The Role of Media Content in

Explaining the Index Futures Market Behaviour

A thesis presented in partial fulfilment of the requirements for the degree

of

Doctor of Philosophy

in

Finance

At Massey University, Palmerston North,

New Zealand

Pok, Wei Fong

2015

ABSTRACT

This study examines sentiment from routine financial news and outlines the impact of the media content on three main index futures contracts of Hong Kong Exchanges and Clearing Limited, Bursa Malaysia and Singapore Exchange. The sample selection is based on high a percentage of English usage, thus enabling cross-country comparison. I generate several news factors from routine financial news and find that factors that represent pessimistic market sentiment are more prevalent. Highly pessimistic news factors (*Pessimism*, *Negative* and *Weak*) predict lower returns on the same day, while the negative impact will reverse within the next five days. The finding is consistent with the noise trader theory that investors initially overreact to negative market news and drive the price lower, before the price corrects itself. Since investors are only obligated to pay maintenance margin when trading futures contracts, the trading strategy based on this finding becomes more economically significant compared to paying full price in the spot market. The second part of this thesis examines the role of sentiment in predicting the mean-variance relationship. I argue that sentiment will affect returns volatility. Risk-averse rational investors require higher returns for holding risky assets, which implies positive mean-variance relationships. During a high sentiment period, noise traders dominate the trading activities, leading to a weaker mean-variance trade off. Evidence points to negative mean-variance relationships during a high sentiment period, but the results are mixed during a low sentiment period. The finding suggests that sentiment can be incorporated into the index futures pricing model, through its interaction with returns volatility.

ACKNOWLEDGEMENTS

Seeing much, suffering much, and studying much, are the three pillars of learning.

—Benjamin Disraeli

Researching and writing this thesis have not been an easy ride. I am grateful to have had remarkable people around, who raised me up when I was lacking self-confidence to complete the journey.

On the top of the list are my supervisors, Professor Martin Young and Dr. Md Humayun Kabir. I would like to express my greatest gratitude for the time you have devoted and the constructive comments given. Thank you for patiently guiding me through the process.

The School of Economics and Finance has provided constant support. Associate Professor Ben Marshall generously assisted in solving data issues. Cameron Rhodes helped to set up my work station and solved several software issues. Fong Mei Chin introduced very useful databases and taught me how to operate them. Maryke Bublitz made the paperwork easy for me. I gratefully appreciate the assistance, which helped me to move forward.

I am grateful to Universiti Tunku Abdul Rahman for granting financial support and study leave to complete my study. I will make myself a greater asset to the university.

Also, I am deeply indebted to Peir Peir and Phua, See Yee and Wei Seng, Maggie Zou and her mother. They have given me a home in New Zealand and see me as one of the family. Thanks to Eric who always lend an empathetic ear when I am down. Special thanks goes to Tuck, who spent three years reading, analyzing and writing together with me. We had food and sport therapy to keep our sanity.

To my friends, I am blessed to have your friendship. The joy and laughter that you bring mean a lot to me.

Although they will not enjoy reading it, I dedicate this thesis to
my parents Pok Kim Seong and Wong Kuan Thai
grandparents Pok Tom Seng, Hew Ting Hiong and Lee Yu Mei
siblings Wai Kin, Wei Ling, Tim, Wei Keong, Maggie and Wei Lan
for their constant support and unconditional love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 : INTRODUCTION

1.1 INTRODUCTION

This thesis investigates the role of media content extracted from routine financial news in explaining the index futures market behaviour. I extract news sentiment factors to predict the stock index futures returns. I also create news sentiment regimes to investigate the index futures returns mean-variance relationship. This study adds to the literature of price formation process by investigating the role of news content as information and non-information and its relationships with trading activities of stock index futures. The trading strategy, which is designed base on the findings, is more likely to be economically significant because the average trading costs of trading stock index futures is much lower than trading in the cash market.

Section 1.2 provides the background and motivation of the study. Investors use financial news as a source of reference to make investment decisions. Market participants are likely to misinterpret the news, yet survive in the long-run and cause the asset prices to deviate from its fundamentals. Section 1.3 discusses the delimitations and the scope of study and the key assumptions that justifies the research design. Section 1.4 and 1.5 explain the significance and contributions of the research.

1.2 BACKGROUND AND MOTIVATION

Newspapers have evolved over time since their first introduction. They contain news from both local and international scenes, sports, entertainment, fashion and classifieds as part of the commonly featured sections. Investors look for information from the investment section.

From the perspective of financial studies, information can be divided into two types: technical and fundamental. Technical information ranges from the simplest price and trading volume to the more sophisticated statistical models. Fundamental information encompasses news that affects the valuation of a financial asset, for example, news on a firm's R&D, project's progression, state of economy and political stability.

This has prompted some researchers to investigate the impact of news on the market event. Shiller (2005) devotes a whole chapter in his book to discuss this diligently. He wrote:

"…news stories rarely have a simple, predictable effect on the market events. Indeed, in some respects, they have less impact than is commonly believed. However, a careful analysis reveals that the news media do play an important role both in setting the stage for market moves and in instigating the moves themselves."(Shiller, 2005, p. 105)

He further explains the role and process of news media in gathering public's attention and stimulating their reaction in detail:

"….news media are fundamental propagators of speculative price movements through their efforts to make news interesting to their audience. They sometimes strive to enhance such interest by attaching news stories to price movements that the public has already observed, thereby enhancing the salience of these movements and focusing greater attention on them. Or they may remind the public of past market episodes, or of the likely trading strategies of others. Thus the media can sometimes foster stronger feedback from past price changes to further price changes, and they can also foster another sequence of events, referred to here as an attention cascade. This is not to say that the news media are a monolithic force pushing ideas onto a purely passive audience. The media represent a channel for mass communication

and the interpretation of popular culture, but popular culture has an inherent logic and process of its own." (Shiller, 2005, p. 105)

These opinions are supported by empirical evidence. In a survey conducted by Oberlechner and Hocking (2004), traders were given a questionnaire to rate the importance of the information sources. The scale of 1 denotes very important and 4 denotes unimportant. Traders rate wire services as the most important information source (Mean=1.48, σ_M = 0.05), while daily newspapers are considered somewhat important (Mean=2.17, σ_M = 0.1).

Financial columns consistently report what has happened in the stock markets. However, not all investors are able to interpret the news correctly, update beliefs accordingly and make the right investment decision. They are called noise traders. Classical asset pricing models safely ignore the impact of irrational noise traders on asset pricing based on two arguments. The aggregation argument, which, suggests that investors' expectations are random, so the impact of over-expectation and under-expectation should be cancelled out at the aggregate level. The arbitrage argument can be traced back to Friedman (1953) who claims that irrational traders cannot survive in the long run. Irrational traders buy when the prices are too high and sell when the prices are too low, eventually losing all their wealth in the long-run. At the same time, arbitrageurs trade against this irrationality, and offset the price impact of irrational trades.

These two arguments are duly challenged in literature. The theory of noise traders suggests that rational agents who face a finite time horizon are concerned about fundamental risks and noise trader risks that their irrational beliefs last for too long or are even further aggravated. There are limits to arbitrage making rational traders reluctant to take a large a position against noise traders (Campbell & Kyle, 1993; DeLong, Shleifer, Summers, &

Waldmann, 1990b(DSSW hereafter); Figlewski, 1984; Shiller, 1981; Shiller, Fischer, & Friedman, 1984). Consequently, the asset prices deviate significantly from the fundamentals. In addition, informed rational traders can take advantage of the positive feedback traders who chase the market trend. Rational traders buy on good news, with the anticipation of selling to positive feedback traders when the price is higher than warranted by the fundamentals (DeLong, Shleifer, Summers, & Waldmann, 1990a)

Motivated by Tetlock's (2007) investigation on the relationship among media pessimism, aggregate stock market returns and trading volume, this study aims to address the following issue: ***What is the role of routine media content in explaining the daily index futures market behaviour?*** Thus, the interaction among media content and stock index futures markets will be examined. The first part of this thesis attempts to capture the sentiment from the routine news and outlines the impact of media content on returns and trading volume in the index futures markets.

The second part of this thesis focuses on the role of the news sentiment on index futures returns mean-variance relationship. This is motivated by the empirical evidences on deviations between the theoretical futures prices estimated by futures pricing model from the actual futures prices. The Cost of Carry Pricing Model is widely used to estimate the theoretical price of stock index futures. The model assumes that riskless borrowing can be obtained at a constant rate and continuous dividends are paid constantly in a frictionless market.

The persistence of mispricing implies the inadequate specification of pricing models or the futures market itself is inefficient[1]. Researchers have attributed the mispricing to short-sales restrictions (Fung & Draper, 1999; Gay & Jung, 1999), taxes (Cornell and French, 1983 a, b; as cited in Cornell, 1985) and transaction costs (Brailsford & Cusack, 1997). Mispricing and inefficient markets accelerate arbitrage activities, hence intensifying market volatility. Fung and Patterson (1999), Gay and Jung (1999), and Yadav and Pope(1994) find that mispricing of index futures is positively correlated with index volatility. Hence, there is ground to believe that incorporating volatility in the pricing model can mitigate the mispricing problem. In financial studies, volatility usually refers to the difference between expected return and actual return; the greater the difference, the greater the uncertainty faced by investors. Standard deviation or variance of an asset's returns over a certain period, together with its probability distribution, is usually used to quantify the risk faced by investors (Poon & Granger, 2003).

Investment risk is commonly considered in asset pricing models because investors require compensation for bearing risky assets. Market participants face various risks, namely liquidity risk, default risk, interest rate risk, currency risk, inflation risk, operational risk and political risk. Volatility as a general measure of risk, however, is not taken into account in the cost of carry pricing model. The costs of carry pricing model's underlying assumptions imply that returns volatility should have no impact for futures prices.

[1] The cost of carry pricing model is built upon the no-arbitrage argument. According to Sutcliff (2006), the underlying assumptions of no-arbitrage conditions are no marking to market, single and constant riskless interest rate, no taxes, certain dividends, no transaction costs, no initial margin, no continuous compounding, holding a portfolio that replicates the underlying index, not restriction for short-selling, simultaneous trading is feasible, no delivery price risk, no changes in the definition of the index, perfect asset divisibility, no price effect except in aggregate, shares are paid for immediately, capital gain and loss are paid at liquidation, no risk of default, common currency for shares and multiplier, zero value of voting and other rights, immoral investments, costless storage, and independence of the index. The author concludes that " …there are no assumptions which are invalid, cannot be relaxed by using as more general no-arbitrage condition, and where violations of the assumptions have an important effect".

Alternatively, Hemler and Longstaff (1991) propose a Closed-form General Equilibrium Pricing Model, which explicitly models the stochastic market volatility and stochastic interest rates[2]. In the context of the Nikkei 225, the Hang Seng, the KOSPI 200, the TAIFEX and the SGX Taiwan stock index futures, this model outperforms the cost of carry pricing model (Wang, 2007, 2009). Wang (2009) compares five volatility estimators and concludes that the forecasting performance of the general equilibrium pricing model is affected by the volatility estimates.

In summary, these studies suggest that volatility can be used as a parameter in the index futures pricing model and there is a positive mean-variance relationship provided traders have rational expectations. However, the empirical investigations on the mean-variance relationship yield mixed results. Several explanations are proposed, including model specification, sampling period, sampling frequency and the selection of conditional variables. However, no consensus has emerged. Other than these methodological issues, there is evidence showing a weaker or even negative mean-variance relationship because traders' risk-aversion can be influenced by biased beliefs (Yu & Yuan, 2011).

Motivated by the mixed findings, I look into sensitivity of the mean-variance to different model specifications of variance and sampling frequency. This study attempts to address the question of: ***What is the role of sentiment in explaining the index futures' mean-variance relationship?*** I propose that the investor sentiment weakens the mean-variance relationship when the sentiment is high, based on the assumption that noise traders who have poor timing dominate the market in a high sentiment period.

---

[2] The Hemler-Longstaff model can be written as:

$$L_t = \alpha + \beta r_t + \gamma V_t + \xi_t$$

Where $L_t = \log(F_t^a / S_t)$; F_t^a is the theoretical futures price; S_t is the stock price; r_t is the risk free rates to maturity; V_t is the level of volatility.

A large body of literature examines the sentiment-return relationship (see section 2.5) or risk-return relationship (see section 2.6.4) separately. These studies documented how sentiment affects returns but we have little knowledge about the inherent mechanism that underlies the relationship among these variables. This study attempts to create a link among these three variables, specifically: how sentiment affects return by altering the mean-variance relationship. Following the DSSW noise trader framework, Lee, Jiang, and Indro (2002) uses GARCH-in-mean model to test the impact of investor sentiment on conditional volatility and excess return, covers the DJIA, NASDAQ and S&P500 from 1973 to 1995. The role of investor sentiment is modelled with the variance equation, which in turn is used to predict the excess return in the mean equation. Yu and Yuan (2011) let investor sentiment to interact with the return variance, to predict excess return of NYSE-Amex portfolio from 1963 to 2004. Extending the ideas from these papers, I focus on three Asian index futures contracts, namely the Hang Seng Index Futures, the Kuala Lumpur Stock Index Futures and the MSCI Singapore Free Index Futures from 1996 to 2008. In addition, Yu and Yuan (2011) examine the returns mean-variance relationship in the high-sentiment regime. This study examines both high-sentiment regime and low-sentiment regime.

1.3 DELIMITATIONS OF SCOPE OF THE STUDY

This is a cross-country analysis of news media content's impact on investor sentiment and three Asian index futures markets, namely Hong Kong, Malaysia and Singapore. As former colonies of the British, their large English-speaking population enabled fast dissemination of market information. Shiller (2005) explains this in his book:

> *"One of the reasons why U.S. stocks market appears to have a disproportionate effect on markets of other countries is the United States uses the English language, which has emerged as a world language. It is much easier for foreign reporter, who invariability*

knows English, to respond to stories from the United States or the United Kingdom than to stories from Germany or Brazil. Producing news stories is a business with tight deadlines. And it requires fast actions. A lot of reporters have the ability to pick up a story from another country in English, and turn it into a local story in pinch." (Shiller, 2005, p. 104).

Cross-country analysis of media content based on one language is important to enable meaningful cross-country comparison. This property ascertains consistency and reliability of the measurement when different word count software and word categories dictionaries are used. However, this limits the scope of the study.

Hong Kong's official languages are Chinese and English. Before 1974, English was the sole official language. The amendment of Official Language Ordinance in 1987 makes Chinese and English language as the languages for legislation. As part of the preparation to handover Hong Kong to China in 1997, the Hong Kong Basic Law declared English as a co-official language with Chinese. According to the Hong Kong population by-census main report in 1996, 200,000 people indicated English as their first language while 2,300,000 people reported English as their second language. Hong Kong recorded a 35.9% English speaking population in that year.

Malaysia's official language is the Malay language that is used in government functions and is the medium of instruction in schools. However, 32% of the population is English speaking (Crystal, 1995, p. 109). The Cambridge Encyclopaedia of the English Language tabulates this figure based on the Malaysian total population in 1990. Nair-Venugopal (2001) stated that "nowhere is the use of English more entrenched in Malaysia than in the private sector domains of corporate business and industry, banking and finance." The globalisation forces have pressured the Malaysian government to reverse the policy in order to meet the demand for higher proficiency in English. Now English has become the main medium of instruction in colleges and universities besides the Malay language.

Singapore has four official languages: English, Mandarin, Malay and Tamil. Singapore's national anthem is written in Malay, but all the four languages are used in schools, government offices and courts. However, proceedings are written in English only. Based on the population census in 2000, 665,087 of the population in Singapore use English as their first language, and 3,257,906 or approximately 71% of the population is English speaking.

English is widely used as a medium to disseminate market information in Hong Kong, Malaysia and Singapore, resulting from historical and current policies. Well established and tested word count software developed in English warrant the common use of English in the sample countries to ensure reliability of measurement of media content. This is the main reason used to exclude other developed markets (Japan and Korea) and emerging markets (China, Thailand and India) from this study.

The large English speaking population in these counties has contributed to a high-level of readership of English newspaper. The Nielsen Media Index survey in 2008[3] reported New Straits Times (NST) ranking second place after The Star, with readership of 2% over a 14.29 million populations. However, in an earlier survey conducted by Synovate in 2005 (as cited in Davis, 2005), half of NST readers were found to be white-collar workers, of which 60% were professionals, managers and businesspersons. Also, the majority of NST readers were males and above 30 years old (Davis, 2005). These reader profiles make it an appropriate resource to collect news and study the role of news media in index futures trading. Singapore AC Nielsen 2008 Media Index survey (Chua, 2008) reported that the Straits Times (ST) newspaper lead by

[3] Information obtained from Nielsen Malaysia, through http://my.nielsen.com/news/20081023.shtml.

capturing 1.44 million readers, which is 39% over the readership base. In Hong Kong, the leading paid English newspaper, South China Morning Post (SCMP), recorded an average net circulation of 118,622 per effective publishing day in January 2008[4]. A free newspaper, The Standard (TS), dethroned South China Morning Post from its number one position. The Standard's executive editor, Steve Shellum asserted that its readership grew from 50,000 to 250,000 in IFRA Publish Asia conference in Macau on April 2, 2008 (Thomascrampton, 2008). I collect news from SCMP because TS is unable to provide news for the entire sampling period, from 1996 to 2008.

I extract news sentiment from SCMP (Hong Kong), NST (Malaysia) and ST (Singapore) based on its' readership and credibility. The reliability of news sentiment can be tested in different market environments in light of its relationship with index futures returns, volume and volatility.

1.4 SIGNIFICANCE OF THE STUDY

Qualitative studies on index futures are relatively limited as compared to those conducted on spot markets. The study based on index futures market is meaningful for several reasons.

First, the average transaction cost in futures market is much lower than the cost incurred in spot market. Investors are required to maintain a margin rather than paying the full nominal amount stated in futures contracts. If there is any relationship existing between sentiment, returns and volatility, the trading strategy derived from such relationships is more likely to be economically significant. As shown in Table 1.1, the stock index futures becomes

[4] Information obtained from Audit Bureau of Circulation, Hong Kong, through http://www.hkabc.com.hk/en/index.htm

increasingly active due to its relatively low transaction cost as compared to share trading in spot markets.

Table 1.1 Summary of Stock and Index Futures Trading Value for Selected Asia-Pacific Exchanges in 2012
This table summarises the stock and index futures trading value in millions of USD. The data is excerpted from World Federation of Exchanges 2012 Annual Report. The trading of stocks and index futures is carried out by the same exchange for most of the markets except for China and Taiwan.

Derivatives Exchange	Index Futures Trading Value (USD millions)	Stocks Trading Value (USD millions)	Stock Exchange
Australian Securities Exchange	1,099,110	936,584	Australian Securities Exchange
Bombay Stock Exchange Limited	42,680	109,459	Bombay Stock Exchange Limited
Bursa Malaysia Derivatives	53,888	124,246	Bursa Malaysia Derivatives
China Financial Futures Exchange	12,034,382	4,970,009	Shanghai Stock Exchange & Shenzen Stock Exchange
Hong Kong Exchanges and Clearing	4,027,771	1,106,086	Hong Kong Exchanges and Clearing
Osaka Securities Exchange	3,717,203	142,905	Osaka Securities Exchange
Tokyo Stock Exchange	1,478,459	3,471,884	Tokyo Stock Exchange
Korea Exchange	7,086,514	1,517,496	Korea Exchange
National Stock Exchange of India	530,206	522,396	National Stock Exchange of India
Singapore Exchange	NA	255,929	Singapore Exchange
Taiwan Futures Exchange	1,525,454	678,210	Taiwan Stock Exchange Corp
total	31,595,667	12,992,204	

Source: http://www.world-exchanges.org/files/statistics/pdf/WFE2012%20final.pdf

Second, futures traders are often better trained than retail traders in spot market. Although the average transaction cost is lower in futures trading, the initial outlay to trade a contract is relatively large for retail traders. The initial margin is HKD90,450, RM5,800 and SGD4,750 for Hang Seng Index Futures (HSIF), Kuala Lumpur Composite Index Futures (KLCIF) and Morgan Stanley Singapore Free Index Futures (SiMSCIF) respectively. The amount of final settlement is large, given the contract multiplier is HKD50 per index point for SHIF, RM50 per index point for KLCIF and SGD200 for SiMSCIF. The top ten HKEX participants hold 87.68% long open interest and 78.37% of short interest of Hang Seng Index futures from 30th March to 3rd April 2009. The majority of players in the market are professionally trained institutional traders who are better informed as compared to retail traders. This study indirectly investigates the reactions of the informed-traders to news content that could be used as a proxy for investor sentiment.

Third, derivatives market provides a unique avenue for hedging activities to take place, which cannot be accomplished in spot market. Wang (2008) suggests that Asian stock markets are rather volatile. The futures prices are highly correlated with spot prices. Taking the cost of carry into account, the adjusted spot price to be delivered at a future date becomes the futures price. Under this circumstance, equity index-based derivatives play the role of cost effective risk management instruments. The study of index futures will in turn shed light on spot market price behaviour.

Fourth, Sutcliffe (2006) reviews literature on lead-lag relationship between spot and futures, and concludes that returns of index futures lead returns of spot index by a few minutes. There is not much evidence pointing to leads or lags for a day. The reviewed studies tend to understate the contemporaneous relationship between spots and futures prices. These markets actually react to relevant information simultaneously. By predicting returns of index futures using news factors, the lead-lag relationship between spot and futures also implies predictability of cash index's returns. Adding on to Tetlock (2007), this study includes contemporaneous relationships between news content and trading activities.

1.5 CONTRIBUTIONS OF THE RESEARCH

Building on the method of Tetlock (2007), this study attempts to examine whether news content explains trading activity consistently in markets of different structures. Investors' preference could change with the arrivals of information. Although we can compile high-frequency tick-by-tick data, we are unable to measures investor sentiment within minutes or seconds before the transaction. Therefore, as far as possible, a daily gauge of sentiment is preferred. A carefully planned survey yields more reliable and less biased results, but conducting a daily sentiment survey is beyond the capacity of most traders and financial research institutions. There are limitations that are costly to overcome.

First, is the trade-off between the timeliness of a survey and its cost. For example, the surveys by American Association of Individual Investor (AAII) and Investor Intelligence (II) are both conducted on a weekly basis. Information in these surveys has already turned "stale" before it is published. Alternatively, daily routine financial analysis in newspaper is one of the possible cost and time effective source to measure investor sentiment in view of its large readership, and readers take it as a source of information to make investment decision.

Second, survey responses are subject to bias. A survey by the AAII categorises individual investors' responses in the weekly survey as bullish, bearish or neutral. The same categories are also used in the Investor Intelligence sentiment index after reviewing 150 market newsletters. These surveys tend to suffer from loss of information and inconsistencies may happen in the process of classifying opinions into three strict categories, especially when the opinions are mixed. Trochim (2001) explains the advantages of content analysis over questionnaire survey, where respondents tend to alter their behaviour to gain a good impression. In addition, bias may arise when respondents interpret the questions differently. Poorly designed questionnaires, which annoy respondents, will produce similar outcomes.

Content analysis has advantage over survey because it is indirect, without the intrusion of interviewer or survey instrument, thus lowering respondent bias.

Third, unavailability of survey results set in a domestic context. Previous studies use survey-based sentiment measures from US that are readily available, for example II and AAII. Evidence show that US stock markets do affect markets in Asia; however, a sentiment measure that can directly gauge the investor sentiment in a particular market is more relevant. Internationale Nederlanden Groep (ING)[5] conducts the ING Investor Sentiment Dashboard survey for 13 Asia Pacific countries including China, Hong Kong, India, Indonesia, Korea, Malaysia, Philippines, Singapore, Taiwan, Thailand, Japan, Australia and New Zealand. Investor sentiment scores ranging from 0 to 200, the higher the score, the more optimistic investors are in a particular market. Surveys are conducted quarterly by face-to-face interview or internet survey. There are 1,343 respondents in total, from 13 countries for the 2008 fourth quarter survey. A quarterly survey is less informative than daily measures. This study adopts a replicable procedure to derive current and backdated daily investor sentiment, provided that the news archive is available.

This study contributes to the existing body of knowledge by examining the reliability of a high frequency sentiment measure. News content analysis that is set in a domestic context enables a more comprehensive exploration of investor sentiment's impact on index futures returns and returns volatility. This study adds to the literature on price formation processes by investigating the role of news content as information and non-information and its relationship with trading activities of index futures.

[5] http://www.ingnz.com/WEB/website.nsf/content/press+releases/$FILE/ING%20Investor%20Sentiment%20Index%20FINAL_613030.pdf

The contribution of the mean-variance relationship analysis is two-fold. First, based on rational expectation theory, investors require a higher return for holding riskier assets. However, the empirical evidence on the risk-return relationship is mixed, and extant research is yet to find a conclusive explanation for these observed mean-variance patterns. Sentiment is one of the possible explanations. Second, volatility is not explicitly incorporated into the index futures pricing model[6]. If the relationship between volatility, returns, and sentiment can be established, sentiment can be incorporated into the pricing model indirectly and eventually improve its forecasting power.

