott L. Zeger (1986): Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, **73**, 13-22.
- Lo, A. and J. Wang (2006). Trading volume: Implications of an intertemporal capital asset pricing model. *Journal of Finance*, **61**(6), 2805-2840.
- Loughran, T., and B. McDonald (2011). When is a liability not a liability?. *Journal of Finance*, **66**, 35-65.
- Loughran, T., and B. McDonald (2013). Measuring Readability in Financial Disclosures. Available at SSRN: <http://ssrn.com/abstract=1920411>.
- Manganelli, S. and G. Wolswijk (2009). What drives spreads in the Euro area government bond market? *Economic Policy*, **24**(58), 191-240.
- Manning, C., and H. Schutze (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: The MIT Press.
- McCallum, A. (1996). Bow: A toolkit for statistical language modeling, text retrieval, classification and clustering. *Working paper*, School of Computer Science, Carnegie-MellonUniversity.
- Mitchell, M. L. and J. H. Mulherin (1994). The impact of public information on the stock market. *Journal of Finance*, **49**, 923-950.

- Mody, A. (2009). From Bear Stearns to Anglo Irish: How Eurozone sovereign spreads related to financial sector vulnerability. IMF Working Paper, No.09/108. Available at SSRN: <http://ssrn.com/abstract=1424844>.
- Nelson, D.(1991): Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, **59**, 347–370.
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.
- Pagano, M. and E.-L. VonThadden (2004). The European bond markets under EMU. *Oxford Review of Economic Policy*, **20**(4), 531–554.
- Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. Cambridge Working Papers in Economics, 0435, University of Cambridge.
- Rees, L. and B. Twedt (2012). Reading between the lines: an empirical examination of qualitative attributes of financial analysts' reports. *Journal of Accounting and Public Policy*, **31**, 1-21.
- Schuknecht, L., J. vonHagen, and G. Wolswijk (2010). Government bond risk premiums in the EU revisited the impact of the financial crisis. ECB Working Paper, No.1152. Available at SSRN: <http://ssrn.com/abstract=1540653>.
- Schwarz, K. (2009). Mind the gap: disentangling credit and liquidity in risk spreads. Working Paper, Department of Finance, The Wharton School, University of Pennsylvania.
- Sgherri, S. and E. Zoli (2009). Euro area sovereign risk during the crisis. IMF Working Paper, No.09/222. Available at SSRN: <http://ssrn.com/abstract=1490223>.
- Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends'. *American Economic Review*, **71**, 421-436
- Shleifer, A. (2000). *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford, UK: Oxford University Press.
- Sinha, N. R. (2010). Underreaction to News in the US Stock Market. Available at SSRN: <http://ssrn.com/abstract=1572614>.
- Spyrou S. (2011). Are sovereign bond yield spreads affected by investor sentiment? Evidence from the Greek budget crisis. Available at SSRN: <http://ssrn.com/abstract=1856943>.
- Stock, J. H. and M. W. Watson (2001). Vector autoregressions. *Journal of Economic Perspectives*, **15**(4), 101-115.
- Stone, P. J., D. C. Dunphy, M. S. Smith, and D. M. Ogilvie (1966). *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge: MIT Press.

Tauchen, G. E. and M. Pitts (1983). The price variability-volume relationship in speculative markets, *Econometrica*, **51**, 485-505.

Tetlock, P.C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, **62**, 1139-1168.

Tetlock, P.C., M. Saar-Tsechansky, and S. Macskassy (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, **63**, 1437-1467.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 17-38.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, **50**, 1-25.

Wright, S. E., and Budin, G. (2001). *Handbook of terminology management* (volumes 1 and 2). Amsterdam and Philadelphia: John Benjamins publishing company.

Zakoian, J.M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*. **18**, 931-955.