I find that during high sentiment periods, noise traders dominate the trading activities, leading to a weaker mean-variance trade off. Evidence points to a negative mean-variance relationship during high sentiment periods, but the results are mixed during the low sentiment period. This is consistent with the theory of noise traders dominating the market during the high sentiment period. The findings have implications for the volatility estimates and the selection of the pricing model. Sentiment can be incorporated into the index futures pricing model, through its interaction with return volatility.

1.6 ORGANISATION OF THE THESIS

I review the literature with respect to varies proxies of investor sentiment, theories and evidences of the investor sentiment on assets returns and trading volume, theories and evidences of the role of investor sentiment on returns mean-variance relationship in Chapter 2. Chapter 3 outlines the hypothesis development, model specification, data and sample. Chapter 4 elaborates on how to generate sentiment news factors and examine their consistency over time. A structural vector autoregressive model is used to examine the

[6] See Brailsford and Cusack (1997) for comparison of three futures pricing models, namely cost of carry model, Ramaswamy-Sundaresan model and Hemler-Longstaff model.

relationship among index futures returns, news sentiment factors and trading volume. Chapter 5 presents the analysis on the role of sentiment in moderating the mean-variance relationship. Chapter 6 concludes the study.

CHAPTER 2 : LITERATURE REVIEW

2.1 INTRODUCTION

There are extensive surveys on behavioural finance studies that offer comprehensive summaries of how limit to arbitrage and cognitive biases lead to trading behaviour and pricing pattern that are inconsistent with the Subjective Expected Utility notion. Shefrin (2000) illustrates how investors commit costly mistakes in their investment decisions, while Hirshleifer (2001) reviews psychology-based asset pricing theories. Shiller (2002) discusses the evolution of behavioural studies in finance and focuses on the discussion of feedback models and limits to arbitrage. Ritter (2003) provides discussions on cognitive biases and illustrates how these biases are related to inflation and under-pricing of IPOs. Barberis and Thaler (2003) conduct a very comprehensive survey of behavioural finance. While reviewing a number of behavioural models[7] that explain the financial anomalies[8], my aim, in this chapter, is also to synthesise the extant measures of investor sentiment to establish empirical relationship among observed variables and sentiment measures in relationship to the theoretical models' prediction and finally to examine the current stand of investor sentiment studies.

Section 2.2 in this chapter discusses behavioural theories and models that explain investor behaviour, reviews the definitions of investor sentiment and outlines the definition in this research context. Section 2.3 elaborates the commonly used measures of investor sentiment in spots and derivatives markets. Section 2.4 discusses the theory and evidence of investor sentiment on assets returns and trading volume. Section 2.5 reviews the theory and

[7] The paper classifies the model as in Barberis, Shleifer and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998),De Long, Shleifer, Summers and Waldman(1990b), , Hong and Stein (1999) as belief-based models; Barberis, Huang and Santos (2001) and Barberis and Huang(2001) as preferences-based models.

[8] The size premium, long-term returns reversals, predictive power of scaled-price ratios, price momentum, closed-end fund discounts, comovements of returns among assets, post-earning announcement drift, share repurchase effect and primary and secondary offerings effects.

evidence of the role of investor sentiment on the returns mean-variance relationship. Section 2.6 concludes.

2.2 INVESTOR SENTIMENT AND BEHAVIOURAL FINANCE

2.2.1 Investor sentiment definition and rationale

Behavioural finance complements the traditional finance framework by relaxing the assumption of agents' *rationality* and proving that the *limits to arbitrage* enable the survival of sentiment traders. Traditional finance theories assume that the price impact of irrational agents only lasts for a short period, because rational traders' arbitrage activities eventually lead to price correction. However, behavioural financial models predict that mispricing can be persistent due to limits to arbitrage[9].

Barberis and Thaler (2003) highlight two important assumptions of rationality. First, economic agents update their beliefs as stated by Bayes' law as they receive new information. Second, agents make decisions conditional upon their beliefs, and the decisions made are supposed to comply with the theory of subjective expected utility. There are two competing theories that can be used to explain financial anomalies (Brav & Heaton, 2002). The "rational structural uncertainty" theories relax the first assumption and the "behavioural" theories relax the second assumption. The concept of bounded rationality underlies the behavioural theories. Dumas, Kurshev and Uppal (2009) define sentiment as "fluctuations in the probability beliefs of overconfident agents relative to agents with the proper beliefs". Investor sentiment is the beliefs formed based on heuristics rather than Bayesian rationality. Such heuristic biases of sentiment driven investors or noise traders tend to over-price or under-price assets, deviating from the intrinsic value warranted by fundamentals.

[9] This including (but not limited to) transaction cost, short-sale constraints, risk-aversion arbitrageurs and agency consideration.

Classical financial theories argue that irrational investors have no important impact on price formation because arbitraging activities by rational investors offsets the influence of those irrationals. As irrational investors bid prices above fair values, rational arbitrageurs grab the profit opportunities by short selling the overvalued assets. This process brings the assets values back to their intrinsic values. In the real word, limits to arbitrage do exist and hinders rational investors from betting against irrational investors.

Psychological theories may lend an explanation to why many investors make investment decisions that cannot be justified by fundamentals. Cognitive psychology explains how heuristics lead to systematic biases when people form beliefs. Tversky and Kahneman (1974) introduce three heuristics: representativeness, availability, and anchoring-adjustment. Representativeness leads to sample size neglect. Noise traders believe that the gains in a few days in a row can represent gains in a longer period. Availability bias suggests that recent event will gain more weight when assigning probability for an event to happen. This implies that investors tend to be more confident that recent trend will repeat or persist. In psychological experiments, subjects tend to be overly optimistic, believing themselves to have above-the-average chance to experience positive events and below-the-average chance to experience negative event relative to others (Weinstein, 1980). Investors also tend to extrapolate time series by exponential smoothing method (Andreassen & Kraus, 1990) and they believe in trend continuation.

2.2.2 News, sentiment and behavioural models

DeLong, Shleifer, Summers, and Waldmann (1990b) argue that rational arbitrageurs are risk-averse and fundamental risk is one of the factors limiting arbitraging activities. On the other hand, the unpredictability of irrational investors' beliefs creates "noise trader risk". It is hard to predict when irrational investors' beliefs revert to the fundamentals and whether their beliefs become more extreme or long lasting. Arbitrageurs need to pay interest for the resources they borrow to implement the arbitrage trading strategies. At times, arbitrageurs are forced to liquidate their position before the prices approach the fair values[10].

Rational arbitrageurs understand that assets are over-valued when a noise trader's sentiment is optimistic. Since limits to arbitrage make arbitrage riskier and less attractive, arbitrageurs make rational decision to buy instead of short selling the overvalued assets. Demand from arbitrageurs pushes the prices higher and whet the appetite of noise traders to chase the trend. Arbitrageurs' buying spree triggers feedback trading and prices display momentum to an even higher level. Just before the bubbles are large enough to make noise traders realise that they are actually paying too much for the assets, arbitrageurs would have sold the assets in hand, and finally the prices start to adjust gradually to the levels warranted by the fundamentals. The positive feedback models proposed by DeLong, Shleifer, Summers, and Waldmann (1990a) predict that the overreactions to news trigger positive feedback.

If noise traders consistently lose money to arbitrageurs, no matter if arbitrageurs bet against or follow the trend, why do noise traders not eventually lose all their money and be forced to exit the market? Why do noise traders do not learn from their mistakes, and become arbitrageurs instead? DeLong et al. (1990a) and Shleifer and Summers (1990) give answers to these questions. First, in the case of arbitrageurs, they are reluctant to bear the noise traders' risk and fundamental risks. Noise traders are actually being rewarded with higher expected returns for bearing these risks, even though noise traders buy at high prices as they are selling at even higher prices. Second, learning is difficult for noise traders because every episode of trend chasing and price bubbles are not exactly the same. The experience from one bubble

[10] Moreover, riskless arbitrage opportunities might not exist due to the availability of perfect substitutes. In addition, transaction costs, short sell constraints, and agency considerations further limit the feasibility of arbitrage activities.

may not be relevant enabling noise traders to exit at the right time during the next bubble. Third, the time gap between bubbles is long enough for inexperienced noise traders to join the market while the experienced noise traders decide to quit. Fourth, noise traders who lose in the last episode of trend chasing may have saved enough money and return to be active again in the feedback trading.

In real life, not all investors are smart enough to distinguish real information from noise. It is likely that they receive information but interpret it with bias, or receive non-information and bet on it as if it is real information. Recommendations made by brokers or so-called financial experts (Shleifer & Summers, 1990) and trend chasing strategies (DeLong et al., 1990a) are among the examples. The initial price changes may be due to changes in fundamentals, the gains made in the early stage of the trend that feeds the appetite of noise traders, as they are overconfident that the trend will persist. Optimistic sentiment motivates noise traders to chase the trend, buy when the prices increase and sell when the prices decrease. The market remains untouched if there are only a few irrational traders acting in the market.

News is one of the major channels of information dissemination. The content of news varies. Macroeconomic news such as interest rates and foreign exchange policies affect securities prices by changing the forecasts of the market. Company-specific news forms the expectation on future dividends and its growth rate. Generally, changes in securities prices exhibit momentum in the short run and reversal in the longer run. There are theories based on different assumptions trying to explain this stylised fact. Positive feedback theory proposed by DeLong et al. (1990a) and the overconfidence model by Daniel, Hirshleifer, and Subrahmanyam (1998) are built on the assumption that prices overreact to news. Prices overshoot initially and the overconfident investors chase the trend. The direction of price changes continues for a while before it reverses in the long-run. However, another strand of behavioural models assumes investors to underreact to news. In the model of Barberis, Shleifer, and Vishny (1998), investors do not update their beliefs accordingly upon the arrival of new information due to representative and conservatism bias. Prices adjust too slowly and this causes the prices to move in the same direction in consecutive days. In the Hong and Stein (1999) model, news does not arrive at the same time. Every "news watcher" observes different pieces of private information and prices underreact to diffused news. In this situation, trend-chasing strategies by momentum traders can be profitable. However, in the long-run, positive feedback trading leads to overreaction. Even though these behavioural models start from different positions (under- or overreaction), these models still reach the same conclusion that prices tend to exhibit momentum in the short horizon and reversals in the longer horizon.

De Long et al. (1990b) popularise the concept of the "hold more" effect, "price pressure" effect, "Friedman" effect and "create space" effect, and bridges the link between investor sentiment and stock returns. As bullish irrational investors hold more risky assets than rational arbitrageurs do during the period of high sentiment, higher expected returns is considered as a form of reward for noise traders bearing the risk. If this is true, we should expect a positive relationship between sentiment and returns. At the same time, bullish noise traders demand more stocks. Greater demand drives the stock prices to higher levels and expected returns are lowered consequently. Hence, expected returns are negatively related to investor sentiment. These two effects interact to determine the impact of sentiment on returns. If the hold-more effect dominates the price pressure effect, one would expect that bullish sentiment to lead to higher returns. When the price pressure effect outweighs hold-more effect, bullish sentiment predicts lower returns. This explains the short run impact of the shift in sentiment to excess returns.

Conversely, the Friedman effect and create space effect are long-run concepts and attribute to variability in noise trader beliefs. A greater shift in sentiment is associated with greater future returns volatility (higher risk) and lower expected returns. Investors are advised to buy at low prices and sell at high prices. Unfortunately, noise traders have poor market timing, follow the footsteps of other noise traders, and end up buy-high-sell-low, thus earning poor returns. This is called the Friedman effect. In this case, noise traders' returns are negatively related to the variability of their beliefs. Risk-averse arbitrageurs avoid betting on noise traders' mispricing when there is high variability in noise trader beliefs. This is the so-called "create space" effect. A large enough create-space effect offsets the negative impact of the price pressure effect and the Friedman effect on returns.

2.3 MEASURES OF INVESTOR SENTIMENT

Researchers exploring alternative measures of investor sentiment have helped to improve the understanding of how investor sentiment relates to asset valuation. Investor sentiment is a latent concept that is not directly observable but explainable using cognitive psychology. There are two broad categories of investor sentiment proxies.

First, subjective or "direct" measures of sentiment are obtained through survey. Survey instruments are used to obtain investors' or advisors' direct point of view about the movements of the market or investment recommendation. The opinion is subjective, not everyone will have exactly the same forecast towards the same market event. Sentiment measures derived from media content discussed earlier fall into this category.

Second, objective or "indirect" measures that are available after the investor sentiment is reflected in trading activities. This includes retail investor trades, mutual fund flows, trading volume, dividend premium, closed-end fund discount, option implied volatility, IPO first day return, IPO volume and equity issues.

2.3.1 Subjective measures of investor sentiment for the derivatives market

Sources to gauge investor sentiment for the derivatives markets are similar to those in the spot markets[11], for example public media, newsletters, and surveys. The Market Vane collects buy or sell recommendations made by market advisers and derives monthly forecast of the commodity market returns. It is viewed as being a bullishness indicator. Brown and Cliff (2004) include this measure to construct a composite measure of sentiment.

The Consensus, Inc. publishes weekly Consensus Bullish Sentiment Index (CBSI), which is similar to the II and the AAII. While the latter two measure the sentiment of newsletter writers and individual investors in the spot equity market respectively, the CBSI represents the opinion of newsletter writers about the futures markets. The CBSI is a ratio of the bullish newsletters to all the newsletters in the sample. Sanders, Irwin and Leuthold (2003) examine the effectiveness of the CBSI as a contrary market indicator. The study covers 28 futures markets, including grain, livestock, food/fibre, financial and metal/energy futures contracts. However, the evidence is too weak to support the idea of the CBSI as a contrary opinion indicator. In addition, their study finds that the sentiment index predicts both returns continuations and returns reversals.

Qiu and Welch (2006) explore correlations among survey-based sentiment measures and conclude that these indices “measure some form of generic sentiment”. In short, the UBS/Gallup survey is highly correlated with the MCCI, and the AAII is negatively correlated with the UBS/Gallup Index. Fisher and Statman (2000) report correlations between the surveys

[11] A discussion on measures of investors’ sentiment is presented in Appendix B.

of small (AAII), medium (II, newsletter writers) and large (Merrill Lynch Wall Street sell-side strategist sentiment) investors. Their study concludes that “individual investors grow bullish when newsletter writers grow bullish, but not lockstep” and “changes in the sentiment of Wall Street strategists are virtually unrelated to changes in the sentiment of individual investors or newsletter writers”.

2.3.2 Objective measures of investor sentiment for the derivatives market

The put-call ratio is a sentiment proxy derived from options, and investors view it as a fear indicator. Bullish investors who believe prices are going to increase buy call options. This keeps the put-call ratio at lower levels. Bearish investors, who predict prices will decrease, tend to trade puts and driving the put-call ratio to a higher level. Bullish sentiment among noise traders imposes a price pressure effect and causes assets to be overvalued, thus some investors view it as a contrarian indicator. Luo and Li (2008) use put-call ratio derived from the Taiwan Stock Exchange Capitalisation weighted Index Options as a proxy for investor sentiment and relate it to trading behaviour of domestic and foreign traders in spot markets. Foreign traders act rationally and gain in the long-run. They sell when the market sentiment is high (i.e. when assets are overvalued) and buy when the market sentiment is low (i.e. when assets are undervalued). Domestic traders practice the opposites and suffer losses.

The Volatility Index (VIX) is widely recognised as the ‘fear’ benchmark for the U.S. Stock Markets. The Chicago Board Options Exchange (CBOE) introduced the VIX in 1993. The initial objective was to estimate the implied volatility of 30-day at-the-money S & P 100 Index Options prices. The VIX is the estimate of possible fluctuation of S&P 100 stock prices in the next 30 days. The VIX methodology was revised in 2003 to make it more practical for deriving trading and hedging strategies. The new calculation is a weighted average of S&P 500 Index options puts and calls of multiple strike prices. The inclusion of a wide range of the S&P 500 options makes the VIX a more representative market fear or confidence indicator. Generally, the VIX is higher when the investors are less confident or are concerned about market conditions and the VIX is usually lower when the sentiment is bullish and confident. Before 15 September 2008, the VIX fluctuated around 20 to 30. As the credit crisis hit Wall Street, the VIX rose sharply, approached 80 on 27 October 2008, reflecting a panic mood in the market.

The Trading Index (TRIN) was developed by Richard Arms, and is also known as the ARMS Index. It is derived from spot market data, but is widely used by investors in futures markets as a contrary indicator. The formula is as follows:

$$\mathrm{TRIN} = \frac{\text{Number of advancing issues / Number of declining issues}}{\text{Volume of advancing issues / Volume of declining issues}} \tag{2.1}$$

Some rules of thumb to interpret TRIN have been developed over the years. TRIN=1 indicates that the buy and sell volume are in balance. A TRIN below 1 indicates that the volume of advancing issues is more than the volume of declining issues, implying that the market is in an up-trend due to being overbought. A TRIN above one shows the volume of declining stocks is more than the volume of advancing stocks, and the market is in a down-trend due to being oversold. The current over-bought or over-sold scenario is expected to be reversed in the near future. The market is considered oversold when the 10-day moving average of the TRIN is more than 1.2 and is overbought when it is below 0.8.

Simon and Wiggins lll (2001) use the VIX, put-call ratio and trading index as proxies for market sentiment and conclude that these sentiment proxies are able to forecast S&P futures returns. Simulations with contrarian strategies record greater returns even after risk adjustment. The finding can be explained by the contrarian belief that in the period of low sentiment, assets are under-priced and the stock market will adjust to fair price again in the subsequent period. Most studies find these three sentiment measures as contrarian indicators for future price movements.

The Commodity Futures Trading Commission (CFTC) publishes the Commitment of Traders (COT) reports[12]. Traders are categorised as large speculators (those taking non-commercial positions), large hedgers (those taking commercial position) and small traders (those positions not exceeding the threshold set by CFTC). Wang (2003a) views the COT index to have resembled private information as compared to other opinion-based sentiment proxies and defines the COT sentiment index as:

$$SI_t^i = \frac{NP_t^i - Min(NP_t^i)}{Max(NP_t^i) - Min(NP_t^i)} x100 \tag{2.2}$$

where NP_t^i represents the net position (long position less short position) taken by traders type *i* in week *t*. $Max(NP_t^i)$ and $Min(NP_t^i)$ are the maximum or minimum net trading positions over a three-year moving window up to week *t*. In the same study, sentiment indices that are derived from six agricultural commodities are regressed on respective commodities' returns. As a conclusion, the large speculators' sentiment predicts price continuation and the large hedgers' sentiment predicts price reversals with no role for the small traders. Wang (2003b) extends the study to the S&P 500 index futures market, and comes to the same conclusion as in Wang (2001) for the three traders' types.

[12] The COT sentiment index is derived from trading position of traders in ten U.S. agricultural or commodity futures exchanges. While the report has been published since 1924, it became a monthly issue in 1962, shortened to fortnightly in 1992 and finally weekly from 2000 onwards. See http://www.cftc.gov/MarketReports/CommitmentsofTraders/AbouttheCOTReports/index.htm

Open interest is the total number of contracts entered into, to be delivered or cash settled. It is similar to liquidity indicators in the spot markets. When traders are actively trading the futures contracts, open interest increases, and this implies that traders are bullish. Wang, Keswani and Taylor (2006) use the OEX put-call open interest as the measure of sentiment and argue that this measure gives better information than volume data does. Wang (2001; 2003a; 2003b; 2004) uses open interest from the COT report to proxy sentiment for the large hedgers, the large speculator and the small traders in commodity futures markets.

2.3.3 Composite measures of investor sentiment

There are pros and cons of using a single sentiment measure. Some might be leading indicators while others could be lagging. Composite measures based on different sources could be a better comprehensive sentiment measure.

Brown and Cliff (2004) find a significant relationship among subjective sentiment measures and objective sentiment measures. There is a trade-off between keeping the analysis parsimonious or including all possible sentiment measures to avoid information loss. However, letting too many variables to enter the regression might cause co-linearity problems. As a result, the Kalman Filter and principal component analysis are used to generate a single measure of sentiment from a set of potential measures. These include the percentage change in margin borrowing, the percentage change in short interest, the ratio of short sales to total sales, the ratio of odd-lot sales to purchases, CBOE put-call ratio, trading positions reported by Commodities Futures Trading Commissions, and monthly forecasts by Market Vane.

Baker and Wurgler (2007) consider existing proxies, including investor surveys, investor mood, retail investor trades, mutual fund flows, trading volume, dividend premium, closed-end fund discount, option implied volatility, IPO first day returns, IPO volume, equity issues over total new issues and insider trading. Each of the proxies is regressed on a set of macroeconomics variables to remove the influence of economic fundamentals. The residuals from the regressions are used to create an index of sentiment. The dividend premium appears to have the most influence on investor sentiment, followed by IPO first day returns, number of IPOs, share turnover, and closed-end fund discounts. Equity share in new issues (i.e. equity issues over total new issues) has the least weight. In another study, Baker, Wurgler, & Yuan (2012) while examining global and local sentiment, reduce the number of proxies from six to four: volatility premium, total volume of IPOs, IPOs' first day returns and market turnover.

However, the investor sentiment is reflected on these proxies at a different pace given the same shock. Some take longer to adjust and others quickly adjust to the sentiment shock. Lei (2005) points out that composite sentiment measures might be noisier than a single sentiment proxy if investor sentiment does not drive the sentiment proxies at the same time or with the same time lag. Similarly, Simon and Wiggins lll (2001) argue that a specific proxy might have predictive power at specific point in the market cycle, which could be diluted or obscure individual information content if aggregated to create an overall index.

Lemmon and Portniaguina (2006) use the sentiment component of consumer indices[13] and find that the consumer confidence is negatively and weakly related to the former. The sentiment measures used in both studies are an optimistic indicator that theoretically should be positively correlated. One potential reason for the result is the frequency of the measures, while the composite sentiment index is constructed annually and the consumer indices are constructed quarterly. These two sentiment proxies are based on the same country but capture sentiment over different time intervals. The fluctuations in the consumer confidence

[13] University of Michigan Survey of Consumer Sentiment and Conference Board survey of consumer confidence. The sentiment component is the residuals from the regression of the consumer confidence against macroeconomics variables.

indices over quarters are reflected in the correlation coefficients with the composite sentiment index.

2.3.4 Media/News measures of investor sentiment

Early study of investor sentiment based on media content begins with the classification of financial newsletters or advisory comments into bullish, bearish, or having mixed opinion and expecting correction. Investor Intelligence (henceforth II) has provided the fortnightly US Advisor Sentiment Index since 1963, becoming weekly in 1969. Over the publication period, only the same four editors consistently work on classifying over one hundred independent market newsletter authors' opinions. The index is extensively used in research works related to investor behaviour and market anomalies. Solt and Statman (1988) calculate a bearish sentiment index as a percentage of bearish newsletters to the total of bearish and bullish newsletters. Siegel (1992) uses the bull-to-bear ratio to explore the association between sentiment and asset pricing. The ratio inversed from 2 to 1 before October 1987 crash, to 1 to 2 after the market crash. Lee, Jiang, and Indro (2002) argue that newsletters written by investment advisors and read by individual investors, will change individual sentiment at the end. The number of bullish newsletters relative to the total of number of bullish and bearish newsletters is used as a proxy of individual investor sentiment. On the other hand, Brown and Cliff (2004; 2005) calculate bull-bear spread (percentage of bullish newsletters minus percentage of bearish newsletters) as an indicator of investor sentiment for institutional investors or professionals.

The development of information technology enables dissemination of information through the internet. Investors gather real time trading data, obtain investment specialist advice, and even exchange opinions on-line. Wysocki (1999) started another possible way to measure investor sentiment in a rather primitive way by counting the number of postings in stock message boards. The study uses the number of posting to predict next-day trading volume and next-day abnormal stock returns. Tumarkin and Whitelaw (2001) employ event study techniques to investigate the impact of abnormal message activity on stock prices and trading volume. They argue that the messages may contain insider proprietary information, influencing overall market sentiment. Those participating in the message board discussion influence each other in making investment decisions. Dewally (2000) constructs the Thread Direction Index (range from -1 to +1) to indicate the sentiment in messages. If all the messages posted are in negative about the company, the index is -1 and if all the messages posted are positive, the formula will yield +1.

In the early stages, researchers transformed newsletters, messages or news into categories of bullish, bearish or uncertain. This approach is time consuming thus limiting the sample size of the studies. In addition, consistencies of classification done are questionable due to subjectivity and too much reliance on human effort.

Recent studies overcome the drawback by using artificial intelligence to categorise the media content, because computer algorithms can perform consistent grouping. Antweiler and Frank (2004) employ the naive Bayes and support vector machine to interpret more than 1.5 million messages from internet message board. Tetlock (2007) uses the General Inquirer to scan through news article from a long-standing column in the Wall Street Journal; approximately 4000 articles in total. These methods are proven reliable, but the issue of validity remains an obstacle yet to be overcome. Durant (2008) proves that distance between the timing of model training[14] and the timing of the model's prediction is another factor that

[14] "Supervised machine learning uses statistics to build mathematical models to perform a particular performance task based on example data, known as training data. The task of the model is to generalize from the training examples in a reasonable way so that it may classify unseen examples as accurately as the training examples. Once the machine learning classifier is created, it can be applied to other" (Durant, 2008, p.1)

influences the predictive power of classification models. Other factors reported in earlier studies including machine learning and induction techniques, class composition, sample size, and features selection. Table 2.1 gives an overview of investor sentiment measures.

2.3.5 Fundamentals information versus sentiment

Researchers adopt two approaches to account for the fundamental components subsumed in the sentiment measures. First, raw sentiment measures are regressed on a set of fundamental proxies, and the residuals from the regression are considered as pure sentiment measures (Baker & Wurgler, 2006; Glushkov, 2006; Verma et al., 2008). The fundamentals proxies include the monthly changes in the industrial production index; yield on 1-month US Treasury bills; difference in monthly yields on 3-month and 1-month treasury bills; spread on the 10-year US treasury bond and 3-month treasury bills; Baa and Aaa corporate bond default spread dividend yield; monthly changes of the consumer price index; market portfolio's excess returns; small-minus-big and high-minus-Low premium (Fama & French, 1993); momentum factors (Jegadeesh & Titman, 1993); currency fluctuations, growth in consumption and dummy variables for recessions.

Alternatively, the fundamental variables are added to the model specified for hypothesis testing as control variables. Brown and Cliff (2005) use a similar set of fundamental variables as above. Wang, Keswani and Taylor (2006) argue that sentiment arises due to market behaviour, thus lag returns and lag volatilities are included as control variables.

Besides approaching this issue from the perspective of econometric methods, researchers justify whether a variable proxies for fundamentals or sentiment by examining its co-movement patterns with price changes or returns. Tetlock (2007) does not explicitly remove or control for the fundamental variables. Instead, the researcher justifies that the pessimism factor is a measure of sentiment because its initial negative impact on stock index returns reverse after a few days. If the pessimism factor reflects fundamental information, the negative impact should persist and should have no reversal.

2.3.6 Conclusion

Other than sentiment measures based on trading data in spot markets, sentiment measures are derived from media content and surveys. These measures can represent investor sentiment in the derivatives markets even though the word “derivatives” is not explicitly mentioned. For example, the Abreast of the Market summarises the US market in general; the Yale Sentiment Index and the Michigan Consumer Confidence Index reflect how the respondents perceive the outlook for the future economy. Valuation in the derivatives markets is closely related to the spot markets and economic outlook. However, researchers still find the needs of having sentiment measures pertaining to derivatives markets, in order to minimise the noise from other markets.

Table 2.1 Summary of Commonly Used Investor Sentiment

This table gives an overview of investor sentiment measures. These measures are categorised by market (spot or derivatives), measurement type (subjective or objective), indicator type (bullish, bearish, confidence or fear) and the Target Market. Over the years, these measures have been reconstructed for other countries, based on similar surveys or methods.

Market	Measurement Type	Indicator	Indicator Type	Target Market
Spot	Subjective/ Survey-based	Investor Intelligence Survey	Bullish	US
		Yale Buy on Dips Index	Confidence	US
		Yale Crash Index	Confidence	US
		Yale Valuation Index	Confidence	US
		Michigan Consumer Confidence Index	Confidence	US
		American Association of Individual Investor Survey (AAII)	Bullish	US
	Objective/ Trading-based	Closed-end funds discounts	Bearish	
		Buy-Sell Imbalance	Bullish	
		Mutual-fund flows		
		Market liquidity	Bullish	
		IPO price	Bullish	
Derivatives	Subjective/ Survey-based	Consensus Bullish sentiment Index	Bullish	US Futures market
		Market Vane Buy-Sell Recommendations	Bullish	US Futures Market
	Objective/ Trading-based	Put-Call ratio	Bearish	
		CBOE Volatility Index (VIX)	Fear	US stock index options
		TRIN index / ARMS Index	Bearish	
		Commitments of Traders Index (COT)	Bullish	US Agricultural and Commodity Futures

From the perspective of chronological order, news-based sentiment measures should have advantage over the trading-based sentiment measures. The logic behind this is simple. Asset valuation is based on the expected future cash flows, and macroeconomic announcements and company specific earnings information usually made public in the form of news releases. Upon receiving news, investors adjust their beliefs and finally make investment decisions. Finally, this “information” is reflected in the trading data, for example price and trading volume. Hence, the news-based sentiment measures should lead the trading-based sentiment measures. The challenge to construct a good news-based sentiment measure relies on the efficiency of computer algorithms to classify sentiment. In addition, the news contains real information and investors themselves interpret the news with sentiment. Thus, ideally one should be able decompose the sentiment measures generated from news to irrational and rational components or fundamental and noise components by some means.

2.4 THEORY AND EVIDENCE OF THE INVESTOR SENTIMENT ON ASSETS RETURNS AND TRADING VOLUME

Different groups of investors tend to have different perceptions about the fundamentals and future market prospects. In addition, it is possible that they will interpret the same information differently. Wongchoti, Wu and Young (2009) posit that institutional investors take past trading information as a continuation indicator and the retail traders are interpreting past trading information as a contrary indicator. Wang (2003a) finds that the large speculator sentiment is a price continuation indicator and the large hedger sentiment is a contrary indicator. Small investors are assumed to be more inclined to sentiment in the previous studies. However, recent evidence indicates that institutional traders are acting on the sentiment rather than on the fundamentals.

The substance that makes sentiment measures a research interest is that researchers are yet to reach a mutual conclusion on why investor sentiment has a prolonged impact on asset prices. The role of investor sentiment in price formation, and how it affects trading activities are yet to be agreed upon.

2.4.1 Impact of investor sentiment on cross-section of average returns

Neal and Wheatley (1998) examine the returns predictability of the closed-end fund discounts and net mutual fund redemptions. The study is based on the first decile and the tenth decile of the value weighted NYSE-AMEX portfolios. The close-end fund discounts and net redemptions are positively related to the small firms' expected returns. Conversely, the large firms show evidence to the contrary that the net redemptions are negatively predicting the expected returns. However, there is little evidence to suggest that the odd lot ratio is able to explain the expected returns.

While most of the studies are focused on the near-term sentiment predictive power and documented little supportive evidence, Brown and Cliff (2005) believe that the sentiment is exuberant, and escalates over time. The study uses the II bull-bear spreads and relates it to the long horizon cumulative returns and pricing errors. In order to examine the effect of size or value to returns predictability of sentiment, the study includes 36 portfolios selected based on size and book/market sorts. The results suggest that high sentiment predicts negative futures returns of the larger firms, or those firms with low book-to-market value. This finding disputes the conventional beliefs that smaller firms are more sensitive to sentiment. Next, the study finds that market pricing errors are positively related to sentiment even after controlling for the rational factors. Combined together, these findings imply that overvaluation during the period of high sentiment will be corrected over the longer horizon.

Baker and Stein (2004) propose a model to explain how the expected returns are related to the liquidity proxy that serves as a measure of sentiment. In a market facing limits to arbitrage, irrational investors keep themselves away from the market when they are bearish. When they are over confident about the information that they received, they tend to over value stocks and actively trade in the market. This is reflected in the market liquidity measures, for example, high share turnover and low bid-ask spreads. Therefore, the expected returns of these stocks are lower. More evidence has built upon the foundation of this model. Baker and Wurgler (2006) argue that sentiment only affects the stock prices when uninformed demand shocks and limits to arbitrage exist. Investor sentiment as a measure of propensity to speculate gives rise to demand shocks in the cross-sections at different magnitudes, given the same difficulty to arbitrage. In another scenario, limits to arbitrage vary in the cross sections, while the demand shocks are constant. The authors suggest that the conventional behavioural explanations over-simplify the concept of sentiment by categorising it into a few biases (e.g. overconfidence, representativeness, and conservatism) and combining it with limits to arbitrage and relating it to stock prices. The study using an aggregate measure of sentiment finds stocks that are hard to arbitrage, for example, "stocks of low capitalisation, younger, unprofitable, high volatility, non-dividend paying, growth companies or firms in financial distress are likely to be disproportionately sensitive to a broad wave of investor sentiment". In a follow-up study, Baker and Wurgler (2007) consider the research question focusing on how to quantify the impacts of sentiment. Average future returns of speculative stocks are, on average, higher than bond-like stocks in the periods of low sentiment, which is consistent with the classical asset pricing models that the risk of bearing speculative stocks is priced. However, in a period of high sentiment, the speculative stocks are over-priced; its average future returns are on average lower than the bond-like stocks, which is consistent with behavioural models.

Kumar and Lee (2006) tests the role of individual investor sentiment in price formation, using the Buy-Sell Imbalance as the measure of sentiment. The finding proves that individual investors tend to buy or sell the same group of stocks during the same period. This systematic trading pattern is proven to be able to explain the returns co-movements of the same basket of stocks. Evidence shows that individual investors prefer to invest in those stocks of smaller, lower-priced, higher book-to-market ratio, and lower institutionally owned firms. Consequently, these stocks are sensitive to changes in the individual investor sentiment. In addition, those stocks facing higher cost of arbitrage are exhibiting greater sensitivity to sentiment. The sentiment-returns relationship for individual investors is positive; excess returns are higher when individual investors are relatively bullish. Moreover, the BSI measure outperforms the popular closed-end fund discounts in explaining the returns variability. Generally, the conclusions are in line with Lee, Shleifer and Thaler (1991) and Barberis, Shleifer, and Wurgler (2005).

2.4.2 Impact of investor sentiment on aggregate market returns

Fisher and Statman (2000) examine the relationship between sentiment of individual (AAII), newsletter writers (II) and Wall Street strategist (sell-side strategist sentiment by Merrill Lynch) with future S&P returns. The study documents that individual investors and Wall Street strategists' sentiment are reliable contrary indicators because they are negatively related to future S&P returns. Sentiment of newsletter writers is negatively related to sentiment but this is not statistically significant. Interestingly, returns of large stocks show higher correlation with the shift in individual sentiment than with large investor sentiment. In addition, returns of small stocks show higher correlation with the change in large investor sentiment than that in individual investor sentiment.

Brown and Cliff (2004) examine returns predictability by constructing a composite measure of sentiment. Despite the strong contemporaneous relationship between sentiment and returns, there is insufficient evidence to determine the ability of sentiment in predicting near-term stocks returns. The study also finds a positive association between institutional investor sentiment and stock returns. This is counter intuitive because conventionally, individual investors are believed to be more prone to noise.

In the context of the S&P 100 index, Wang, Keswani and Taylor (2006) use investor sentiment derived from the S&P options market, i.e. OEX put-call trading volume ratio (PCV), OEX put-call open interest ratio (PCO), and the ARMS index as proxies of daily investor sentiment. In addition to PCO and PCV, the study also uses AAII and II as weekly proxies for individual investor sentiment and institutional investor sentiment, respectively. Consistent with Fisher and Statman (2000) and Brown and Cliff (2004), the study finds evidence that returns Granger-cause sentiment but little evidence that sentiment Granger-cause returns.

Tetlock (2007), using news as proxy for investor sentiment, finds that the pessimism factor derived from Abreast of the Market, a column in the Wall Street Journal, has negative influence on the next day's Dow Jones returns. However, the patterns reverse in the subsequent four days. He explains that the initial negative impact of the pessimism factor on returns is due to the adjustment of the Dow Jones Index to negative news or due to negative investor sentiment. The later reversals rule out the information explanation because the impact of information should persist rather than reverse. In addition, the finding is also consistent with DeLong et al. (1990b) feedback trading hypothesis because evidence also shows that negative returns predict higher pessimism the next day.

Early studies on investor sentiment treat sentiment as fully irrational. However, Verma, Baklaci and Soydemir (2008) believe there are two components to sentiment: rational sentiment and irrational sentiment. Rational sentiment can be attributed to fundamental and risk considerations, while irrational sentiment is purely noise. AAII represents individual investor sentiment and II is the proxy for institutional investor sentiment. In an attempt to segregate irrational sentiment from AAII and II, sentiment proxies are regressed on a set of fundamental variables. Residuals from these regressions are used to proxy irrational sentiment. The residuals, AAII, II, S&P returns, and Dow Jones Industrial index returns enter a VAR model. Impulse response functions capture the response of one variable to a shock from another variable respectively. The study finds that rational sentiment has a greater impact on stocks returns as compared to irrational sentiment, implying that rational sentiment induces trading based on fundamentals. On the other hand, past stock returns have positive effects on stocks returns and the positive impact of irrational sentiment on stock returns is reversed in the subsequent period. This is consistent with the conclusion that sentiment is a contrarian indicator.

2.4.3 Impact of investor sentiment on derivatives market returns

Simon and Wiggins lll (2001) examine the predictive power of the VIX, put call ratio and the TRIN on S&P futures returns, spanning from January 1989 through June 1999. The percentage change of returns over 10-day, 20-day, and 30-day are regressed on each of the sentiment indicators. Generally, the coefficients of the regressions are positive and statistically significant. This implies that the VIX, put-call ratio, and the TRIN are reliable contrary indicators. The study also finds that when the sentiment is more extreme (top 10 decile), the impact of sentiment is greater over a longer period, that is, 30-day return as compared to 10-day return.

Instead of explicitly determining the sentiment-return relationship, Fung, Mok, and Lam (2000) and Fung and Lam (2004), focus on the overreaction hypothesis. De Bondt and Thaler (1985) suggest that investors attach too much weight to recent information and too little weight on the prior data, finally causing systematic pricing error. If this is true, the price reversals will follow an extreme initial price changes. Then, greater initial price changes will lead to greater a reversal, which is known as the magnitude effect. Although studies do not elaborate the role of investor sentiment pertaining to overreaction, it is safe to associate extreme price changes with investor sentiment. Fung, Mok, and Lam (2000) examine price reversals of the S&P 500 Index Futures and the Hang Seng Index Futures (HSIF) using tick-by-tick data. The results support the overreaction hypothesis, and the HSIF demonstrates stronger evidence of this. However, implementing strategy to track these systematic patterns does not lead to economically and statistically significant positive returns. Fung and Lam (2004) calculate pricing error as the difference between a futures price and its theoretical value, and suggest that pricing error can serve as a proxy for investor sentiment. Investors are over confident when the pricing errors are positive (over-priced) and under confident when the pricing errors are negative (under-priced). Evidence supports the overreaction hypothesis. Subsequent returns are positive when the HSIF is under-priced and conversely, negative when it is over-priced.

Wang (2001, 2003b; 2004) extensively study the sentiment-returns relationship for commodity futures, index futures, and foreign exchange (FOREX) futures. The sentiment index for the large hedger, the large speculator and the small trader are derived from the COT report. Wang (2001) uses total open interest to compute the sentiment index while the net trading position, i.e. long open interest less short open interest is used in Wang (2003b) and Wang (2004). Generally, these studies conclude that sentiment of the large speculators predicts price continuation while sentiment of the large hedgers is a contrary indicator. Traders long contracts when they are bullish (high sentiment) and short contracts when they are bearish (low sentiment). The findings suggest two possible market timing strategies. First, long contracts when the large speculators are bullish or the hedgers are bearish. Second, short contracts when the large speculators are bearish or the hedgers are bullish. Evidence shows that the large speculators possess better timing ability and earn higher subsequent returns in all the three categories of futures contracts. One possible explanation is the hedging pressure theory. The hedgers pay to transfer risks to the speculators and the speculators price the risk they bear. Market timing strategies based on extreme sentiment (top 2 deciles) outperform strategies based on sentiment alone. In addition, the strategy that considers sentiment of the large speculator and the large hedger simultaneously also demonstrates higher returns than the strategy that considers solely sentiment of the large speculator or the large hedger. Wang (2001) and Wang (2003b) do not account for market risk, thus it is unrealistic to conclude that the sentiment-based market timing strategies are “profitable” when the returns are positive. Wang (2004) includes risk-adjusted returns and the conclusion is still consistent with the earlier literature.

As opposed to the conventional investment ‘Buy-Low-Sell-High’ rule, there are investors following trend chasing strategies; buying when the prices rally and selling when the prices fall. Kurov (2008) finds that the intensity of feedback trading increases with sentiment, and conjectures that uninformed expectations of noise traders are at least a part of the driving force. The study uses the AAII and the II as proxies for investor sentiment. The impact of sentiment prevails in both the E-mini NASDAQ 100 index futures and the E-mini S&P 500 index futures. The intensity of positive feedback trading increases when investors are bullish and is less intense when investors are bearish. The same study also provides evidence on sentiment-returns volatility relationships.

Sanders, Irwin, and Leuthold (2003) focus on the CBSI that represents sentiment of newsletter writers and targeted readers. Their study covers 28 US futures markets with little evidence supporting the theory of contrary opinion. Past returns are found to predict subsequent sentiment, while past sentiment does not significantly predict subsequent futures returns.

2.4.4 Impact of investor sentiment on liquidity

Early studies do not investigate whether investor sentiment predicts liquidity. Instead, these studies view change in liquidity as evidence of investor sentiment affecting trading activities, and aim to explore the role of investor sentiment in the price formation process.

Campbell, Grossman and Wang (1993) examine the relationship between trading volume and serial correlation in stock returns. The model assumes there are two types of investors. Public information, for example, negative news about the stock market create negative sentiment among non-informed traders who will try to liquidate their positions at low prices. On the other hand, the risk-averse utility maximisers are willing to buy these stocks at lower price with higher expected returns. Prices increase on subsequent days due to higher expected returns. Buying and selling between the non-informed and the risk-averse utility maximisers are reflected in unusually high trading volume.

Coval and Shumway (2001) examine the relationship between the sound level at the trading floor and volume. When traders expect that the cost to execute trades will increase in the future, they tend to raise their voice, and demand execution of their trades immediately.

Baker and Stein (2004) assumes irrational investors will overweigh private signals they received and underweight the decisions made by other investors. Positive market signals invoke positive sentiment among irrational traders. Short-sales constraints keep them away from the market unless they value the assets more highly than their fundamental value. Increasing liquidity measures are an indication of irrational investors swayed by positive sentiment present in the market. The study shows that liquidity measures, i.e. equity issuance and share turnover predict future returns of a CRSP equal-weighted portfolio.

With respect to derivative markets, Bessembinder and Seguin (1993) investigate the evidence from currencies, metals, agricultural commodities and financial futures markets and finds that trading volume is positively related to volatility while open interest has a negative impact on returns volatility. Chan, Fung, and Leung (2004) draw the same conclusion for four futures contracts (metal and agricultural commodities) traded on three Chinese futures exchanges. Fung and Patterson (1999) find the volatility-volume and volatility-open interest relationships are positive for five currency futures markets.

Liquidity indicators are strongly associated with investor sentiment, making it a widely used proxy for investor sentiment. Lei (2005) uses trading volume trend as a measure of investor sentiment as a solution to the nonstationary problem of trading volume series. Wang, Keswani and Taylor (2006) use the OEX put-call open interest as the measure of sentiment and argue that this measure is a better indicator than volume. Wang (2001; 2003a; 2003b; 2004) use open interest from the COT report to proxy sentiment for the large hedgers, the large speculator and the small traders in commodity futures markets. Baker and Wurgler (2007) include IPO volume and IPO first day returns in the creation of a composite sentiment index.

2.5 THEORY AND EVIDENCE OF THE ROLE OF INVESTOR SENTIMENT ON THE RETURNS MEAN-VARIANCE RELATIONSHIP

2.5.1 Introduction

The Risk-return relationship is the tenet of capital asset pricing theories. The Capital Asset Pricing Model (CAPM) by Sharpe (1964), Lintner (1965), Mossin (1966) and the Intertemporal Capital Asset Pricing Model (ICAPM) by Merton (1973) assume that investors are rational, therefore, risk-averse. Investors require higher compensation for bearing greater risks. This implies a positive risk-return relationship.

Researchers adopt a time series approach to examine the relationship between the market portfolio conditional returns and its conditional variance, while others use a cross-sectional approach to examine the relationship between the assets' expected returns and systematic market risk. Based on the CAPM, one regresses excess expected returns of an asset on market portfolio returns. The coefficient of the regression is viewed as a proxy for market risk. Alternatively, the covariance between marginal utility and some risk factors in the framework of ICAPM are used as a risk measure.

However, literatures fail to attain consensus on the returns mean-variance relationship. Previous studies suggest that various volatility model specifications and frequency of data used are the main reasons of the disagreement. Other studies attributed the disagreement to variation in the sample period, omitted variables in estimating the unobserved volatility, or the mean-variance relationship itself. I begin with various volatility model specifications, sampling periods, sampling intervals, and finally omitted conditioning variables. I discuss these issues in the following sections.

2.5.2 Commonly used volatility measures for time series data

Risk is uncertainties faced by the investors, specifically downside or negative returns. There are a few commonly used measures of risk. Given the time series nature of the market portfolio index, variance or standard deviation of the market portfolio returns are common measures of risk. The basic formula of variance is average squared deviations of excess return from its mean.

$$\sigma^2 = \frac{1}{N-1}\sum_{t=1}^{N}(R_t - \bar{R}_t)^2$$

Where *N* is the sample size, R_t is the returns and $\bar{R}_t$ is the average returns.

The modified versions of this formula are proven to improve the volatility forecasting power. Figlewski (1997) suggests that the sample mean is an inaccurate estimate of true mean and ignoring the $\bar{R}_t$ in the above formula can improve the forecast power. In another study on

Deutsche Mark-Dollar volatility , Andersen and Bollerslev (1998a) assume $\bar{R}_t = 0$. Alternatively, Ding, Engle and Granger (1993) suggest the use of absolute returns instead of squared returns and this is further supported by Ederington and Guan (2005). Others suggest removal of extreme values (Garman & Klass, 1980; Parkinson, 1980). These formulas are relatively tractable, but fundamentally problematic because they capture both negative and positive sides of the deviations (Poon & Granger, 2003). Thus, taking only the returns deviation below the mean or the downside variance is suggested (Harlow & Rao, 1989; Sortino & Van Der Meer, 1991). Nonetheless, this method is not widely adopted by the professionals because downside risk is operationally complicated to obtain and the target return parameter is hard to determine. These issues make the portfolio selection process tedious (Grootveld & Hallerbach, 1999) .

Another form of model-free volatility measure is realised variance that uses high frequency intraday data. The sum of squared of high frequency intraday returns make a good estimate of realised volatility, but it is subject to market microstructure noise. An optimal sampling interval mitigates this problem. The forecast error decreases as the sampling interval increases. The 5-minutes returns are optimal, because of low levels of autocorrelation and the zero mean property (Andersen & Bollerslev, 1998b; Andersen, Bollerslev, & Lange, 1999; Bollerslev, Gibson, & Zhou, 2011). The mathematical representation of the sum of squares of returns is as follow:

$$RV_t = \sum_{i=1}^{n} (R_{t,i})^2$$

Where *n* is the number of returns observed in day *t*, $R_{t,i}$ is the *i*th return during day *t.*

The ARCH (Autoregressive Conditional Heteroskedasticity) class models forecast one-step-ahead variance, conditioning on information available at the point of time of estimation.

GARCH-type models (Bollerslev, 1986) are widely used because these models are relatively parsimonious as compared to the ARCH models. For instance, GARCH-in-Mean includes conditional variance in the mean equation (Bollerslev, Engle, & Wooldridge, 1988; Engle, Lilien, & Robins, 1987); Exponential GARCH considers the conditional variance as an exponential function and includes the asymmetric effects of good news and bad news (Nelson, 1991); GJR-GARCH (Glosten, Jagannathan, & Runkle, 1993) introduces a asymmetric term for the leverage effect; Threshold-GARCH (Zakoian, 1994) estimates conditional standard deviation and allows for positive and negative shocks; Quadratic-ARCH (Sentana, 1995) is a quadratic function that captures dynamic asymmetries; Component-GARCH(Engle & Lee, 1999) allows the short- and long-term components in the variance equation; Continuous time-GARCH (Klüppelberg, Lindner, & Maller, 2004) allows for unevenly spaced data due to non-trading day; Heterogeneous Autoregressive-GARCH (Corsi, Mittnik, Pigorsch, & Pigorsch, 2008) captures long memory behaviour of variance and accounts for volatility clustering. Poon and Granger (2003) perform an extensive review on various versions of GARCH-type volatility models. The criticisms on these procedures generally focus on the underlying parametric assumptions and the trade-off between degrees of freedom and the number of conditioning variables that can be included in the model.

The historical volatility, realised volatility and conditional volatility use historical stock prices as the inputs. Alternatively, we can calculate implied volatility from options prices. This is a forward-looking measure because the option prices incorporate the investors' expectation on the asset prices up to the option's expiry.

2.5.3 The sentiment-variance relationship

Researchers try to uncover whether investor sentiment measures are able to explain returns volatility and whether the investors price the volatility induced by sentiment. The latter implies that investor sentiment is a systematic risk because only risks that cannot be eliminated through diversification are priced. Besides, there are attempts to investigate the impact of investor sentiment on investors' propensity to take risk.

Investor sentiment lends an explanation to excess returns volatility. Shiller (1989) finds that the volatility of closed-end fund discounts is greater than the volatility of its dividends and attributes to the phenomena to investors' irrationality. Pontiff (1997) extends the study, and compares the returns volatility of a closed-end fund with its' underlying assets. The closed-end funds' returns volatility is 64% more volatile than its underlying assets. In addition, the funds' stock price returns in excess of net asset value returns can be better explained by sentiment risk instead of the Fama and French (1993) measures of risks for market, value and size. These early evidences indirectly attribute to returns volatility to fluctuations in investor sentiment. Following these prior studies and motivated by Delong et.al (1990b), Brown (1999) uses sentiment index constructed from the AAII survey, to directly test the sentiment-volatility hypothesis. The findings support the proposition that returns volatility increases due to noise traders trading on sentiment.

Baker and Wurgler (2006, 2007) suggest that hard-to-value and hard-to-arbitrage stocks are highly volatile and tend to record lower returns following high sentiment period. These studies not only provide empirical evidence on the sentiment-return relationship, but also indirectly explain why the risk-return trade off does not hold at times. Noise traders tend to overprice assets during high sentiment periods. The returns are lower due to a high purchase price and subsequent price reversals due to over-pricing. This explains the reason why there is no compensation for sentiment-induced volatility.

Incorporating the changing market dynamics into the volatility model is expected to produce better estimates for risk. However, the market data and macroeconomic information are not always available in high frequency. Mitra, Mitra and Dibartolomeo (2009) propose an alternative solution. They include two common measures of investor sentiment: options implied volatility and news sentiment, into the model to improve the estimation of the portfolio risk (volatility). Options implied volatility is a forward-looking measure of investors' expectations for future volatility. It can be calculated up to per minute basis, while news can be converted into sentiment measure on daily basis. These variables improve the volatility estimates.

Based on the above studies, several sentiment measures are identified as explanatory factors to returns volatility. Yang and Wu (2010) construct nine sentiment measures from four categories of trading data, including overall trading, margin trading, TAIEX options and foreign plus institutional investors. They conduct Grey analysis to determine the sequential relationships among these factors, and rank them according to the domination power in determining the relationship with price volatility. The resulted ranking is as follow: short-sales volumes, open interest, put-call ratios, trading volume and finally buy-sell orders.

However, not all studies corroborate the sentiment induce volatility paradigm. Since there are evidences on a bi-directional sentiment-returns relationship (see Brown & Cliff, 2004; Solt & Statman, 1988), it is reasonable to suspect that there is a two-way sentiment-volatility relationship. Wang, Keswani and Taylor (2006) suggest that the prior evidence could be spurious due to the role of returns on predicting volatility being omitted. There is evidence that lags of returns and lags of volatility Granger-cause sentiment, but no countervailing proof is found in their study.

Based on the present evidence, it is reasonable to maintain that at least some portion of the return volatility can be attributed to investor sentiment. This has motivated further tests on whether the volatility arising from sentiment trading is priced. Some research investigates a more general question, for example whether sentiment can predict returns or if sentiment factors enter the returns generating process. The role of investor sentiment as a risk factor is implicitly implied if sentiment is priced, evidenced by investor sentiment significantly predicting returns (Elton et al., 1998; Kumar & Lee, 2006; Sias, Starks, & Tinic, 2001) .

Another line of research explicitly models the investor sentiment into the volatility equation. This enables the examination of the role of investor sentiment on volatility as a measure of risk. However, very few studies have tested this. The theory was first proposed by Lee, Shleifer and Thaler (1991), which states that closed-end funds and small stocks face noise trader risk. There should be compensation for bearing the risk. This motivates Lee, Jiang and Indro (2002) to carry out empirical tests. They explicitly model the sentiment factor in the mean and variance equation in a GARCH model. Fitting the model to U.S. aggregate data, they find that a positive (negative) change in sentiment leads to the following week's lower (higher) conditional volatility. Moreover, the positive relationship between returns and change in sentiment suggests that investors systematically price the sentiment as if it is a risk measure. Beaumont, Daele, Frijns, Lehnert, & Muller (2008) apply the same model to US market daily data and draw the same conclusion. In spite of this, the impact of sentiment on volatility is not confirmed by Samsell (2007) when the empirical model is tested with monthly data.

2.5.4 Empirical evidence of mean-variance relationship

The literature uses the term 'risk-return' and 'mean-variance' relationships interchangeably because variance is widely used as a measure of risk. There are intense investigations on this building block of modern rational asset pricing theory.

The Capital Asset Pricing Model predicts a positive mean-variance relationship. Extant studies employ different measures of mean and variance, span over different sampling periods and adopt various estimation methods, and find a positive mean-variance relationship (Bali, Demirtas, & Levy, 2009; Bali & Peng, 2006; Bollerslev et al., 1988; Bollerslev et al., 2011; Chou, 1988; Darrat, Gilley, Li, & Wu, 2011; De Santis & Imrohoroğlu, 1997; French, Schwert, & Stambaugh, 1987; Ghysels, Santa-Clara, & Valkanov, 2005; Goyal & Santa-Clara, 2003; Guo & Neely, 2008; Guo & Whitelaw, 2006; Harrison & Zhang, 1999; Ludvigson & Ng, 2007; Lundblad, 2007; Scruggs, 1998; Tang & Shum, 2004; Yu & Yuan, 2011). On the other hand, there is empirical evidence that does not square with the capital asset pricing theory, and supports counter-intuitive negative relationship between mean and variance of returns (Brandt & Kang, 2004; Campbell, 1987; Glosten et al., 1993; Harvey, 2001; Li, 2011; Whitelaw, 1994).

The debate is further complicated by evidence of an insignificant mean-variance relationship (Baillie & DeGennaro, 1990; Bali, Cakici, Yan, & Zhang, 2005; Bollerslev & Zhou, 2006; Campbell & Hentschel, 1992; Chan, Karolyi, & Stulz, 1992; Nelson, 1991; Theodossiou & Lee, 1995). Furthermore, Turner, Startz, and Nelson (1989) and Müller, Durand, and Maller (2011) yield mixed findings. Thus the literature does not reach a consensus on the direction of the mean-variance relationship and propose arguments based on rational expectations, biased beliefs or methodological issues.

2.5.5 Justifications for mixed mean-variance relationship

The empirical evidence for mean-variance relationship is mixed. There are two broad categories of explanations: methodological issues and other mechanisms. Methodological issues include the measurement of variance, mean-variance model specification, sample period, and linear assumption of the mean-variance trade-off. Other mechanisms include leverage aversion, dividends, and investor sentiment.

2.5.5.1 Variance measures

The main area of debate for the mixed mean-variance relationship lies on various variance measures as it is proven to be sensitive to the variance measure used in estimating the relationship.

The discussions on model free historical variance measures focus on the weighting methods. Goyal and Santa-Clara (2003) use equal-weighted average of cross-sectional variance instead of market portfolio's variance and find a positive mean-variance relationship. Adopting the same model specification, Bali, Cakici, Yan, and Zhang (2005) use value weighted average of cross-sectional variance and find no significant relationship. Based on the contradicting findings, they suggest the positive trade-off between return and variance is partly driven by the liquidity premium from holding small stocks.

Earlier studies employ the *GARCH-M* model that relates mean returns to conditional variance in the mean equation, but their conclusions are somewhat inconsistent. Baillie and DeGennaro (1990) , Chan, Karolyi and Stulz (1992) , Theodossiou and Lee(1995) find no significant mean-variance relationship; French et al. (1987) and Scruggs (1998) find a positive relationship. The modified versions of ARCH yield a more consistent positive mean-variance relationship, for example Chou (1988) uses IGARCH; Bollerslev, Engle and Wooldrige (1988) use multivariate GARCH; Guo and Neely (2008) use component GARCH; Yu and Yuan (2011) use asymmetric GARCH.

Besides the various volatility ARCH types models, prior research suggests that omitted variables in estimating volatility also contribute to the mixed mean-variance relationships. The literatures find a positive relationship after some macroeconomics or state financial variables are included in the volatility model. Scruggs (1998) and Glosten et al. (1993) include long-term government bond returns. Bollerslev, Gibson and Zhou (2011) include Moody AAA bond spread, housing starts, S&P price-earning ratio, industrial production, the producer price index and payroll employment. Whitelaw (1994) finds that the spread between Treasury bills and commercial paper has predictive power over volatility, which will subsequently affect the mean-variance relationship. Glosten et al. (1993) include seasonal dummy variables for January and October.

Other than the above mentioned potentially omitted variables, Glosten et al. (1993) show that if asymmetric responses are allowed in the variance equation of the *GARCH-M* model, the initial positive mean-variance relationship turns negative (Harvey, 2001).

2.5.5.2 Mean-variance model specification

Guo and Whitelaw (2006) include hedging components (investment opportunities) and find a significant positive relationship between excess market returns and its conditional variance. Yu and Yuan (2011) propose an alternative method to approach the question. They argue that sentiment alters the supposedly positive mean-variance relationship. Taking into account the interaction with sentiment, the positive mean-variance relationship becomes flatter. The implications of the results are two-fold. First, the propensity to take risk is higher during high sentiment periods. Second, investors tolerate lower rate of returns during high sentiment periods given the same level of volatility.

2.5.5.3 Sample period, sampling frequency and sample size

The state variables are time variant and investors adjust their perception of the market variance to synchronise with state variables. If there is an unexpected increase in variance, the investors ask for a higher risk premium and supress prices to a lower level, leading to a negative-mean variance trade-off (French et al., 1987; Schwert, 1989; Turner et al., 1989). This suggests that the level and innovations of variance drive the mean-variance relationship: one will observe positive mean-variance trade-off if the study covers a low variance period and vice versa. For example, the positive mean-variance relationship as formed in Goyal & Santa-Clara (2003) for the period 1963 to 1999 turned out to be negative when the sample period is extended for another two years to 2001 by Bali et al. (2005). Rossi and Timmermann (2010) find that the mean-variance relationship is either flat or negative during the period of high volatility, for example October 2007 and financial crisis during 2007 and 2008; while the mean-variance relationship remains positive during the period of low or normal volatility.

The majority of empirical evidence on return mean-variance relationships opt for daily or monthly data. The choice of the sampling interval is a compromised decision between noise and information content. A shorter interval captures more information and more noise at the same time. In order to rule out the effect of portfolio rebalancing, transaction costs and immediate consumption needs, Harrison and Zhang (1999) analyse data for longer intervals, monthly, quarterly, annually and bi-annual. They find a positive mean-variance relationship for annual and bi-annual data.

2.5.5.4 Linear assumption of the mean-variance trade-off

Among all available methods, ordinary least squares and ARCH class models are widely adopted in the investigation of the mean-variance relationship. These methods assume a linear relationship between the first and second moments of returns. The fact remains that the mean-variance relationship is non-stationary and time varying. Consequently, the linear assumptions lead to a problematic inference of the mean-variance relationship (Chou, 1988; Schwert, 1989; Whitelaw, 1994). Backus and Gregory (1993) state that "It can be increasing, decreasing, flat, or even non-monotonic. The shape depends on both the preferences of the representative agent and the probability structure across states". Rossi and Timmerman (2010) concur with this but propose that there should be positive mean-variance if the risk is properly measured.

2.5.5.5 Other theories

The negative mean-variance relationship is closely related to asymmetric volatility literature. The past negative returns have greater impact on future volatility as opposed to past positive returns. The leverage hypothesis proposes that lower stock prices reduce equity value and lead to higher leverage. Consequently, the stock is riskier with greater volatility (Black, 1976; Christie, 1982; Duffee, 1995). Alternatively, the volatility feedback hypothesis suggests that in case of increased volatility, investors expect higher returns, thus, lowering the stock price. Allowing volatility feedback effects in the model lead to detection of negative risk-return relationships (Campbell & Hentschel, 1992; French et al., 1987). The leverage hypothesis contends that returns lead the changes in volatility while the volatility feedback hypothesis implies that the changes in volatility leads the returns.

Other studies attribute the mixed mean-variance relationship to the behavioural aspects of the investors. Shefrin (2008) attribute the negative risk-return relationship to investor error arising from representativeness and affect heuristics: " stocks of good companies are representative of good stocks". Investors perceive that a reputable company is historically sound and has good future prospects thus being less risky. Investors expect higher returns from a good company. Investors rarely verify their thoughts with facts. These misperceptions elicit a negative correlation between risk and return. Hibbert, Daigler, and Dupoyet (2008) suggest that negative returns instigate fear and positive returns instill confidence. Representativeness, affect heuristics and extrapolation that are biased induce the momentum effect. The change in the implied volatility is negatively associated with returns (contemporaneous and lag terms) and lags of change in implied volatility. The evidence confirms the momentum effects and supports the behavioural explanation of the negative risk-return relationship.

The noise trader explanation suggests that noise traders are more dominant during periods of high sentiment. Rational arbitrageurs are reluctant to bet against the mispricing because the timing of price adjustment is unpredictable. Yu and Yuan (2011) prove that the mean-variance relationship is positive in low-sentiment periods. However, there is a weaker mean-variance relationship during low-sentiment periods.

2.6 SUMMARY AND CONCLUSION

The shortcomings of classical pricing models to explain excessive returns have motivated the burgeoning of investor sentiment literature. Section 2.2 sets out the role of investor sentiment in the financial literature. The review, evaluation, comparison and empirical evidence (Sections 2.2 to 2.5) of the commonly used proxies of investor sentiment inform the selection of sentiment proxies, to answer the first research question: What is the role of media content in explaining the daily index futures market behaviour? The review on theory and evidence of the role of investor sentiment on the mean variance relationship (section 2.6) provides guidelines to the second research question.

While advocates of classical finance theory try to relate the predictability of investor sentiment for market activities to investors' risk consideration, I am convinced that the inherent irrationality of investors carve out the role of investor sentiment in explaining market behaviour. The objective measures (i.e. trading data) are lagged indicators because sentiment is reflected in trading activity. Subjective measures (i.e. newsletter and surveys) are leading indicators, but too costly to conduct and suffer from respondent bias. News sentiment is supposed to reflect future expectations and macroeconomic information. I examine contracts with high trading volume to avoid the non-synchronous trading problem, and employ vector auto regressive models to account for returns' positive autocorrelation and market microstructure noise. This study adds to the investor sentiment literature by examining the reliability of news sentiment in predicting index futures trading activity across three index futures contracts of different sizes and structures. This thesis fills the research gap by investigating whether the general stock market news is able to predict the index futures trading activities.

Section 2.6 reviews the literature that examines the theory and evidence of the role of investor sentiment on the returns mean-variance relationship. Although the risk-return relationship is the tenet of capital asset theories, no consensus has emerged from the extant literature. Previous studies attributed the mixed evidence to volatility model specifications, sampling issues, and omitted conditioning variables. Other studies justify the negative mean-variance relationship with the leverage hypothesis, volatility feedback hypothesis and noise trader theory. In addition, there is empirical evidence on the association between sentiment and volatility. The second research question fills this gap. I add to the discussion of mixed mean-variance relationship by proposing that the investor sentiment alters the mean-variance relationship during periods of extreme sentiment. Previous studies investigate this issue based on monthly data, I use daily volatility and daily sentiment regime, to narrow down the prediction interval.

CHAPTER 3 : HYPOTHESES AND DATA

3.1 INTRODUCTION

Tetlock (2007) uses the General Inquirer and the Harvard IV-4 psychosocial dictionary. The first principal component extracted from the 77 word categories is a linear combination of four word categories: *Negative, Fail*, *Weak* and *Fall*. This is called media pessimism. It is likely that the optimistic word categories outweigh the pessimistic word categories. Henceforth, high investor sentiment is referring to highly optimistic news content and low investor sentiment is referring to highly pessimistic news content. In addition, the news sentiment is recoded into a nominal scale: *Good, Newhigh, Low* and *NewLow*. This study tests the hypotheses pertaining to the relationship of Index futures returns, volume, and volatility with news sentiment, in the context of the Hong Kong, Malaysia, and Singapore index futures markets.

3.2 HYPOTHESIS DEVELOPMENT AND MODEL SPECIFICATION

3.2.1 News sentiment and index futures returns

Financial news serves as an input for investment decisions. Financial columns consistently report what has happened in the stock markets. Reporters may use more optimistic words, when reporting gains in the stock markets. Conversely, readers see more adjectives that are negative in the news when the market performs poorly. Thus, the recent stock market returns may forecast the news sentiment.

Hypothesis 1

H_0: Lag returns do not forecast bad news factors.

H_1: Lag returns negatively forecast bad news factors.

Previous studies emphasise the timing of media sentiment. First, if news sentiment predicts investor sentiment, low media sentiment forecasts low near-term returns and reversals in the longer term. Second, if the news sentiment reflects investor sentiment, one would see low news sentiment after a series of low returns; while the low returns continue in near-term with reversals in future (Tetlock, 2007). The question of whether the news sentiment forecast the investor sentiment or reflects the past investor sentiment resembles the chicken and egg debate. The study suggests that both explanations could be true; news sentiment reflects the past and forecast the future investor sentiment. If news sentiment is a proxy for investor sentiment, we observe low news sentiment after a series of low returns, and low returns persist in near-term, and then reverse to higher returns in the longer term. The same study also proposes that news factors are a proxy for fundamental information that has not been incorporated into prices.

Hypothesis 2

H_0: Bad news factors do not forecast returns.

H_1: Bad news factors negatively forecast returns

The behavioural theory views this as an overreaction of investors and the prices will reverse to fundamental values in the longer-term. Information theory predicts that negative information about future cash flows results in pessimistic sentiment that lower prices. Both theories imply that low sentiment will forecast low short-term returns. However, the information theory implies that prices should reflect new information and the impact will persist, thus no reversal is expected.

Hypothesis 3

H_0: There is no returns reversal as predicted by pure information theory.

H_1: There is returns reversal as predicted by sentiment theory.

In reality, the news may contain new information with some noise. Investors overreact to noise and the prices will reverse later, but the adjustments based on information will persist. It would be realistic to test for the Hypothesis 4.

Hypothesis 4

H_0: There is no returns reversal as predicted by pure information theory.

H_1: There is partial returns reversal as predicted by noise and sentiment theory

If the evidence supports sentiment theory, then Hypothesis 5 will be tested to examine if the prices reverse to their fundamental values.

Hypothesis 5

H_0: The returns reversals do not offset the initial change in returns.

H_1: The returns reversals exactly offset the initial change in returns as predicted by stale information theory.

Stale information theory assumes the news sentiment is a proxy for information that is already incorporated into prices and no-information theory assumes that news sentiment purely reflects noise. Both theories predict that the news sentiment should not have any predictive power over futures returns.

3.2.2 News sentiment and index futures trading volume

Coval and Shumway (2001) examine the relationship between the sound level at the trading floor and volume. When traders are expecting the cost to execute trades to increase in the future, they tend to raise their voice, hope to execute their trade immediately. Based on the same rationale, Tetlock (2007) suggests if the news sentiment is a proxy for trading costs, then high news sentiment should forecast an increase in trading volume. Similarly, Baker and Stein (2004) model the relationship between investor sentiment and volume. When the investor sentiment is extremely high, irrational investors are overconfident and dominate the market, causing the liquidity and trading volume to increase sharply. The opposite happens when the sentiment is low.

Hypothesis 6

H_0: News sentiment does not forecast trading volume.

H_1: Pessimistic news negatively forecasts trading volume.

Campbell, Grossman and Wang (1993) examine the relationship between trading volume and serial correlation in stock returns. The model assumes there are two types of investors. Public information, such as negative news about the stock market can instil negative sentiment and trigger uninformed traders to liquidate their positions. On the other hand, the risk-averse utility maximisers are willing to buy these stocks at lower prices and higher expected returns. Prices increase on subsequent days due to higher expected returns. High trading volumes reflect the buying and selling between the non-information and the risk-averse utility maximisers. Similarly, positive news can instill positive sentiment and trigger sentiment traders to increase their positions. It is logical to claim that the absolute values of news sentiment can positively forecast the trading volume.

Hypothesis 7

H_0: Absolute values of news sentiment do not forecast trading volume.

H_1: High absolute values of news sentiment forecast high trading volume.

3.2.3 Structural vector autoregressive model (SVAR)

This study will employ a structural vector-auto-regression (SVAR) model of index futures returns and sentiment, modified from the VAR model introduced by Tetlock (2007). This model will include the contemporaneous effect (same day) and up to 5 previous days for all variables.

Brown and Cliff (2004) find a contemporaneous relationship between sentiment measures and market returns for weekly data, only one lag of returns is included in the regressions because the autocorrelation in returns is relatively small and dies out quickly. In this study, the rationale for the contemporaneous relationship is about the timing of the market events and news going public. Investors read newspapers in the morning comprehend yesterday's events and form today's sentiment. It is likely that trading decisions are influence by the information gather in the morning and reflect on the trading activity data on the same day.

I estimate a three variables SVAR model to test the hypotheses. The specification of the model as follows:

$$N_t = \alpha_1 + \sum_{i=1}^{5} \beta_{1i} N_{t-i} + \sum_{i=0}^{5} \gamma_{1i} R_{t-i} + \sum_{i=0}^{5} \delta_{1i} V_{t-i} + \sum_{j=1}^{14} \lambda_{1j} EXOG_j + \varepsilon_{1t} \quad (3.1)$$

$$R_t = \alpha_2 + \sum_{i=0}^{5} \beta_{2i} N_{t-i} + \sum_{i=1}^{5} \gamma_{2i} R_{t-i} + \sum_{i=0}^{5} \delta_{2i} V_{t-i} + \sum_{j=1}^{14} \lambda_{2j} EXOG_j + \varepsilon_{2t} \quad (3.2)$$

$$V_t = \alpha_3 + \sum_{i=0}^{5} \beta_{3i} N_{t-i} + \sum_{i=0}^{5} \gamma_{3i} R_{t-i} + \sum_{i=1}^{5} \delta_{3i} V_{t-i} + \sum_{j=1}^{14} \lambda_{3j} EXOG_j + \varepsilon_{3t} \quad (3.3)$$

where

R_t= Daily returns of the Hang Seng Index Futures (HSIF), Kuala Lumpur Composite Index Futures (KLCIF) and Morgan Stanley Singapore Free Index Futures (SiMSCIF) on day *t*.

N_t = News factor score obtain by performing principal component analysis (PCA) for day t. The GI category variables are demeaned by day of week using the prior year mean to ensure the media factor generated in PCA does not systematically capture the day of the week variation in the news. The sources of news are Wall Street Journal Asia, New Straits Times, South China Morning Post and The Straits Times.

V_t= Detrended log volume obtain by subtract a 60-day backward moving average[15] for day t. The volume of HSIF, KLCIF and SiMSCIF are proxy by the number of contract traded and open interest at market close.

$EXOG_j$=Exogenous variables.

Equation (3.1) tests hypotheses 1, I expect the $\gamma_{1i}$to be negative because reporters tend to use more optimistic (pessimistic) words when past returns are high (low). Equation (3.2) tests Hypotheses 2 to 5, in which, I expect β_{2i} to be negative when *i*= 0 and 1, and positive when *i*=3, 4, and 5. This is based on the sentiment argument that highly pessimistic news lead to lower returns, but soon investors will realise that they have overreacted and the prices will reverse.

I estimate a four variables SVAR model to test Hypotheses 6 and 7. The specification of the model as follows:

$$N_t = \alpha_1 + \sum_{i=1}^{5} \beta_{1i} N_{t-i} + \sum_{i=0}^{5} \theta_{1i} |N_{t-i}| + \sum_{i=0}^{5} \gamma_{1i} R_{t-i} + \sum_{i=0}^{5} \delta_{1i} V_{t-i} + \sum_{j=1}^{14} \lambda_{1j} EXOG_j + \varepsilon_{1t} \quad (3.4)$$

$$|N_t| = \alpha_2 + \sum_{i=0}^{5} \beta_{2i} N_{t-i} + \sum_{i=1}^{5} \theta_{2i} |N_{t-i}| + \sum_{i=0}^{5} \gamma_{2i} R_{t-i} + \sum_{i=0}^{5} \delta_{2i} V_{t-i} + \sum_{j=1}^{14} \lambda_{2j} EXOG_j + \varepsilon_{2t} \quad (3.5)$$

$$R_t = \alpha_3 + \sum_{i=0}^{5} \beta_{3i} N_{t-i} + \sum_{i=0}^{5} \theta_{3i} |N_{t-i}| + \sum_{i=1}^{5} \gamma_{3i} R_{t-i} + \sum_{i=0}^{5} \delta_{3i} V_{t-i} + \sum_{j=1}^{14} \lambda_{3j} EXOG_j + \varepsilon_{3t} \quad (3.6)$$

$$V_t = \alpha_4 + \sum_{i=0}^{5} \beta_{4i} N_{t-i} + \sum_{i=0}^{5} \theta_{4i} |N_{t-i}| + \sum_{i=0}^{5} \gamma_{4i} R_{t-i} + \sum_{i=1}^{5} \delta_{4i} V_{t-i} + \sum_{j=1}^{14} \lambda_{4j} EXOG_j + \varepsilon_{3t} \quad (3.7)$$

[15] This is to transform the trading volume and open interest into a stationary time series. Campbell, et al. (1993) and Fung and Patterson (1999) employ the same method, but 100-day backward moving average is used. Fung and Patterson (1999) have also tried with 20-day backward moving average, and yield similar results.

I expect β_{4i} to be negative because highly optimistic (pessimistic) news implies over-confidence, causing the liquidity and trading volume to increase (decrease) sharply. I expect θ_{4i} to be positive because both optimistic and pessimistic news triggers buying and selling between the non-information and the risk-averse utility maximisers that increase trading volume.

Index futures returns, bad news factors and volume are endogenous variables in these two SVAR systems. I calculate close-to-close returns ($\log(\frac{P_{settlement,t}}{P_{settlement,t-1}})$). The volume is measure in number of contracts traded daily and daily open interest. Volume series usually contain trend component and are non-stationary. I employ a simple detrending method to address these issues. I subtract a 60-day backward moving average from the log volume series, which is similar to the geometrically declining average of volume growth rates.

EXOG represents all exogenous variables. The exogenous variables are five lags of the detrended squared index futures residuals to proxy for volatility (j=1 to j=5); dummy variable for Chinese New Year effect (j=6); dummy variable for 1997 Asian financial crisis (j=7); dummy variable for 2008 Wall Street financial meltdown (j=8); dummy variable for January effect (j=9); dummy variables for day-of-the-week effect (Monday to Thursday, i.e. j=10 to j=13); and finally dummy variables for four days prior to settlement date and inclusive the settlement day (i.e. j=14 to j=17).

I calculate the proxy for volatility as follows. Firstly, the variable R_t will be regressed on 12 lags of R_t to obtain a residual. The residual is then squared, and a past 60-day moving average of the squared residual is subtracted from the squared residual[16].

Yen, Lee, Chen and Lin (2001) confirm the existence of a Chinese New Year effects in Hong Kong, Japan, South Korea, Malaysia, Singapore and Taiwan. The study finds a consistent up-moving trend from 15 days before Chinese New Year and lasts up to 15 days after that. A dummy variable will be created and 1 will be assigned to 15 days before and after Chinese New Year, and 0 for other days. Table 3.1 lists the month and date for Chinese New Year from 1995 to 2008.

Table 3.1Chinese New Year Days in the Gregorian Calendar

Year	Month	Date
1995	January	31
1996	February	19
1997	February	7
1998	January	28
1999	February	16
2000	February	5
2001	January	24
2002	February	12
2003	February	1
2004	January	22
2005	February	9
2006	January	29
2007	February	18
2008	February	7

Lean, Smyth and Wong (2007) find evidence of weekday effect and January effect in Hong Kong, Malaysia and Singapore Market. I create dummy variables for Monday, Tuesday, Wednesday, Thursday, and the month of January.

[16] Tetlock (2007) define the trend as 60-day moving average; Ciner (2006) define the trend as 200-day moving average and Bessembinder and Seguin (1992) define the trend as 100-day moving average. These studies conclude that the results are robust to the number of days used to calculate the trend.

The onset of the 1997 and 2008 financial crises had a great impact on equity markets. The 1997 crisis period is refers to the period from August 1997 to December 1997, following Hassan, Mohamad, Ariff and Nasir (2007). In addition, the 2008 financial crisis refers to the period of October 2008 to December 2008, because the equity markets in the samples started to plunge from October 2008.

Trading volume is exceptionally high from about four days before the settlement date. Settlement date and four days prior settlement date equal to 1, and 0 for other days

3.2.4 Estimation of structural vector autoregressive model (SVAR)

Since it is possible that the bad news factor has an impact on the trading activities, contemporaneous terms are specified in the equations (3.1), (3.2) and (3.3). This is a VAR model in primitive form that is similar to the structural form for a simultaneous equations model. However, a *standard* or *reduced* from VAR contains only predetermined values on the right hand side. The VAR can no longer be identified recursively if the equations have a contemporaneous feedback term. The primitive form VAR (or structural VAR) is not identified; therefore identifying restrictions must be imposed so that the structural VAR can be identified. Danielsson and Love (2006) suggest that adding instrumental variables will circumvent this problem. However, more variables will reduce degrees of freedom. Moreover, is not easy to identify an appropriate instrumental variable for news sentiment. Alternatively, Brooks (2008, p. 296) suggests to impose zero restrictions so that the VAR model can be identified.

To determine the coefficients to be imposed zero restrictions, the equations (3.1), (3.2), and (3.3) are written as matrices and vectors. The exogenous variables and deterministic terms are omitted at this stage for simplicity. This will not affect the structural VAR estimation at the later stage.

$$\begin{pmatrix} N_t \\ R_t \\ V_t \end{pmatrix} = \begin{pmatrix} \beta_{11} & \gamma_{11} & \delta_{11} \\ \beta_{21} & \gamma_{21} & \delta_{21} \\ \beta_{31} & \gamma_{31} & \delta_{31} \end{pmatrix} \begin{pmatrix} N_{t-1} \\ R_{t-1} \\ V_{t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \beta_{15} & \gamma_{15} & \delta_{15} \\ \beta_{25} & \gamma_{25} & \delta_{25} \\ \beta_{35} & \gamma_{35} & \delta_{35} \end{pmatrix} \begin{pmatrix} N_{t-5} \\ R_{t-5} \\ V_{t-5} \end{pmatrix} + \begin{pmatrix} 0 & \gamma_{10} & \delta_{10} \\ \beta_{20} & 0 & \delta_{20} \\ \beta_{30} & \gamma_{30} & 0 \end{pmatrix} \begin{pmatrix} N_t \\ R_t \\ V_t \end{pmatrix} + \begin{pmatrix} U_{1t} \\ U_{2t} \\ U_{3t} \end{pmatrix} \quad (3.8)$$

The contemporaneous terms from (8) can be taken over to the left-hand-side and written as

$$\begin{pmatrix} 1 & -\gamma_{10} & -\delta_{10} \\ -\beta_{20} & 1 & -\delta_{20} \\ -\beta_{30} & -\gamma_{30} & 1 \end{pmatrix} \begin{pmatrix} N_t \\ R_t \\ V_t \end{pmatrix} = \begin{pmatrix} \beta_{11} & \gamma_{11} & \delta_{11} \\ \beta_{21} & \gamma_{21} & \delta_{21} \\ \beta_{31} & \gamma_{31} & \delta_{31} \end{pmatrix} \begin{pmatrix} N_{t-1} \\ R_{t-1} \\ V_{t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \beta_{15} & \gamma_{15} & \delta_{15} \\ \beta_{25} & \gamma_{25} & \delta_{25} \\ \beta_{35} & \gamma_{35} & \delta_{35} \end{pmatrix} \begin{pmatrix} N_{t-5} \\ R_{t-5} \\ V_{t-5} \end{pmatrix} + \begin{pmatrix} U_{1t} \\ U_{2t} \\ U_{3t} \end{pmatrix} \quad (3.9)$$

The equation (9) can be rewrite in the form of AB-model introduced by Amisano and Giannini (1997).

$$Ay_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_5 y_{t-5} + B\varepsilon_t \quad (3.10)$$

Where y_t is the *k*-element vector of the endogenous variables in which *k* = 3 in this study; A and B are invertible matrices of dimension 3x3; ε_t is a 3x1 vector containing the structural unobservable structural disturbances; u_t is the reduced (or observed) form disturbances. The contemporaneous relationship among the endogenous variables can be estimated through the A matrix. The B matrix captures the structural shocks that enter the system.

The reduced form of a VAR model can be obtained by multiply both sides by A^{-1}.

$$\begin{aligned} y_t &= A^{-1}A_1 y_{t-1} + A^{-1}A_2 y_{t-2} + \cdots + A^{-1}A_5 y_{t-5} + A^{-1}B\varepsilon_t \\ &= A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_5 y_{t-5} + u_t \quad (3.11) \end{aligned}$$

where

$A_1 = A^{-1}A_1, \dots, A_5 = A^{-1}A_5$ and

$$u_t = A^{-1}B\varepsilon_t \tag{3.12}$$

Now the equation (3.11) only consists of predetermined values (lag variables) on the right-hand-side. The relationship between the underlying unobserved shocks ε_t, and the observed disturbance, u_t, is explained in (3.12). The structural innovations ε_t are assumed to be orthonormal, such that $E[\varepsilon_t \varepsilon_t'] = I$ .This assumption enables A and B to be identified by imposing the restriction:

$$A \sum A' = BB' \tag{3.13}$$

The symmetry of both sides of (3.13) imposes k(k+1)/2 (equals to 6 for a three endogenous variables VAR) restrictions on the 2k^2 (equals to 18 for a three endogenous variables VAR) unknown elements of A and B. At least 12 additional restrictions need to be supplied (i.e. 18-6). In this study, R is assumed to have no contemporaneous effect on NEWS, and VOLUME is assumed to have no contemporaneous effect on NEWS and R, in such that δ_{10} and γ_{10}, and δ_{20} are assumed to be zero. Put together the assumptions of ε_t are orthonormal, I impose the following restrictions:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ NA & 1 & 0 \\ NA & NA & 1 \end{bmatrix}, \quad B = \begin{bmatrix} NA & 0 & 0 \\ 0 & NA & 0 \\ 0 & 0 & NA \end{bmatrix}$$

The equations (3.1), (3.2), and (3.3), as a structural VAR system, the endogenous variables are ordered from the most exogenous to least exogenous i.e. *N, R and V*. N_t is available in the morning of day *t* before the markets are open; while R_t and V_t are available after the markets are opened for day *t*. Thus, R_t and V_t have no predictability power over N_t.

The lead-lag relationships between return and volume are well documented. Statman, Thorley and Vorkink (2006) find a positive relationship between aggregate market turnover and lagged market returns.

The implication of the restrictions imposed on the A and B matrices are three-fold. First, the N_t has no contemporaneous interaction with R_t and V_t and the unobserved structural shock is equivalent to the observed reduced-form shock. Second, the R_t is assumed to be contemporaneously affected by itself and the structural shocks from the first equation. Third, the V_t is assumed to be affected by its own lags, and contemporaneously interacting with the first two structural shocks.

Based on the same justification from equation (3.8) to equation (3.13), I impose restrictions for equation (3.4) to (3.7) in the form of 4x4 matrices.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ NA & 1 & 0 & 0 \\ NA & NA & 1 & 0 \\ NA & NA & NA & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} NA & 0 & 0 & 0 \\ 0 & NA & 0 & 0 \\ 0 & 0 & NA & 0 \\ 0 & 0 & 0 & NA \end{bmatrix}$$

3.2.5 News sentiment and the index futures mean-variance relationship

This study intends to test whether investor sentiment is priced and has a role in the risk-return trade off. It should be made clear that there is no intension to evaluate the mean-variance model from the perspectives of different volatility models, mean-variance model specifications, sampling issues and linearity assumptions of the mean-variance trade-off as discussed in section 2.6.5. Specifically, I argue that noise traders dominate the market during high sentiment periods and introduce noise trader risk that should be priced. In addition, high investor sentiment attenuates the otherwise positive mean-variance relationship, or reverses it to become negative as explain by prospect theory. This study fundamentally incorporates the role of sentiment and provides a basis for the discussion of mixed mean-variance relationships.

I first assume that investors are prone to cognitive biases instead of being fully rational. These biases may arise throughout the process of forming beliefs and preferences. People are subject to cognitive biases, for example, overconfidence, optimism, representativeness, conservatism, belief perseverance, anchoring, availability biases, mental accounting, and the framing effect (Barberis & Thaler, 2003; Ritter, 2003).

Second, the cognitive biases lead to a bias expectation for risk and return among sentiment investors. Campbell and Kyle (1993) suggest that excessive volatility of stock prices do not attribute to fundamentals; IPOs (Ritter, 1991) and close-end funds are sold at discount (Chopra et al., 1993; Lee et al., 1991).

Third, there are more noise traders in the market who trade actively when the sentiment is relatively high or low. DeLong et al.(1990b) suggest that sophisticated investors turn passive when sentiment traders dominate the market. Prices become volatile due to lack of arbitrage activities against the noise traders.

The above assumptions imply that there is a role for investor sentiment in asset pricing and mean-variance relationships. During the period of high sentiment, sentiment investors overreact to trading-induced good news and bid the price even higher. They buy at a high price and thus lower the required return. Regardless of the impact of sentiment on risk, sentiment investors are simply willing to be compensated with lower returns. Hence, this study attempts to test the following hypotheses:

Hypothesis 8

H_0: There is no mean-variance relationship

H_{1a}: There is positive mean-variance relationship when create space effect outweigh the Friedman effect.

H_{1b}: There is negative mean-variance relationship when Friedman effect outweigh the create space effect.

The Merton (1973) model suggests that investors demand higher returns for investing in riskier assets. A positive mean-variance relationship implies a risk-return trade off. However, there is mixed evidence for mean-variance relationships. DeLong et al. (1990b) attribute this to noise traders misperceiving the risks of holding the risky assets. A greater shift in sentiment is associated with greater future returns volatility (higher risk) and lower expected returns. On one hand, noise traders have poor market timing, follow the footsteps of other noise traders, and end up buyinghigh and-sellinglow, eventually earning poor returns. This is named as 'Friedman' effect; noise traders' returns are negatively related to the variability of their beliefs. On the other hand, risk-averse arbitrageur will avoid betting on noise traders' mispricing when there is high variability in noise trader beliefs. This is the so-called 'create space' effect. Volatility is positively related to expected returns. In conclusion, the net impact of volatility on excess returns is positive when the create space effect outweighs the Friedman effect and vice versa. I test Hypothesis 8 using the coefficient, β_1, in Equation 14a and 14b.

Hypothesis 9

H_0: Sentiment does not predict excess returns

H_{1a}: High sentiment positively (negatively) predicts the returns when hold-more effect (price pressure) dominate the price pressure (hold-more) effect.

H_{1b}: Low sentiment negatively predicts the excess returns because price pressure and hold-more effect are negative.

DeLong et al. (1990b) explain the impact of noise traders on asset pricing. The direct short-term impact arises from 'price pressure' and the 'hold more' effect. Bullish noise traders demand more stocks. Greater demand drives the stock prices to higher levels and expected returns are lower. Consequently, expected returns are negatively related to investor sentiment. This is called price pressure effect. At the same time, bullish irrational investors hold more risky assets than rational arbitrageurs during the period of high sentiment, higher expected returns is a form of reward for noise traders bearing the risk. I expect a positive relationship between sentiment and returns. This is called hold more effect. These two effects interact to determine the impact of sentiment on returns. If the hold-more effect dominates the price pressure effect, one would expect bullish sentiment to lead to higher returns. When the price pressure effect outweighs (weaker than) the hold-more effect, bullish sentiment predicts lower (higher) returns. The impact of the price pressure effect and hold more effect is always negative when investors are bearish. If the $\boldsymbol{\alpha_2}$in the Equation 14a and 14b is significant, the required return for each unit of risk increases or decreases at a fixed amount at all levels of risk, depending on the magnitude of price pressure and hold more effect.

Hypothesis 10

H_0: Sentiment does not change the sensitivity of return to risk.

H_{1a}: Investors are less responsive to risk in the high sentiment regime.

H_{1b}: Investors are more responsive to risk in the low sentiment regime.

Alternatively, investor sentiment can *interact* with risk and change the sensitivity of return to risk as suggested by Yu and Yuan (2011). This study proposes that investors are less responsive to risk during the high sentiment period. High sentiment following market run-up weakens the risk-return relationship. I extend this thought to the low sentiment period. Sentiment investors overreact to bad trading-induced news and take the risk cautiously. Low news sentiment interacts with risk and makes the risk-return slope steeper. The investors become more sensitive to risk during the low sentiment period. I expect β_2 in the Equation 14a and 14b to be negative in a high sentiment regime and to be positive in a low sentiment regime.

I test the Hypothesis 8,9, and 10 based on the following equations in line with Yu and Yuan (2011):

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \tag{3.14a}$$

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \tag{3.14b}$$

where the dependent variable is the daily excess index futures returns of the HSIF, KLCIF and SiMSCIF.

Excess returns are defined as the daily returns less the risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate defined as the three-month treasury bill discount rate *i* for the HSIF and SiMSCIF and the one-month Kuala Lumpur interbank offer rate for the KLCIF. Var_t is the conditional variance while V_t is the realised variance. *Sentiment* is a dummy variable that refers to sentiment measures namely *Bad, Newlow, Good, Newhigh*. *Bad* = 1 if the daily routine news reports the market fell on the prior day; *Newlow*=1 if the market dipped to a new low; *Good*=1 if the market rose; *Newhigh*=1 if the market hit a new high; otherwise, 0. These sentiment measures apply to the HSIF, KLCIF and SiMSCIF with two exceptions. For the case of the HSIF, *Lowbench* is used instead of *Newlow* while *Highbench* is used instead of *Newhigh*. *Lowbench* =1 if the HSI fell to a lower benchmark and

Highbench=1 if HSI rose and hit a higher benchmark. Section 3.3.5 explains how to derive sentiment regime from daily news.

3.3 DATA AND SAMPLE

3.3.1 Data and sample period

Hong Kong, Singapore and Malaysia index futures will be included based on the following sample selection criteria and justifications. First, the index futures contracts must be actively traded in these countries. Second, English literacy in these countries enable news sentiment to be consistently derived by using an artificial intelligence coding system. This is essential to make meaningful comparisons among countries. The Hang Seng Index has been the major performance benchmark for the Hong Kong equity market since 1969. It is weighted by market capitalisation of 33 constituent stocks. Hong Kong Futures Exchange introduced the Hang Seng Index Futures (HSIF) contracts in May 1986, in order to meet the growing needs of hedging and portfolio management tool. The HSIF soon became one of the most traded index futures contracts in Asia. The HSIF recorded a total trading volume of 21,716,508 contracts in the year 2008. In the same year, the range of average daily trading volume was from 74,311 to 111,456 contacts[17].

There are two major market performance benchmarks in Singapore, namely the Straits Times Index (STI) and the MSCI Singapore Free Index (SiMSCI). The STI consists of 30 top performing stocks, and it is weighted by market capitalisation. The constituent stocks usually made up 5 to 10% of the overall market capitalisation. The Morgan Stanley Capital International constructs the SiMSCI. The index consists of 36 stocks from five sectors, namely financials, industrials, telecommunications services, consumer discretionary and consumer staples. Futures contracts of these indexes are electronically traded on the Singapore Exchange

[17] HKEx Fact Book 2008, http://www.hkex.com.hk/data/factbook/2008/fb2008.htm

Derivatives Trading (SGX-DT). The STI is a more popular performance benchmark relative to the SiMSCI. However, the SiMSCI futures contract is more actively traded as compared to the STI futures contracts. In year 2008, the total trading volume for the SiMSCI futures was 4,635,517 contracts with only 2,734 contracts traded for the STI futures contract.[18] This research is confine to SiMSCI futures contract that has been traded since 7 September 1998.

The Malaysian derivatives market is relatively small as compared to the Hong Kong and the Singapore markets. The benchmark index is the market-capital weighted Kuala Lumpur Composite Index (KLCI) with 100 constituent stocks from the main board. The KLCI futures contract was introduced in December 1995. Total trading volume for year 2008 was 2,920,728[19].

This study covers the 1997 and 2008 crises. Since the three index futures contracts were introduced at different times, the sample period varies between markets. The data for the Hang Seng Index Futures and the Kuala Lumpur Composite Index Futures span 13 years, from 3rd January 1996 to 31st December 2008. The data for the Singapore Morgan Stanley Free Index spans 11 years, from 3rd January 1998 to 31st December 2008. I obtain daily trading data on spot index and index futures from Datastream; including HSI, SiMSCI, KLCI, HSIF, and SiMSCIF. I collect data for KLCIF directly from Bursa Malaysia because Datastream does not compile data for the full sample period. The 5-minute interval data are retrieved from Thomson Reuters Tick History. Only the spot month contracts are actively traded in these markets, thus only spot month contracts will be included in this study. I assume that traders rollover the spot contracts to the next month contract as the spot month contracts are matured.

---

[18] More detail information can be found in http://www.sgx.com/wps/portal/marketplace/mp-en/products/derivatives_products/equity_index/sgx_simsci
[19] Calculate based on data from Datastream.

3.3.2 Source of daily news

I retrieve daily routine financial news from the Lexis-Nexis Academic Universe news archive, the Proquest News Stand, Factiva and the News Bank. The sources of the news are selected based on readership and creditability. I choose the English press from Hong Kong, Singapore and Malaysia with highest readership.

These routine financial news sources are like summaries of market activity. News writers will compose the news content based on early trading sessions when these markets are closed. Some information, for example, prices and volume, will be made available through real time trading data services before the news is published in the press. In addition, the news may be updated to the press's website before the news is published. Hong Kong, Singapore, and Malaysia are in the same time zone.

3.3.3 Deriving news factors from daily news

This section explains how I convert the qualitative news articles into quantitative data that enable hypotheses testing by econometrics models. This involves a few steps. First, the daily news article will be compiled, arranged in ascending order by date in an excel file. Second, the file containing news articles is uploaded to a server based General Inquirer[20] (GI). The GI will classify all words in each news articles individually into 77 word categories, according to the Harvard IV psychosocial dictionary. Third, I perform principle component analysis (PCA) to capture the word categories that explain most of the variations in the news content. This is to avoid including too many variables in the hypotheses testing, which will reduce the degrees of freedom of the test and to mitigate multicollinearity problems among right hand side variables. Fourth, factor scores are generated based on the factor coefficients

[20] http://webgi.stone-center.eu/

estimated from the principal component analysis. These factor scores will become proxies for news sentiment. The whole process is discussed in detail below.

3.3.3.1 Comparisons of content analysis methods

Deffner (1986), Rosenberg, Schnurr and Oxman (1990) and Morris (1994) categorise the content analysis methods into three broad categories. The first category is individual word count systems that divide a single text into word categories and perform rudimentary word count. Second, the artificial intelligence computerised systems that are able to map words or phrases into categories, according to key meaning of surrounding words in the context. The third category is the human-score and phrase-based content analysis systems. The General Inquirer (GI) program, which assigns tags to text words based on the Harvard III psychosociology dictionary, belongs to the first category. A later version of the General Inquirer, using the Harvard IV psychosociological dictionary falls under the second category. A brief discussion of these methods justifies the content analysis tools that I use in this study.

Human-scored systems. Scorers are trained to assign text into specific categories according to coding rules. Reliability of this method is examined by comparing the consistencies of coding of the same text into categories by multiple scorers. Inconsistencies suggest coding rules need to be revised to improve the reliability of coding. The method in Gottschalk , Winget, and Gleser (1969) used to be popular in psychological research; the context, syntax and lexicon will jointly determine the category a word belongs to. Researchers need to create coding rules that meet their own research needs. Comparison among studies is limited because the coding rules are not standardised. Rosenberg et al. (1990) comment, "Hand-scored system is almost by necessity, difficult and expensive to utilise. Only a small number of reliably trained scorers are typically available at any point in time and, even if funding is non-problematic, these scorers may not be equally accessible to researchers from different institutions." Given these limitations, this method is not a good choice for longitudinal financial data.

Individual word count systems. These systems perform exactly the same task as the human score task systems. The systems strictly follow the predetermined coding rule; there are no subjectivity problems as compared to the human-score systems. The superior reliability and coding speed is the main reason for its popularity. However, this primitive rudimentary word count and classification method does not consider the context; for example, a word may not carry the same meaning in different sentences.

Artificial intelligence computerised systems. This is an improvement of rudimentary individual word count systems. "Please" could carry different meanings in different contexts. "Could you *please* do me a favour?" and "They work hard to *please* their boss" illustrate the word "please" is context-sensitive. An artificial intelligence computerised systems can differentiate words according to context. It considers the grammatical arrangement of words in a sentence and takes into account the lexicons when paring the words with category tags. Weber (1990) claims this can enhance the coding accuracy but Rosenberg et al. (1990) finds artificial intelligence systems to have no advantage over individual word count systems.

I use the General Inquirer to perform content analysis for four reasons. First, the current GI version sets the Harvard IV psychosocial dictionary categories, Laswell Dictionary categories, and Semin and Fieldler social cognition categories as default dictionaries. In addition, the GI users can develop their own word categories. These dictionaries are well developed and are being tested over time. This study adopts the Harvard IV-4 dictionaries for its relevance to this study. The dictionary consists of 77 word categories[21]. Second, the new GI is context sensitive in that it is able to handle words according to syntax and lexicons. However, the GI cannot distinguish certain combinations of words in a sentence. Tetlock (2007) gives simple yet clear examples of the limitation: “No, the economy is not strong.” and “It is not that the economy is not strong”. These sentences mean the opposite but the GI will tag these two sentences as if the meanings are the same. Although artificial intelligence systems are designed to mitigate this problem, Rosenberg et al. (1990) finds that it has no advantage over word count system. Third, GI can perform analysis on a large amount of data in reasonable time. This research covers three markets, from 1995 to 2008, approximately 250 news entries per year. Fourth, the use of the same word categories permits cross-comparison and contributes to cumulative research. An example of research in the same strand is Tetlock (2007), who uses the Harvard IV-4 psychosocial dictionary to derive news sentiment from the Wall Street Journal’s (US edition) “Abreast of the market” column.

3.3.3.2 How General Inquirer works

Figure 3.1 provides an example of news as an input to General Inquirer. Most of the news are short, unless there was an unusual event that happened on the day before publication. Generally, the routine financial news will summarise the market movements for yesterday, which of the counters went up or down, and fundamentals or sentiment factors that are deemed to affect stock market valuation. From time to time, market analysts or brokerage houses may give advice about future market movements, even what to buy or sell. The sample news gives an impression that the market condition is somewhat positive. Note the term “somewhat” is ambiguous and subjective. A standard coding rule is needed to systematically measure the degree of sentiment implied by words used in the news content.

[21] The detail explanations of these 77 word categories are available at http://www.wjh.harvard.edu/~inquirer/homecat.htm

This is important because variation in coding rules will introduce variation that is not relevant to the news sentiment and trading itself.

Figure 3.1 Sample of News as Input to General Inquirer
This news article is retrieved from Lexis-Nexis Academic Universe. The news contains title, date, author name, body, section, length, and date loaded to the database. Keywords are highlighted in red.

The Straits Times (Singapore)

Christmas cheer for **stock** investors

December 24, 1999

Edna Koh

IN A breakaway from typical Christmas behaviour, the bourse yesterday ended in record territory due to buying interest in a few key **stocks.**

Situational interest in the bluest of blue chips, namely **stocks** such as Singapore Airlines and Singapore Press Holdings, sent the **Straits Times Index** to an all-time high of 2,424.15, up 40.49 points or 1.35 per cent.

Although the market traditionally winds down in the run-up to the yuletide season, not everyone was surprised by yesterday's break-through.

"It's precisely because it's a quiet week that when a few **stocks** move up, prices are moved along," said Vickers Ballas research head Timothy Wong.

Analysts said that the buying may have come from Asean fund managers, particularly those from Hong kong, as most others from Europe or the US have closed their books.

Most of yesterday's top gainers had also rallied to levels out of easy reach of the typical punter. Top gainer was Pacific Century Regional Developments.

The counter, the parent of Hongkong-listed Pacific Century CyberWorks, jumped a heart-stopping $ 4.40, or 30.6 per cent, to $ 18.80 on speculation that the unit was in talks for a **share** exchange with Microsoft Corp. This was denied by the company.

Contract manufacturers like JIT Holdings and Omni Industries also rose 36 cents and 20 cents each to $ 4.80 and $ 2.80 respectively following Nasdaq's record close on Wednesday.

Topping the actives list was L&M Group Investments on 23.9 million **shares** traded. The Indonesian-linked counter rose 14 cents to $ 1.26 on talk that the company was launching a new technology-related venture.

Turnover rose to 321.22 million **shares** worth $ 563.54 million from the previous day's 282.63 million **shares** valued at $ 474.11 million. Gainers led losers by 238 to 78 with another 182 counters closing unchanged. -Edna Koh.

SECTION: Money; Market Report; Pg. 38
LENGTH: 290 words
LOAD-DATE: December 24, 1999
LANGUAGE: ENGLISH

Copyright 1999 The Straits Times Press Limited

Figure 3.2 Sample of General Inquirer output based on Harvard IV Psychosocial dictionary

This is a sample output generated from Internet General Inquirer, http://www.lbuse.umd.edu:9090/, which only can process article one by one. This study will use sever-based General Inquirer that can simultaneously handle large amounts of data.

tag	N	%	words
Pos	9	3.1	CHEER=1 EASY#1=1 INTEREST#1=2 PRECISE=1 RALLY=1 TRADITIONAL=1 VALUE#2=1 WORTH#2=1
Pstv	7	2.41	EASY#1=1 INTEREST#1=2 PRECISE=1 TRADITIONAL=1 VALUE#2=1 WORTH#2=1
Affil	4	1.38	CHEER=1 CLOSE#1=1 PARENT=1 US=1
Virtue	5	1.72	EASY#1=1 PRECISE=1 TYPICAL=2 WORTH#2=1
Pleasur	1	0.34	VALUE#2=1
Neg	4	1.38	CLOSE#2=2 DENY=1 LOSER=1
Ngtv	4	1.38	CLOSE#2=2 DENY=1 LOSER=1
Hostile	1	0.34	DENY=1
Pain	2	0.69	BLUE#1=2
Strng	19	6.55	ACTIVE#1=1 COMPANY#1=2 CONTRACT#1=1 DEVELOPMENT=1 HIGH#1=1 INDUSTRY=1 LAUNCH=1 MANAGER=1 MANUFACTURER=1 MOST#1=1 MOST#2=1 MOVE#1=2 NEW#1=1 ROSE#1=3 TRADITIONAL=1
Power	3	1.03	MANAGER=1 MAY#1=1 PRESS#2=1
Weak	4	1.38	FEW#1=2 FOLLOW#1=1 LOSER=1
Subm	1	0.34	FOLLOW#1=1
Actv	29	10	ACTIVE#1=1 BUY#1=2 CHEER=1 CLOSE#2=2 COME#1=1 DEVELOPMENT=1 END#2=1 EXCHANGE#1=1 HAVE#1=1 INVESTMENT=1 INVESTOR=1 JUMP#1=1 LAUNCH=1 LEAD#2=1 MOVE#1=2 RALLY=1 RESEARCH=1 ROSE#1=3 SENT=1 SURPRISE#2=1 TALK#2=2 TRADE#2=1 VENTURE#1=1
Psv	4	1.38	FOLLOW#1=1 INTEREST#1=2 QUIET#1=1

Figure 3.2 shows the General inquirer matching each word in the news to the Harvard IV psychosocial dictionary word tags[22]. One word can be matched with more than one tag. For example, the word "rose" in the sample news carry the meaning of "increased in value". In fact, the General Inquirer is able to differentiate the word "rose" in three contexts. First, the "rose" is used as a verb and means increased or expanded. Second, the "rose" is used as a verb and means returned to life. Third, the "rose" is used as a noun and means the flower. In the sample news, the word "rose" falls under the first context. Based on the Harvard IV psychosocial dictionary, "rose" can be matched with three word tags (or categories): strong, active and rise. From Figure 3, the word "rose" is paired with the tag "strong" (label as Strng)

[22] The Harvard IV dictionary word categories or word tag is available at http://www.wjh.harvard.edu/~inquirer/homecat.htm

and the tag "active" (label as Actv). Under the fourth column, "rose#1=3" indicates the word 'rose' has been repeatedly used for 3 times in the first context. The General Inquirer generates the word count fall under each tag and scales it into a percentage. This is important so that comparisons can be made between texts of different words length. Output for the sample news shows 19 words or 6.55% of words that fall under "strong".

I demeaned these Harvard VI psychosocial word categories by prior year's mean of each day-of-the-week respectively. This is to minimise the possibility of factor analysis capturing the day-of-the-week systematic patterns, which will affect the factor selection.

3.3.4 Aggregation of news factors using principal component analysis

I employ Principal Component Analysis (PCA) to derive news sentiment from the demeaned GI output. This factor analysis method can reduce a large number of word categories to a small number of factors. The PCA will derive a news factor by linearly combining the above demeaned word categories that explain most of the variation among all the 77 word categories. Each of the word categories are equally weight in the calculation of variance. Word categories that are highly correlated will be combined to become one factor (or component), that will be used as proxy for news sentiment. Each principal component is orthogonal on each other. I select the principal component with the highest eigenvalue for further analysis.

The writing style of the writer may change over time and lead to different patterns of word usage in news content; time variation has to be taken into consideration. Tetlock (2007) derives the media sentiment proxy for year *t* by fitting factor loadings estimated from a sub sample of year *t-1* into word counts of year *t*. On one hand, this ensures that the construction of the media sentiment proxy is based on known information. For example, if word categories from January to December of year *t* are used to estimate factor coefficients; fitting these factor coefficients to word counts for the month of January in year *t* is based on unknown information from the months of February to December of the same year. On the other hand, the data set will suffer from loss of data for one year. However, the one-year information loss can be avoided since the same study also proves that the factor loadings are stable over the years. Thus, in this study, the PCA will be used to estimate factor coefficients for year *t*. I use the factor coefficients and word counts from the GI to produce a news sentiment measure for year *t* itself. To examine the reliability of this procedure, I generate factor coefficients from each yearly sub-samples and fit into the Harvard IV word categories generated by the GI for all yearly sub-samples. Then I examine the correlations among news factors generated from factor coefficients of different years. The factor coefficients are robust to time variation if these news sentiments factors are positive and highly correlated.

3.3.5 Deriving sentiment regime from daily news

I retrieve routine daily financial news from Factiva. Hong Kong's South China Morning Post, Malaysia's New Straits Times and Singapore's The Straits Times are chosen for the large circulation numbers each has. Donaldson and Kim (1993) find that the Dow Jones Industrial Index is subject to invisible psychological support levels or resistance barriers. I read the headlines and recode the news into dummy variables, which represent high or low sentiment based on the content. For example:

"Caution remains before results of heavy weights; Hang Seng Index dives again ahead of interim earnings from Hutchison and Cheung Kong."

South China Morning Post, August 25, 2005, BUSINESS POST; Markets Report; Pg. 16, 643 words, Fiona Lau

"HSI drops below 15,000 as funds slash holdings; Index loses more than 1,000 points for second time in 3 days"

South China Morning Post, October 11, 2008 Saturday, BUSINESS; Pg. 1, 537 words, Wong Ka-chun

"US recession fears shake markets; HSI dives to 5-month low as Bush's economic rescue plan fails to reassure investors."

South China Morning Post, January 22, 2008 Tuesday, NEWS; Pg. 1, 483 words, Nick Istra

"Surging China plays extend share rally; HSI keeps winning run even as investors cash in property gains."

South China Morning Post, July 22, 2005, BUSINESS POST; Markets Report; Pg. 14, 494 words, Anette Jonsson and Fiona Lau

"Hang Seng Index cracks 22,000 level."

South China Morning Post, June 23, 2007 Saturday, BUSINESS; Pg. 1, 156 words, Wong Ka-chun

"Stock market reaches record high; Investor confidence soared recently as the Hang Seng Index was at its highest level ever, but caution has been advised"

South China Morning Post, October 15, 2007 Monday, YOUNGPOST; liberal studies; Pg. 6, 1009 words, Elaine Yau

The first to sixth headlines are recoded as *Bad, Lowbench, Newlow, Good, Highbench* and *Newhigh* respectively, 0 otherwise. The first recode was completed from 4th August 2010 to 19th August 2010 and the second recode was performed from 29th November to 7th December 2010 to examine the consistency of human coding at a different time.

3.3.6 Measures of volatility

I calculate two measures of volatility from each of these three categories: historical volatility, realised volatility, and conditional volatility. These has enable the examination of the consistency of the mean-variance relationship across various volatility models

3.3.6.1 Historical and Realised volatility

The selection of measures of latent actual volatility depends on its unbiasness, consistency and forecast ability. Common measures of volatility are sample variance and sample standard deviation. Both formulae measures the deviations from the sample mean. However, Figlewski (1997) argues that sample mean is not an unbiased estimate for true mean and suggests that deviation around zero yields higher accuracy. Thus, I calculate squared daily returns as a measure of daily historical volatility (V_t). This is commonly used before the intraday high frequency data was made available[23].

$$V_t = R_t^2$$

I have also compute rolling window volatility (RW_t) using 30-days moving average as another proxy for daily volatility.

$$RW_t = (R_t - \frac{1}{30-1}\sum_{s=t-29}^{t} R_s)^2$$

Nonetheless, Anderson and Bollerslev (1998b) suggest that squared returns are a noisy measure, and propose that high-frequency intraday squared returns reduce the noise. Poon and Granger (2003) conclude that intervals shorter than 5 minutes are highly serially correlated due to market microstructure. As a result, I calculate the sum of squares of intraday returns in 5-minute intervals[24] (*SSR*).

$$SSR_t = \sum_{i=1}^{n} R_i^2$$

where *n*=the total number of 5 minutes interval for a particular trading days.

[23] For example Brooks (1998), Day and Lewis(1992), and West and Cho(1995).

[24] The logarithm of realised volatility as in log form it is much closer to normal distribution (Andersen, Bollerslev, Diebold, & Ebens, 2001), but the log of volatility estimates (usually <0) is always negative, this makes the interpretation become difficult (e.g. negative risk-return relationship will become positive if using log volatility), thus I do not transform the volatility measures into log form.

The value of *n* varies across the three markets due to several changes of trading time over the sample period. On average, *n*=53, 74 and 80 for HSIF, KLCIF and SiMSCIF respectively[25].

The squared returns underestimate the volatility if the closing price moves back near the opening price after experiencing significant fluctuation within the day. Cumulative absolute returns can reasonably deal with this issue (Andersen & Bollerslev, 1998a). I calculate intraday average absolute returns on a daily basis (*AAR*).

$$AAR_t = \frac{1}{n}\sum_{i=1}^{n}|R_i|$$

3.3.6.2 Conditional volatility

I adopt GARCH(1,1) (Bollerslev, 1986) to model conditional variance. I also estimate TGARCH (1,1), which considers the asymmetric impact of positive and negative shocks (Glosten et al., 1993).

The model specification of GARCH(1,1):

$R_t = \alpha_0 + \varepsilon_t$

$\varepsilon_t \sim N(0, h_t)$

$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1}$

The model specification of TGARCH(1,1)

$R_t = \alpha_0 + \varepsilon_t$

$\varepsilon_t \sim N(0, h_t)$

[25] See the Appendix A for detail information on trading hours.

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \varepsilon_{t-1}^2 I_{t-1} + \beta_3 h_{t-1}$$

where R_t is the daily index futures returns; h_t is the conditional variance; I_{t-1}=0 if $\varepsilon_{t-1} \leq 0$ and I_{t-1}=1 if $\varepsilon_{t-1} > 0$.

I am unable to decompose the media sentiment into rational components based on fundamentals and irrational components which are merely noise as in Verma et al. (2008) because the sentiment is a binary measure (1=optimistic , or 1 = pessimistic, 0 otherwise). In addition, the media sentiment is generated from routine financial news that summarises the previous day's market activity. The information contain in these news are historical information and do not reflect future expectation.

CHAPTER 4 : EMPIRICAL RESULTS – THE ROLE OF INVESTOR SENTIMENT IN EXPLAINING THE INDEX FUTURES RETURNS AND TRADING VOLUME

4.1 INTRODUCTION

This section begins with the generation of a bad news factor and the examination of its reliability, and is then followed by descriptive statistics of the variables. The statistical properties of the variables will be taken into consideration during the estimation process. Finally, the analysis results are summarised.

I obtain daily trading data of the Hang Seng Index Futures (HSIF), the Kuala Lumpur Composite Index Futures (KLCIF) and the Singapore Morgan Stanley Free Index Futures (SiMSCI) from DataStream. This study only includes the most actively traded nearest contracts, in order to mitigate the nonsynchronous trading problem. DataStream provides time series of spot month contract, and switch over on the 1st day of new trading month[26].

The exchanges are closed during public holidays, and there is no trading news to be released on the next day of a public holiday. Thus, a one-day public holiday will lead to loss of data for 2 days. The analysis for HSIF and KLCIF span 1996 and 2008, but I include trading data and news articles from 1995 to meet the need for trend adjustments. The analysis of SiMSCIF span from 1999 to 2008, and trading data and news article are collected from 1998 for the same reason.

[26] HSIF and SiMSCIF switch over on last trading day of the nearby contract month, i.e. the second last business day of the month. KLCIF switches over on the last trading day of the nearest contract month, i.e. the last business day of the month. Carchano and Pardo (2009) construct the continuous futures return series by using five different rollover methods and find no significant difference among these returns series. The same study also points out that the simplest rollover method will reach similar conclusion as some complex methods.

The returns, *R*, is defined as the log difference of the settlement price[27]. The trading volume and open interest depicts growing trends. Hence, the trading volume, *V* is defined as the log value of the number of contracts traded (or open interest) each day less the prior 60 days moving average of the same variable.

After all detrending and adjustments for holiday and weekends, the sample size for HSIF, KLCIF and SiMSCIF is 3114, 3102 and 2359 days respectively[28]. Each year consists of about 230 to 240 data points.

4.2 GENERATING SENTIMENT NEWS FACTORS USING PRINCIPAL COMPONENT ANAYSIS

This section discusses the principal component generation using principal component analysis from the 77 Harvard IV word categories. The PCA reduces a large number of word categories into a small number of factors by linearly combining the word categories that explain most of the variation among the 77 word categories. The results indicate that only one or two news factors are generated on a yearly basis for each of the New Straits Times, the South China Morning Post and The Straits Times.

[27] It is unknown how the Datastream compiles the daily settlement price. The HSIF final settlement price is the average of quotation of the Hang Seng Index taken at five-minute intervals during the last trading day. The KLCIF settlement price is the average value, rounded to the nearest 0.5 of an index point (values of 0.25 and 0.75 and above being rounded upwards) of the KLCI for the last half hour of trading on Bursa Malaysia Securities Berhad on the final trading day excepting the highest and lowest values. The final SiMSCI settlement price is the value of the SiMSCI computed based on the Special quotation methodology applied on each component stock of the SiMSCI on the day following the last trading day.

[28] The data for SVAR analysis of HSIF and KLCIF range from January 3, 1996 to December 31, 2008, while the data set for SiMSCI span from January 3, 1999 to December 31, 2008.

Table 4.1 South China Morning Post: Rotated Component Matrix

The data is output from General Inquirer, which consist of 77 word categories as defined in Harvard IV psychosocial dictionary. All word categories have been demeaned by prior year's day of the week mean, to alleviate the day-of-week effect on principal component analysis. The principal component analysis is performed based on rule of thumb. Variables are considered problematic will be removed from the principal component analysis if KMO measure of sampling adequacy less than 0.5; anti image correlation less than 0.5; communality less than 0.5; and having complex structure where a variable has loading of more than 0.4 in more than one component. The same process repeats until all the criteria are met. **Significant at 1% level

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
First Principal Component													
Positive	-0.233												
Strong	0.003	-0.308	-0.084	-0.192	-0.173		-0.141			-0.075	-0.141	-0.088	-0.107
Active		0.087	-0.142	0.038	-0.034		-0.024			0.086	-0.136	-0.034	-0.100
Rise		-0.134											
Negative	**0.850**	**0.880**	**0.890**	**0.840**	**0.870**	**0.846**	**0.910**	**0.835**	**0.888**	**0.880**	**0.900**	**0.880**	**0.850**
Weak	**0.810**	**0.860**	**0.850**	**0.890**	**0.860**	**0.835**	**0.880**	**0.844**	**0.900**	**0.920**	**0.860**	**0.870**	**0.860**
Fail	**0.740**												
Fall		**0.750**	**0.750**	**0.750**	**0.680**	**0.725**		**0.790**	**0.812**	**0.790**	**0.760**	**0.790**	**0.720**
Second Principal Component													
Positive	**0.790**												
Strong	**0.840**	**0.730**	**0.840**	**0.770**	**0.780**		**0.830**			**0.920**	**0.820**	**0.870**	**0.820**
Active		**0.770**	**0.810**	**0.820**	**0.860**		**0.860**			**0.930**	**0.810**	**0.880**	**0.820**
Rise		**0.710**											
Negative	-0.221	-0.167	0.069	0.072	0.096		0.010			0.156	0.081	0.123	0.084
Weak	-0.355	-0.204	-0.249	-0.078	-0.182		-0.196			0.017	-0.235	-0.147	-0.173
Fail	0.111												
Fall		0.024	-0.201	-0.259	-0.284					-0.143	-0.129	-0.154	-0.338
Sample size	258	256	256	257	253	250	254	256	251	245	242	249	251
KMO measures of sampling adequacy	0.669	0.687	0.667	0.639	0.644	0.651	0.537	0.688	0.695	0.61	0.673	0.641	0.681
Bartlett's Test Chi-square	342.48**	433.57**	334.62**	260.89**	274.31**	162.83**	200.40**	193.06**	314.01**	519.11**	368.68**	360.38**	272.85**
Total variance explain in first component	39.57%	36.63%	42.07%	40.71%	39.78%	64.62%	40.79%	67.83%	75.27%	45.09%	43.41%	43.24%	39.75%
Total variance explain in second component	30.22%	28.36%	29.36%	27.00%	29.13%		36.46%			35.09%	28.19%	31.50%	29.89%
cumulative rotation sum of squares	69.79%	64.99%	71.43%	67.71%	68.91%	64.62%	77.25%	67.83%	75.27%	80.18%	71.59%	74.74%	69.64%

Table 4.2 New Straits Times: Rotated Component Matrix

The data is output from General Inquirer, which consists of 77 Harvard IV psychosocial word categories. All word categories have been demeaned by prior year's day-of-the week average, to alleviate the day-of-week effect before the analysis begins. Variables are considered problematic and will be removed from the principal component analysis if: KMO measure of sampling adequacy less than 0.5; anti image correlation less than 0.5; communality less than 0.5; and having complex structure where a variable has loading of more than 0.4 in more than one component. The same process repeats until all the criteria are met. Rotated component matrix shows the factor loadings, which are coefficients of correlation between word categories and principal factors. Principal factor is a linear combination of word categories that are highly correlated, and explain the most of the variation among the 77 word categories. **Significant at 1% level.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
First Principal Component													
Positive	-0.016			-0.048			-0.062		-0.274			-0.105	
Strong	-0.152			-0.219					0.053			-0.119	
Negative	**0.843**	**0.943**	**0.921**	**0.830**	**0.878**	**0.928**	**0.791**	**0.919**	**0.858**	**0.919**	**0.926**	**0.935**	**0.942**
Weak	**0.865**	**0.943**	**0.921**	**0.833**	**0.850**	**0.928**	**0.851**	**0.919**	**0.845**	**0.918**	**0.926**	**0.910**	**0.942**
Fail	**0.734**			**0.837**	**0.754**		**0.765**		**0.730**				
Second Principal Component													
Positive	**0.783**			**0.845**			**0.931**		**0.721**				
Strong	**0.764**			**0.711**					**0.858**				
Negative	-0.196			-0.294			-0.411		-0.151				
Weak	-0.254			-0.357			-0.188		-0.251				
Fail	0.108			0.077			0.287		0.051				
Sample size	242	242	219	232	238	241	248	246	248	247	245	248	242
KMO measures of sampling adequacy	0.637	0.5	0.5	0.719	0.661	0.5	0.605	0.5	0.654	0.5	0.5	0.544	0.5
Bartlett's Test Chi-square	280.40**	221.61**	142.76**	364.79**	206.34**	174.79**	234.57**	155.73**	294.53**	154.21**	174.61**	246.63**	218.70**
Total variance explain in first component	40.42%	88.85%	84.74%	42.71%	68.79%	86.04%	48.49%	84.37%	41.25%	84.20%	85.82%	43.21%	88.69%
Total variance explain in second component	25.00%			28.78%			28.84%		26.89%			32.91%	
cumulative rotation sum of squares	65.42%	88.85%	84.74%	71.48%	68.79%	86.04%	77.33%	84.37%	68.13%	84.20%	85.82%	76.12%	88.69%

Table 4.3 The Straits Times: Rotated Component Matrix

The data is output from General Inquirer, which consists of 77 word categories as defined in Harvard IV psychosocial dictionary. All word categories have been demeaned by prior year's day of the week mean, to alleviate the day-of-week effect on principal component analysis. The principal component analysis is performed based on rule of thumb. Variables are considered problematic and will be removed from the principal component analysis if KMO measure of sampling adequacy less than 0.5; anti image correlation less than 0.5; communality less than 0.5; and having complex structure where a variable has loading of more than 0.4 in more than one component. The same process repeats until all the criteria are met. **Significant at 1% level

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
First Principal Component										
Strong	-0.103	-0.297			-0.322		-0.304	-0.182		
Active	-0.006	-0.022			-0.077		0.080	-0.069		
Negative	**0.831**	**0.878**	**0.417**	**0.887**	**0.865**	**0.860**	**0.910**	**0.915**	**0.852**	**0.858**
Weak	**0.872**	**0.877**	**0.423**	**0.902**	**0.861**	**0.881**	**0.836**	**0.866**	**0.900**	**0.876**
Fall	**0.722**	**0.743**	**0.360**	**0.743**	**0.740**	**0.799**			**0.767**	**0.787**
Second Principal Component										
Strong	**0.798**	**0.753**			**0.785**		**0.734**	**0.801**		
Active	**0.853**	**0.884**			**0.897**		**0.873**	**0.845**		
Negative	0.187	-0.059			-0.046		0.066	-0.045		
Weak	-0.116	-0.131			-0.265		-0.288	-0.241		
Fall	-0.356	-0.268			-0.263					
Sample size	241	244	240	240	247	238	239	235	222	214
KMO measures of sampling adequacy	0.631	0.731	0.659	0.649	0.735	0.69	0.545	0.59	0.651	0.684
Bartlett's Test Chi-square	263.31**	343.51**	216.81**	272.65**	396.66**	236.81**	168.46**	185.53**	224.20**	200.38**
Total variance explain in first component	39.68%	43.62%	69.15%	71.73%	42.94%	71.83%	40.65%	40.64%	70.81%	70.72%
Total variance explain in second component	30.79%	28.82%			34.24%		34.68%	35.39%		
cumulative rotation sum of squares	70.47%	72.44%	69.15%	71.73%	77.18%	71.83%	75.32%	76.03%	70.81%	70.72%

The principal component analysis generates two principal components that have eigenvalues greater than 1. Based on the analysis results, 9 out of 77 the word categories are linearly combined to obtain two factor scores. The first one, *Positive* (related to positive outlook); second, *Strong* (indicating strength); third, *Active* (regarding active orientation); fourth, *Rise* (indicating movements); fifth, *Negative* (associated with negative outlook); sixth, *Weak* (implying weakness); seventh, *Passive* (regarding passive orientation); eighth, *Fail* (indicating goal has not been achieved) and ninth, *Fall* (indicating movements). Positive, *Strong, Active*, and *Rise* can be categorised as words implying optimism. *Negative, Weak, Passive, Fail,* and *Fall* are associated with pessimism. Henceforth, I use the term "optimism news factor" and "pessimism news factor" in the discussions.

Table 4.1 illustrates the rotated component matrix for South China Morning Post, from 1996 to 2008. The extracted first principal components are consistent across the years, which are linearly combining three Harvard IV psychosocial word categories: *Negative, Weak* and *Fall.* The only exception is for year 1996 that consists *of Negative, Weak* and *Fail*. I generate *Pessimism* news factors for this sample. Generally, the second principal component for the same period made of *Strong* and *Active*, generating *Optimism* news factor.

Table 4.2 summarises the principal component analysis results for the New Straits Times. *Negative* and *Weak* words dominate the news sentiment for the period of 1996 to 2008. In addition, *Fail* word is included in the first principal component for 1996, 1999, 2002 and 2004. The PCA also generates second principal component for 1996, 1999, 2002 and 2004, which consist of *Positive* and *Strong* word categories.

Table 4.3 reports the rotated component matrix for The Straits Times ranging from 1999 to 2008. For all years, the first principal component is an approximate linear combination of Negative, *Weak,* and *Fall* with positive weights; except for 2005 and 2006, which only consist of *Negative* and *Weak*. The analysis only extracts one principal component for year 2001, 2002, 2007 and 2008. Similar to as South China Morning Post, *Pessimism* news factor is the first principal factor and *Optimism* news factor is the second principal factor.

Generally, the cumulative rotation sum of squares for these factor analyses exceed 60%, meaning the first and second principal components explain at least 60% of the variations among the 77 Harvard IV psychosocial word categories. *Pessimism* is more prominent as it carries more weight in the first principal component.

4.3 CONSISTENCY OF THE PESSIMISM NEWS FACTOR OVER THE TIME

Table 4.4, Table 4.5 and Table 4.6 presents the reliability of the principal component analysis in generating the news factor over the sample period. First, I construct principal component analysis for each yearly sub-sample. Second, the factor score coefficients from each yearly sub-sample are fitted into the relevant GI categories to generate factor scores. Finally, I match each yearly factor score generated by factor score coefficients estimated from each yearly sub-sample to obtain the coefficients of correlation. The column labelled as 1996 shows the correlation between 1) News factor scores obtained by fitting news factor coefficients generated from the 1996 sub-sample into the 1996 data; and 2) News factor scores obtained by fitting news factor coefficients generated yearly from 1996 to 2008 into the 1996 data. The row labelled as 1996 shows the correlation between 1) News factor scores obtained by fitting news factor coefficients generated from each year's data into the data from the same year itself; and 2) News factor scores obtained by fitting news factor coefficients generated from 1996 data into each yearly data from 1996 to 2008.

If the optimism news factor and pessimism news factor are use simultaneously in the equations, a negative relationship between optimism and pessimism confounds the estimation. In this case, I select only one news factor for further analysis. Factor loadings of pessimistic word categories are generally higher and positive for the South China Morning Post and The Straits Times in the first principal component. Higher positive factor loading implies a higher positive relationship between pessimistic word categories and the news factor. In conclusion, the pessimism news factor is a more representative news factor, therefore it is used for further analysis.

Table 4.4 New Straits Times: Correlations of the Pessimism News Factors Constructed Yearly

The data generated form principal component analysis (PCA). Pessimism news factor score coefficients are generated from each yearly sub-sample, and fit into all yearly sub-samples. Then correlations among pessimism news factors derive from factor loadings of different years are examined. The pessimism news factors are robust to time variation if these news sentiments factor scores are highly correlated. The factor analysis method can reduce a large number of word categories into a small number of factors. PCA derives a news factor by linearly combine the above demeaned word categories that explain most of the variation among all 77 word categories. Each word category is equally light in calculation of variance. Word categories that are highly correlated will be combine become one factor (or component). Each principal component is orthogonal on each other. Only the principal component with eigenvalue more than one will be selected for further analysis. *Significant at 5% level ;**Significant at 1% level.

Yearly factor score	Sample												
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
1996	1	.963**	.933**	.983**	.982**	.929**	.987**	.957**	.982**	.936**	.955**	.956**	.934**
1997	.916**	1	.999**	.883**	.945**	1.000**	.897**	1.000**	.916**	1.000**	1.000**	.987**	1.000**
1998	.915**	.999**	1	.882**	.949**	.998**	.893**	.997**	.919**	1.000**	.999**	.990**	.999**
1999	.975**	.894**	.835**	1	.964**	.827**	.989**	.886**	.961**	.842**	.870**	.864**	.837**
2000	.987**	.964**	.928**	.968**	1	.927**	.988**	.964**	.984**	.921**	.957**	.948**	.925**
2001	.916**	1.000**	.998**	.883**	.945**	1	.897**	1.000**	.916**	.999**	1.000**	.987**	1.000**
2002	.985**	.936**	.888**	.989**	.992**	.882**	1	.936**	.974**	.884**	.921**	.916**	.885**
2003	.916**	1.000**	.997**	.882**	.943**	1.000**	.898**	1	.915**	.998**	1.000**	.985**	.999**
2004	.981**	.949**	.915**	.963**	.975**	.920**	.956**	.945**	1	.914**	.953**	.959**	.927**
2005	.915**	1.000**	1.000**	.883**	.948**	.999**	.895**	.998**	.918**	1	1.000**	.989**	1.000**
2006	.916**	1.000**	.999**	.883**	.946**	1.000**	.896**	1.000**	.917**	1.000**	1	.988**	1.000**
2007	.926**	.991**	.991**	.902**	.939**	.990**	.896**	.988**	.929**	.990**	.992**	1	.992**
2008	.916**	1.000**	.999**	.883**	.947**	1.000**	.896**	.999**	.917**	1.000**	1.000**	.988**	1
Sample Size	242	242	219	232	238	241	248	246	248	247	245	248	242

Table 4.5 South China Morning Post: Correlations of the Pessimism News Factors Constructed Yearly

The data generated form principal component analysis (PCA). Pessimism news factor score coefficients are generated from each yearly sub samples and fit into all yearly sub samples. Then correlations among pessimism news factors derive from factor loadings of different years are examined. The pessimism news factors are robust to time variation if these news sentiments factor scores are highly correlated. The factor analysis method can reduce a large number of word categories into a small number of factors. PCA derives a news factor by linearly combine the above demeaned word categories that explain most of the variation among all 77 word categories. Each word category is equally light in calculation of variance. Word categories that are highly correlated will be combine become one factor (or component). Each principal component is orthogonal on each other. Only the principal component with eigenvalue more than one will be selected for further analysis. *Significant at 5% level; **Significant at 1% level.

Yearly factor scores	Sample												
	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
1996	1	.865**	.877**	.852**	.889**	.842**	.896**	.804**	.880**	.890**	.930**	.892**	.891**
1997	.818**	1	.985**	.988**	.987**	.976**	.909**	.976**	.984**	.961**	.986**	.986**	.986**
1998	.860**	.990**	1	.992**	.995**	.987**	.935**	.986**	.993**	.974**	.999**	.993**	.998**
1999	.856**	.992**	.994**	1	.998**	.989**	.944**	.991**	.995**	.986**	.992**	.996**	.995**
2000	.871**	.989**	.996**	.997**	1	.981**	.960**	.984**	.992**	.976**	.995**	.990**	.998**
2001	.821**	.984**	.987**	.989**	.979**	1	.895**	1.000**	.999**	.998**	.984**	.998**	.982**
2002	.902**	.936**	.955**	.951**	.960**	.928**	1	.933**	.951**	.931**	.950**	.949**	.956**
2003	.822**	.984**	.986**	.989**	.978**	1.000**	.895**	1	.998**	.998**	.983**	.998**	.981**
2004	.836**	.983**	.991**	.992**	.988**	.998**	.919**	.996**	1	.999**	.986**	.997**	.987**
2005	.835**	.983**	.986**	.993**	.987**	.998**	.921**	.998**	.999**	1	.980**	.996**	.985**
2006	.855**	.993**	.999**	.993**	.995**	.983**	.932**	.983**	.990**	.964**	1	.990**	.998**
2007	.834**	.990**	.993**	.995**	.987**	.998**	.909**	.998**	.998**	.994**	.992**	1	.989**
2008	.873**	.988**	.998**	.994**	.998**	.981**	.952**	.983**	.989**	.966**	.998**	.989**	1
Sample Size	258	256	256	257	253	250	254	256	251	245	242	249	251

Table 4.6 The Straits Times: Correlations of the Pessimism News Factors Constructed Yearly

The data generated form principal component analysis (PCA). Factor scores are generated from each yearly sub samples and fit into all yearly sub samples. Then correlations among news factors derive from factor loadings of different years are examined. The factor loadings are robust to time variation if these news sentiments are highly correlated. The factor analysis method can reduce a large number of word categories into a small number of factors. PCA derives a news factor by linearly combine the above demeaned word categories that explain most of the variation among all 77 word categories. Each word category is equally light in calculation of variance. Word categories that are highly correlated will be combine become one factor (or component). Each principal component is orthogonal on each other. Only the principal component with eigenvalue more than one will be selected for further analysis. *Significant at 5% level; **Significant at 1% level.

Yearly Factor Scores	Sample									
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
1999	1	.994**	.990**	.994**	.991**	.992**	.963**	.929**	.986**	.992**
2000	.994**	1	.979**	.980**	1.000**	.979**	.955**	.907**	.977**	.980**
2001	.982**	.980**	1	.998**	.971**	.998**	.911**	.844**	.999**	.999**
2002	.992**	.983**	.998**	1	.975**	.999**	.935**	.881**	.996**	1.000**
2003	.994**	1.000**	.978**	.978**	1	.977**	.952**	.911**	.976**	.979**
2004	.990**	.983**	.998**	.999**	.975**	1	.932**	.872**	.997**	.999**
2005	.945**	.963**	.939**	.938**	.956**	.930**	1	.992**	.936**	.927**
2006	.945**	.962**	.936**	.936**	.955**	.933**	.993**	1	.932**	.925**
2007	.975**	.977**	.999**	.995**	.969**	.997**	.898**	.821**	1	.997**
2008	.989**	.982**	.999**	1.000**	.974**	.999**	.927**	.869**	.998**	1
Sample Size	241	244	240	240	247	238	239	235	222	214

The news factor generated for the New Straits Times is consistent over the years, as Table 4.4 shows. The first principal components consist of negative words with positive loadings. Therefore, I name it as the *Pessimism* news factor. The factor scores are highly correlated with a positive sign. This further supports the conclusion made earlier.

Table 4.5 and Table 4.6 depict the correlations analysis for the *Pessimism* news factor generated for the South China Morning Post and The Straits Times respectively. Generally, the coefficients of correlation are positive and almost close to one. This implies that the *Pessimism* new factors generated using factor loadings of different years are highly correlated. This further confirms the robustness of the *Pessimism* factor score over time.

4.4 DESCRIPTIVE STATISTICS

Table 4.7 presents the average trading volume, open interest and settlement price by year. Generally, the average daily volume (in number of contracts), open interest (in number of contracts) and settlement price (in domestic currency) depict an increasing trend. Despite the long-run trend, the average daily settlement price for 2002 and 2003 are lower than previous years. However, the positive time trend starts again for the period 2004 to 2007. All three variables show negative growth in 2008 except the HSIF and MSCI volumes. This suggests the need for trend adjustment.

Table 4.7 Average Daily Trading Volume, Open Interest and Settlement Price, 1995 - 2008

The daily statistics are summarised from data that has been obtained from Datastream. The trading volume(*V*) and open interest (OI) are in number of contracts. The settlement price (SP) is in Hong Kong Dollar, Singapore Dollar and Malaysian Ringgit respectively for Hang Seng Index Futures (HSIF), Singapore Morgan Stanley Free Index Futures (SiMSCIF) and Kuala Lumpur Composite Index Futures (KLCIF).

	HSIF			SiMSCI			KLCIF		
	V	OI	SP	*V*	OI	SP	*V*	OI	SP
1996	18398	51911	11507.76	NA	NA	NA	229	986	1130.00
1997	18583	57782	13389.67	NA	NA	NA	1239	4187	967.00
1998	26307	70466	9375.63	322	870	117.34	2455	11155	516.00
1999	28300	49396	12758.96	1110	2984	270.57	1398	2757	706.56
2000	16200	35151	16112.38	1869	4447	273.07	1173	2534	844.08
2001	18062	37053	12563.66	1874	6769	204.94	1183	2489	975.59
2002	19595	48957	10431.79	2478	13873	197.22	701	2276	713.40
2003	27344	76278	10249.62	3555	23702	183.45	975	3624	705.78
2004	34619	107377	12872.74	5108	32547	230.02	3735	11860	854.97
2005	40114	114366	14334.29	6169	38152	263.42	4458	16795	896.05
2006	51455	128187	16885.98	8475	46383	303.97	6632	28187	945.93
2007	69716	134697	23190.15	15442	61995	420.24	12736	35314	1298.05
2008	88618	110170	20830.96	17511	61412	328.61	11601	33834	1134.58

Table 4.8 summarises the close-to-close trading activity by contract, and breaks down the sample into two sub-periods. All the variables are significantly autocorrelated up to five lags. The first order autocorrelation for returns are negative for all three index futures contracts and all sample periods. These results warrant the need to incorporate up to five lags of returns, news sentiment and trading volume in the structural vector autoregressive model.

Table 4.8 Summary of Trading Activity by Contract

This table summarises the daily returns (24-hour returns), detrended log volume and detrended open interest of Hang Seng Index Futures (HSIF), Kuala Lumpur Composite Index Futures (KLCIF) and Morgan Stanley Singapore free Index Futures (SiMSCIF). The full sample period of HSIF and KLCIF covers 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for SiMSCI from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 span from 1/2004 through 12/2008. Panel A reports daily returns in percentage which are calculated as log difference of the price. Panel B reports daily detrended log volume, in which the number of contracts traded is transform into its natural logarithm and less 60 days moving average of the log contract volume. Panel C reports daily detrended log open interest. The detrending method is the same as in panel B. $^{+}$ all the autocorrelations are significant at α=0.01

	Mean	Max	Min	Std. Dev	Skewness	Kurtosis	Autocorrelation				
							1^{+}	2^{+}	3^{+}	4^{+}	5^{+}
Panel A: Daily returns (in percentage)											
Hang Seng Index Futures (HSIF)											
Full period	0.002	0.230	-0.161	0.020	0.281	13.890	-0.472	-0.016	-0.010	-0.099	0.219
Sub-period 1	0.006	0.230	-0.161	0.023	0.407	13.120	-0.077	-0.040	0.096	-0.069	-0.015
Sub-period 2	-0.002	0.113	-0.116	0.017	-0.004	10.847	-0.050	-0.029	-0.007	-0.005	-0.011
Kuala Lumpur Composite Index Futures (KLCIF)											
Full period	-0.009	0.290	-0.387	0.021	-0.971	63.124	-0.074	-0.013	0.016	-0.060	0.016
Sub-period 1	-0.033	0.290	-0.387	0.027	-0.843	44.689	-0.078	-0.019	0.012	-0.083	0.023
Sub-period 2	0.013	0.060	-0.076	0.012	-0.407	6.514	-0.057	0.014	0.032	0.052	-0.018
Morgan Stanley Singapore Free Index Futures (SiMSCIF)											
Full period	0.001	0.073	-0.077	0.015	-0.205	5.796	-0.053	0.017	-0.018	0.048	0.017
Sub-period 1	0.001	0.062	-0.071	0.016	-0.017	4.557	-0.025	0.027	0.012	0.080	-0.009
Sub-period 2	0.000	0.073	-0.077	0.015	-0.469	7.592	-0.092	0.004	-0.059	0.007	0.048
Panel B: Detrended log volume											
Hang Seng Index Futures (HSIF)											
Full period	0.002	8.362	-2.481	0.369	3.350	83.483	-0.097	-0.057	-0.087	-0.055	-0.042
Sub-period 1	0.002	2.264	-2.481	0.353	-0.214	7.207	-0.136	-0.062	-0.084	-0.071	-0.036
Sub-period 2	-0.001	0.002	1.287	-1.429	0.343	-0.164	0.008	-0.132	-0.156	-0.047	-0.072
Kuala Lumpur Composite Index Futures (KLCIF)											
Full period	0.038	1.850	-2.255	0.570	-0.061	3.148	0.663	0.458	0.314	0.211	0.156
Sub-period 1	0.031	1.590	-2.255	0.542	-0.539	3.766	0.656	0.487	0.402	0.323	0.272
Sub-period 2	0.045	1.850	-1.751	0.596	0.269	2.674	0.668	0.437	0.246	0.124	0.065
Morgan Stanley Singapore Free Index Futures (SiMSCIF)											
Full period	0.042	2.574	-4.694	0.675	0.614	6.565	0.609	0.305	0.093	-0.010	-0.069
Sub-period 1	0.053	2.179	-1.865	0.569	0.759	4.269	0.612	0.322	0.128	0.016	-0.060
Sub-period 2	0.031	2.574	-4.694	0.771	0.545	6.601	0.606	0.295	0.073	-0.026	-0.076
Panel C: Detrended log open interest											
Hang Seng Index Futures (HSIF)											
Full period	0.009	0.586	-0.665	0.128	0.116	4.192	0.861	0.732	0.651	0.600	0.561
Sub-period 1	0.018	0.415	-0.281	0.106	0.180	3.340	0.855	0.698	0.594	0.530	0.487
Sub-period 2	0.001	-0.008	0.586	-0.268	0.118	0.742	0.749	0.530	0.410	0.357	0.314
Kuala Lumpur Composite Index Futures (KLCIF)											
Full period	0.037	0.836	-2.961	0.292	-2.700	24.144	0.927	0.867	0.812	0.767	0.733
Sub-period 1	0.031	0.815	-2.961	0.357	-3.139	21.579	0.944	0.907	0.869	0.832	0.802
Sub-period 2	0.042	0.836	-0.468	0.213	0.556	3.445	0.884	0.761	0.660	0.596	0.552
Morgan Stanley Singapore Free Index Futures (SiMSCIF)											
Full period	0.048	0.621	-0.343	0.127	0.846	4.504	0.824	0.667	0.580	0.538	0.513
Sub-period 1	0.075	0.621	-0.252	0.137	0.872	3.739	0.839	0.709	0.633	0.584	0.551
Sub-period 2	0.020	0.521	-0.344	0.108	0.456	4.914	0.776	0.553	0.436	0.399	0.385

Table 4.9 Summary of News Factor by Source of News

This table summarises the daily Pessimism, Negative and Weak news factor generated by General Inquirer and Principal Component Analysis for Asia Wall Street Journal (Abreast of the Market), South China Morning Post, The New Straits and The Straits Times. The full sample period of Abreast of the market, the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times span from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. Panel A reports daily Pessimism news factor, Panel B reports Negative news factor and Panel C reports Weak news factor. The day-of-the-week variations are taken care of by demeaning the news factor with prior year's day-of-the-week's mean.

	Mean	Max	Min	Std. Dev	Skewness	Kurtosis	Autocorrelation				
							1	2	3	4	5
Panel A: Pessimism news factor (Psm)											
South China Morning Post											
Full period	0.004	5.035	-3.084	0.952	0.717	3.997	0.114	0.083	0.043	0.020	0.033
Sub-period 1	0.001	3.504	-2.622	0.983	0.622	3.432	0.161	0.121	0.114	0.050	0.081
Sub-period 2	0.009	4.258	-2.142	0.988	0.742	3.677	0.072	0.049	0.008	0.012	0.002
The New Straits Times											
Full period	-0.001	4.218	-3.091	0.999	0.552	3.352	0.182	0.122	0.116	0.066	0.081
Sub-period 1	-0.007	4.218	-3.091	0.992	0.502	3.145	0.215	0.123	0.116	0.061	0.076
Sub-period 2	0.005	4.077	-2.615	1.006	0.599	3.540	0.151	0.121	0.115	0.070	0.085
The Straits Times											
Full period	0.001	4.138	-2.921	0.998	0.504	3.257	0.162	0.113	0.112	0.105	0.078
Sub-period 1	0.002	3.913	-2.717	0.998	0.534	3.336	0.145	0.103	0.118	0.121	0.099
Sub-period 2	0.000	4.138	-2.921	0.998	0.473	3.173	0.179	0.122	0.104	0.088	0.058
Panel **B**: Negative news factor (NGV)											
South China Morning Post											
Full period	0.019	5.472	-2.796	1.075	0.598	3.781	0.225	0.206	0.164	0.131	0.187
Sub-period 1	-0.093	3.984	-2.590	1.107	0.486	3.270	0.298	0.270	0.255	0.195	0.256
Sub-period 2	0.136	4.615	-2.796	1.132	0.601	3.491	0.175	0.158	0.134	0.091	0.152
The New Straits Times											
Full period	0.038	4.641	-3.051	1.068	0.526	3.619	0.256	0.210	0.195	0.156	0.184
Sub-period 1	-0.061	4.641	-3.051	1.054	0.486	3.419	0.299	0.222	0.208	0.185	0.195
Sub-period 2	0.136	4.418	-2.858	1.073	0.565	3.782	0.203	0.184	0.169	0.115	0.161
The Straits Times											
Full period	0.090	4.688	-3.114	1.106	0.589	3.304	0.215	0.170	0.171	0.191	0.171
Sub-period 1	0.094	4.688	-3.114	1.131	0.550	3.298	0.183	0.141	0.154	0.192	0.157
Sub-period 2	0.085	4.043	-2.506	1.080	0.633	3.294	0.251	0.203	0.188	0.189	0.187
Panel C: Weak news factor (*Weak)*											
South China Morning Post											
Full period	-0.009	4.691	-2.968	1.039	0.584	3.745	0.187	0.152	0.125	0.122	0.123
Sub-period 1	-0.102	4.691	-2.504	1.045	0.654	3.844	0.184	0.163	0.184	0.116	0.134
Sub-period 2	0.080	4.651	-2.968	1.103	0.468	3.206	0.202	0.168	0.113	0.163	0.123
The New Straits Times											
Full period	0.052	4.988	-2.557	1.046	0.593	3.361	0.230	0.178	0.173	0.136	0.143
Sub-period 1	-0.013	4.988	-2.471	1.022	0.591	3.262	0.265	0.182	0.172	0.119	0.136
Sub-period 2	0.116	4.313	-2.557	1.066	0.587	3.422	0.192	0.167	0.167	0.145	0.142
The Straits Times											
Full period	0.046	4.525	-2.815	1.050	0.551	3.331	0.197	0.172	0.154	0.121	0.103
Sub-period 1	0.057	4.318	-2.374	1.025	0.557	3.326	0.153	0.138	0.160	0.110	0.097
Sub-period 2	0.035	4.525	-2.815	1.076	0.547	3.323	0.239	0.202	0.146	0.130	0.108

Table 4.9 summarises the news factor by source of news.

4.5 STRUCTURAL VECTOR AUTOREGRESSIVE (SVAR) ESTIMATES AND CROSS-COUNTRY COMPARISON

This section discusses the relationship between news factors, index futures returns and trading volume. Only *Pessimism*, *Negative* and *Weak* news factors are included in the analysis. According to the principal component analysis, the latter two explain most of the variation of the news sentiment to form the *Pessimism* news factor. Henceforth, for the brevity of discussion, "bad news factors" is used when the discussion about the *Negative, Weak* and *Pessimism* news factors is made as a whole.

4.5.1 Stationary test

Table 4.10 Augmented Dickey Fuller Tests for Trading Data and Bad News Factors
This table reports the t-statistics for Augmented Dickey Fuller test for the time series that enter the structural vector autoregressive model. The null hypothesis tests for a unit root in the series. (i.e. non-stationary). The full sample period of the South China Morning Post and The New Straits Times cover 1/1996 to 12/2008.The full sample period for the Straits Times spans 1/1999 through 12/2003. ** indicates the test statistics significant at 1% level.

	t	Probability	t	Probability	t	Probability
Contracts	HSIF		KLCIF		SiMSCIF	
Returns	-59.892**	0.000	-59.976**	0.000	-51.182**	0.000
Open Interest	-7.293**	0.000	-10.751**	0.000	-5.539**	0.000
Trading volume	-11.713**	0.000	-24.817**	0.000	-7.291**	0.000
Source of news	South China Morning Post		New Straits Times		The Straits Times	
Pessimism	-34.738**	0.000	-25.839**	0.000	-19.168**	0.000
Negative	-10.684**	0.000	-11.234**	0.000	-11.991**	0.000
Weak	-12.361**	0.000	-17.870**	0.000	-21.123**	0.000
|*Pessimism*|	-54.708**	0.000	-56.208**	0.000	-48.153**	0.000
|*Negative*|	-54.041**	0.000	-55.893**	0.000	-20.291**	0.000
|*Weak*|	-53.393**	0.000	-56.260**	0.000	-45.954**	0.000

Table 4.10 presents the t-test statistics of Augmented Dickey Fuller tests, to examine the stationarity of the time series used in the SVAR. Including non-stationary variables in a regression model will result in spurious regression problem. The null hypothesis tests for a

unit root in the series to determine whether the series is non-stationary. All the endogenous variables included in the SVAR model are stationary with p-value less than 0.01. The results confirm that it is unnecessary to perform differentiation on the variables.

4.5.2 Predicting the bad news factor using index futures returns

Table 4.11 presents the SVAR estimates of the coefficient, γ_{1i}, for equation (3.1), to test the Hypothesis 1. I expect the sign of the coefficients to be negative. The estimates show the impact of a 1% increase in HSIF, KLCIF or SiMSCIF close-to-close returns on each of the *Pessimism*, *Negative* and *Weak* news factors in the unit of percentage of one standard deviation. The proxy for trading volume is open interest. The news articles are the daily summary of the stock market movements. It is natural for journalists to use more pessimistic words in the news articles if the market is closed at a lower price.

The impact of negative returns on *Pessimism, Negative and Weak* news factors last up to three days. The impact only lasts for one day for Dow Jones Industrial Index (Tetlock, 2007). For the full sample period analysis, a 1% decrease in HSIF returns today is associated with the increase of 17.63% of a standard deviation of the *Pessimism* news factor on the next day. The p-value for the joint null hypothesis that all five lags of returns do not predict the bad news factor are all less than 0.01. These findings provide strong evidence on the association between lag returns and the bad news factors. In addition, the results for the individual coefficient tests and joint coefficients test are consistent over the three index futures contracts.

This study includes two proxies for trading volume, daily open interest and number of contracts traded daily. The Table 4.12 depicts the estimates of the coefficient, γ_{1i}, for equation (3.1). The endogenous variables are the three bad news factors (*Pessimism, Negative*, and *Weak*), close-to-close returns and number of contracts traded. The result is similar to that of the estimates using open interest.

Table 4.11 Predicting Bad News Factor Using Regional Index Futures Returns and Open Interests

This table presents the coefficients, γ_{1i}, estimated for equation (1) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and trading volume in open interest.

$$N_t = \alpha_1 + \sum_{i=1}^{5} \beta_{1i} N_{t-i} + \sum_{i=0}^{5} \gamma_{1i} R_{t-i} + \sum_{i=0}^{5} \delta_{1i} V_{t-i} + \sum_{j=1}^{14} \lambda_{1j} EXOG_j + \varepsilon_{1t} \quad (3.1)$$

The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008.The full sample period for the Straits Times spans 1/1999 through 12/2003. The coefficients denote the impact of 1% decrease in HSIF, KLCIF and SiMSCIF returns on bad news factor in the unit of percentage of standard deviation. The $\gamma_{1i}$are expected to be negative. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, test for $\sum_{i=0}^{5} \gamma_{1i} = 0$; i=1,2,3,4,and 5; with degree of freedom equals to 1. The t values are reported in parentheses [ ] *, ** and *** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent variable	Bad News Factor			Bad News Factor			Bad News Factor		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Returns(-1)	-17.639***	-16.017***	-16.310***	-15.629***	-15.033***	-16.715***	-28.085***	-26.664***	-26.361***
	[-21.866]	[-17.959]	[-18.634]	[-18.648]	[-17.175]	[-19.467]	[-23.327]	[-20.033]	[-20.703]
Returns(-2)	-6.062***	-5.248***	-6.305***	-5.30**	-4.721***	-5.667***	-5.870***	-6.346***	-5.757***
	[-6.813]	[-5.464]	[-6.65]	[-5.595]	[-4.812]	[-5.825]	[-4.377]	[-4.400]	[-4.144]
Returns(-3)	-1.186	-0.479	-2.406**	-2.012**	-0.275	-2.240**	-2.246*	-2.423*	-2.580*
	[-1.324]	[-0.497]	[-2.526]	[-2.135]	[-0.282]	[-2.314]	[-1.665]	[-1.670]	[-1.846]
Returns(-4)	-0.639	-0.247	0.008	-0.875	-0.876	-0.765	-0.458	-0.708	-1.550
	[-0.712]	[-0.256]	[0.008]	[-0.927]	[-0.898]	[-0.789]	[-0.338]	[-0.487]	[-1.105]
Returns(-5)	-1.670*	-1.138	-0.678	-0.795	-0.528	-0.445	0.073	0.962	-0.497
	[-1.875]	[-1.186]	[-0.716]	[-0.841]	[-0.540]	[-0.459]	[0.054]	[0.670]	[-0.359]
$X^2(1)$Joint	506.830***	339.335***	374.848***	371.667***	310.510***	405.580***	554.876***	413.274***	439.156***
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.12 Predicting Bad News Factor Using Regional Index Futures Returns and Number of Contracts Traded

This table presents the coefficients γ_{1i} estimated for equation (1) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative* and *Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and trading volume in number of contracts traded.

$$N_t = \alpha_1 + \sum_{i=1}^{5} \beta_{1i} N_{t-i} + \sum_{i=0}^{5} \gamma_{1i} R_{t-i} + \sum_{i=0}^{5} \delta_{1i} V_{t-i} + \sum_{j=1}^{14} \lambda_{1j} EXOG_j + \varepsilon_{1t} \quad (3.1)$$

The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The full sample period for the Straits Times spans 1/1999 through 12/2003. The coefficients denote the impact of 1% increase in HSIF, KLCIF and SiMSCIF returns on bad news factor in standard deviation. The γ_{1i} are expected to be negative. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, test for $\sum_{i=0}^{5} \gamma_{1i} = 0$; i=1,2,3,4,and 5; with degree of freedom equals to 1.The t values are reported in parentheses [ ] *, ** and *** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent variable	Bad News Factor			Bad News Factor			Bad News Factor		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Returns(-1)	-17.630***	-15.929***	-16.239***	-15.574***	-15.000***	-16.582***	-27.832***	-26.494***	-26.068***
	[-21.858]	[-17.860]	[-18.548]	[-18.61]	[-17.155]	[-19.335]	[-23.149]	[-19.894]	[-20.495]
Returns(-2)	-6.098***	-5.238***	-6.271***	-5.243***	-4.658***	-5.526***	-5.778***	-6.198***	-5.673***
	[-6.855]	[-5.455]	[-6.612]	[-5.547]	[-4.756]	[-5.693]	[-4.323]	[-4.301]	[-4.095]
Returns(-3)	-1.266	-0.475	-2.354**	-2.095**	-0.309	-2.188**	-2.016	-2.251	-2.284
	[-1.415]	[-0.494]	[-2.474]	[-2.226]	[-0.317]	[-2.263]	[-1.498]	[-1.552]	[-1.638]
Returns(-4)	-0.539	-0.055	0.142	-0.934	-0.907	-0.680	-0.559	-0.733	-1.651
	[-0.601]	[-0.057]	[0.149]	[-0.992]	[-0.932]	[-0.703]	[-0.415]	[-0.505]	[-1.179]
Returns(-5)	-1.652*	-1.085	-0.590	-0.845	-0.571	-0.397	-0.033	0.752	-0.638
	[-1.854]	[-1.131]	[-0.623]	[-0.896]	[-0.585]	[-0.410]	[-0.025]	[0.524]	[-0.462]
$X^2(1)$Joint	507.433***	336.053***	372.288***	370.529***	309.530***	399.819***	545.830***	406.703***	430.015***
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.13 Predicting Regional Index Futures Returns Using Bad News Factor and Open Interests

This table presents the coefficients β_{2i} estimated for equation (2) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and trading volume in open interest.

$$R_t = \alpha_2 + \sum_{i=0}^{5} \beta_{2i} N_{t-i} + \sum_{i=1}^{5} \gamma_{2i} R_{t-i} + \sum_{i=0}^{5} \delta_{2i} V_{t-i} + \sum_{j=1}^{14} \lambda_{2j} EXOG_j + \varepsilon_{2t} \qquad (3.2)$$

The full sample period the South China Morning Post and The New Straits Times cover 1/1996 to 12/2008. The full sample period for the Straits Times spans 1/1999 through 12/2003. The coefficients denote the impact of one standard deviation change in bad news factor on the index futures return in basis points. I expect the $\beta_{2i}$are negative when i =0 and 1, the $\beta_{2i}$are positive when i=2, 3, 4 and 5. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, with 1 degree of freedom. The t values are reported in parentheses [ ]. *, ** and *** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent Variable	HSIF			KLCIF			SiMSCIF		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News measure	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
NEWS	-10.680***	-13.160***	-11.300***	-16.680***	-15.270***	-15.500***	-8.970***	-4.840	-7.960**
	[-2.691]	[-3.665]	[-3.089]	[-4.213]	[-4.029]	[-4.009]	[-2.520]	[-1.504]	[-2.365]
NEWS(-1)	-1.600	-3.410	-6.570*	-8.070**	-12.110***	-6.420	2.930	-1.880	-1.350
	[-0.401]	[-0.949]	[-1.787]	[-2.009]	[-3.159]	[-1.639]	[0.817]	[-0.583]	[-0.398]
NEWS(-2)	8.480**	2.620	5.810	-5.610	-5.020	-8.960**	-2.080	0.212	-3.800
	[2.119]	[0.725]	[1.580]	[-1.389]	[-1.298]	[-2.276]	[-0.581]	[0.066]	[-1.122]
NEWS(-3)	8.750**	4.660	8.290**	-4.640	-4.790	-5.320	1.690	-0.111	-0.278
	[2.186]	[1.287]	[2.257]	[-1.151]	[-1.240]	[-1.356]	[0.474]	[-0.035]	[-0.082]
NEWS(-4)	7.560*	3.050	0.066	1.510	-0.463	4.560	6.250*	4.240	6.560*
	[1.904]	[0.851]	[0.018]	[0.376]	[-0.120]	[1.165]	[1.755]	[1.329]	[1.949]
NEWS(-5)	1.860	-4.060	0.620	4.610	6.570*	1.790	-2.560	-4.930	-1.640
	[0.498]	[-1.183]	[0.177]	[1.220]	[1.808]	[0.488]	[-0.785]	[-1.632]	[-0.524]
sum (-2) to (-5)	26.650	6.270	14.786	-4.130	-3.703	-7.930	3.300	-0.589	0.842
$X^2(1)$	14.285***	1.243	6.032**	0.366	0.362	1.548	0.306	0.013	0.024
p-value	0.000	0.265	0.014	0.545	0.547	0.213	0.580	0.908	0.877
Sum (0) to(-5)	14.370	-10.300	-3.084	-28.880	-31.083	-29.850	-2.740	-7.309	-8.468
$X^2(1)$	2.727*	2.268	0.178	11.757***	17.148***	14.590***	0.142	1.410	1.618
p-value	0.099	0.132	0.673	0.001	0.000	0.000	0.707	0.235	0.203

Table 4.14 Predicting Regional Index Futures Returns Using Bad News Factor and Number of Contract Traded

This table presents the coefficients β_{2i} estimated for equation (2) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and number of contract traded

$$R_t = \alpha_2 + \sum_{i=0}^{5} \beta_{2i} N_{t-i} + \sum_{i=1}^{5} \gamma_{2i} R_{t-i} + \sum_{i=0}^{5} \delta_{2i} V_{t-i} + \sum_{j=1}^{14} \lambda_{2j} EXOG_j + \varepsilon_{2t} \qquad (3.2)$$

The full sample period South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The full sample period for the Straits Times spans 1/1999 through 12/2003. The coefficients denote the impact of one standard deviation change in bad news factor on the index futures return in basis points. I expect the β_{2i} are negative when i=0 and 1, the β_{2i} are positive when i=2, 3, 4 and 5. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, with 1 degree of freedom. The t values are reported in parentheses [ ]. *, ** and *** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent Variable	HSIF			KLCIF			SiMSCIF		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News measure	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
NEWS	-10.920***	-13.300***	-11.070***	-16.620***	-15.210***	-15.180***	-8.640**	-4.420	-7.580**
	[-2.749]	[-3.701]	[-3.025]	[-4.188]	[-4.006]	[-3.917]	[-2.421]	[-1.372]	[-2.247]
NEWS(-1)	-1.520	-3.210	-6.370*	-7.730*	-11.890***	-5.890	3.060	-1.940	-0.984
	[-0.381]	[-0.894]	[-1.733]	[-1.920]	[-3.098]	[-1.503]	[0.853]	[-0.605]	[-0.291]
NEWS(-2)	8.950**	3.440	6.240*	-5.190	-4.670	-8.250**	-2.200	0.118	-3.810
	[2.237]	[0.955]	[1.697]	[-1.282]	[-1.206]	[-2.095]	[-0.615]	[0.037]	[-1.124]
NEWS(-3)	8.540**	4.550	8.620**	-4.520	-4.750	-4.850	1.400	-0.467	-0.463
	[2.132]	[1.258]	[2.347]	[-1.119]	[-1.227]	[-1.233]	[0.391]	[-0.146]	[-0.137]
NEWS(-4)	7.160*	2.850	0.069	1.430	-0.517	4.770	6.350*	4.540	6.910**
	[1.804]	[0.795]	[0.019]	[0.355]	[-0.134]	[1.219]	[1.786]	[1.425]	[2.054]
NEWS(-5)	1.740	-3.850	0.410	5.060	6.880*	2.330	-2.720	-5.080*	-1.840
	[0.467]	[-1.122]	[0.117]	[1.339]	[1.889]	[0.633]	[-0.834]	[-1.684]	[-0.587]
Sum (-2) to -(5)	26.390	6.990	15.339	-3.220	-3.057	-6.000	2.830	-0.889	0.797
$X^2(1)$	14.036***	1.548	6.499**	0.220	0.247	0.887	0.227	0.030	0.022
p-value	0.000	0.213	0.011	0.639	0.619	0.346	0.634	0.862	0.882
Sum(0) to (-5)	13.950	-9.520	-2.101	-27.570	-30.157	-27.070	-2.750	-7.249	-7.767
$X^2(1)$	2.570	1.937	0.083	10.661***	16.105***	12.012***	0.141	1.389	1.358
p-value	0.109	0.164	0.774	0.001	0.000	0.001	0.707	0.239	0.244

Table 4.15 Predicting Open Interest Using Bad News Factor

This table presents the coefficients β_{4i} and θ_{4i} estimated for equation (13) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and open interest.

$$V_t = \alpha_4 + \sum_{i=0}^{5} \beta_{4i} N_{t-i} + \sum_{i=0}^{5} \theta_{4i} |N_{t-i}| + \sum_{i=0}^{5} \gamma_{4i} R_{t-i} + \sum_{i=1}^{5} \delta_{4i} V_{t-i} + \sum_{j=1}^{14} \lambda_{4j} EXOG_j + \varepsilon_{3t} \qquad (3.7)$$

The full sample period of The New Straits Times cover 1/1996 to 12/2008. The full sample period for the Straits Times spans 1/1999 through 12/2003. The coefficients denote the impact of one-standard deviation change in bad news factor and change in absolute bad news factors respectively on the percentage change of open interest of HSIF, KLCIF and SiMSCIF. I expect the β_{4i} are negative and the θ_{4i} are positive. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, with 1 degree of freedom. The t values are reported in parentheses [ ]. *, ** and *** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent variable	HSIF detrended log open interest			KLCIF detrended log open interest			SiMSCIF detrended log open interest		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
NEWS	-0.215*	-0.269***	-0.131	0.302	0.270	0.186	0.094	0.067	0.091
	[-1.892]	[-2.635]	[-1.261]	[1.457]	[1.360]	[0.918]	[0.803]	[0.616]	[0.804]
NEWS(-1)	-0.101	-0.083	-0.122	-0.179	-0.140	-0.208	0.022	0.087	-0.037
	[-0.885]	[-0.810]	[-1.171]	[-0.855]	[-0.702]	[-1.019]	[0.188]	[0.800]	[-0.327]
NEWS(-2)	-0.035	-0.004	-0.120	-0.310	-0.182	-0.319	-0.022	-0.059	-0.027
	[-0.303]	[-0.035]	[-1.153]	[-1.471]	[-0.905]	[-1.553]	[-0.186]	[-0.542]	[-0.238]
NEWS(-3)	-0.007	-0.020	-0.034	-0.048	0.106	-0.048	-0.133	-0.122	-0.088
	[-0.061]	[-0.197]	[-0.328]	[-0.229]	[0.524]	[-0.234]	[-1.131]	[-1.124]	[-0.777]
NEWS(-4)	-0.049	-0.139	0.063	-0.043	0.154	-0.037	-0.028	-0.011	-0.046
	[-0.428]	[-1.361]	[0.614]	[-0.207]	[0.769]	[-0.182]	[-0.242]	[-0.104]	[-0.406]
NEWS(-5)	0.037	-0.021	-0.029	-0.053	-0.149	0.059	0.092	0.139	0.167
	[0.339]	[-0.214]	[-0.293]	[-0.267]	[-0.785]	[0.308]	[0.856]	[1.346]	[1.584]

Table 4.15 (continued)

Dependent variable	HSIF detrended log open interest			KLCIF detrended log open interest			SiMSCIF detrended log open interest		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
ABSNEWS	-0.343**	-0.055	0.075	-0.001	0.124	-0.188	-0.153	-0.315**	0.004
	[-1.985]	[-0.361]	[0.478]	[-0.004]	[0.438]	[-0.63]	[-0.877]	[-1.965]	[0.024]
ABSNEWS(-1)	0.015	-0.002	0.099	0.153	-0.213	0.107	0.009	0.075	0.233
	[0.087]	[-0.010]	[0.629]	[0.484]	[-0.751]	[0.356]	[0.053]	[0.463]	[1.389]
ABSNEWS(-2)	-0.062	0.007	-0.061	0.265	0.253	0.011	0.148	0.169	0.127
	[-0.358]	[0.045]	[-0.389]	[0.837]	[0.891]	[0.037]	[0.847]	[1.039]	[0.757]
ABSNEWS(-3)	0.014	-0.082	0.071	-0.539*	-0.536*	-0.359	0.142	-0.044	0.038
	[0.082]	[-0.536]	[0.448]	[-1.704]	[-1.887]	[-1.194]	[0.809]	[-0.269]	[0.228]
ABSNEWS(-4)	-0.079	0.063	-0.028	-0.072	-0.107	-0.185	-0.033	0.161	-0.294*
	[-0.457]	[0.414]	[-0.179]	[-0.228]	[-0.378]	[-0.615]	[-0.190]	[0.990]	[-1.753]
ABSNEWS(-5)	0.015	0.191	0.026	0.232	0.241	-0.390	-0.089	-0.015	-0.134
	[0.089]	[1.255]	[0.164]	[0.733]	[0.849]	[-1.299]	[-0.511]	[-0.093]	[-0.805]
X2(1)Joint [NEWS]	1.216	3.703	3.786	3.762	2.667	4.318	1.983	3.799	3.084
p-value	0.943	0.593	0.581	0.584	0.751	0.505	0.852	0.579	0.687
X2(1)Joint [ABSNEWS]	0.357	2.014	0.767	4.469	5.688	3.549	1.700	2.369	6.246
p-value	0.996	0.847	0.979	0.484	0.338	0.616	0.889	0.796	0.283

Table 4.16 Predicting Number of Contract Using Bad News Factor

This table presents the coefficients β_{4i} and θ_{4i} estimated for equation (13) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and contract volume.

$$V_t = \alpha_4 + \sum_{i=0}^{5} \beta_{4i} N_{t-i} + \sum_{i=0}^{5} \theta_{4i} |N_{t-i}| + \sum_{i=0}^{5} \gamma_{4i} R_{t-i} + \sum_{i=1}^{5} \delta_{4i} V_{t-i} + \sum_{j=1}^{14} \lambda_{4j} EXOG_j + \varepsilon_{3t} \qquad (3.7)$$

The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The coefficients denote the impact of one-standard deviation change in bad news factor and change in absolute bad news factors respectively on the percentage change of open interest of HSIF, KLCIF and SiMSCIF. I expect the β_{4i} are negative and the θ_{4i} are positive. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, with 1 degree of freedom. The t values are reported in parentheses [ ]. *, ** and *** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent variable	HSIF detrended log contract volume			KLCIF detrended log contract volume			SiMSCIF detrended log contract volume		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
NEWS	0.052	-0.353	0.401	0.198	0.455	-0.093	0.069	-0.389	-0.376
	[0.078]	[-0.583]	[0.653]	[0.244]	[0.586]	[-0.117]	[0.068]	[-0.409]	[-0.382]
NEWS(-1)	-0.140	-0.246	-0.146	-0.069	-0.127	-0.706	1.464	0.650	1.608
	[-0.207]	[-0.406]	[-0.237]	[-0.084]	[-0.163]	[-0.883]	[1.426]	[0.685]	[1.630]
NEWS(-2)	0.198	-0.189	-0.549	-0.566	-0.904	-0.438	-0.312	-0.970	-0.163
	[0.293]	[-0.311]	[-0.891]	[-0.687]	[-1.148]	[-0.545]	[-0.304]	[-1.022]	[-0.165]
NEWS(-3)	-0.557	-0.659	-0.085	-0.223	0.376	-0.376	-1.161	-0.531	-0.432
	[-0.821]	[-1.079]	[-0.138]	[-0.272]	[0.477]	[-0.47]	[-1.133]	[-0.562]	[-0.438]
NEWS(-4)	0.868	0.402	0.642	0.012	0.866	0.017	-0.108	-0.292	-0.236
	[1.290]	[0.665]	[1.051]	[0.014]	[1.105]	[0.021]	[-0.106]	[-0.309]	[-0.241]
NEWS(-5)	0.201	0.197	0.146	0.993	0.601	0.705	1.278	1.502*	1.105
	[0.314]	[0.338]	[0.247]	[1.292]	[0.813]	[0.942]	[1.366]	[1.677]	[1.209]

Table 4.16 (Continued)

Dependent variable	HSIF detrended log contract volume			KLCIF detrended log contract volume			SiMSCIF detrended log contract volume		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
ABSNEWS	6.355	0.238	-0.852	23.657	1.025	0.207	0.954	0.265	2.199
	[-1.376]	[0.264]	[-0.919]	[-0.620]	[0.928]	[0.177]	[0.632]	[0.189]	[1.523]
ABSNEWS(-1)	-0.687	-1.502*	-0.025	-2.255*	-2.993***	-0.866	-2.608*	-1.383	-0.750
	[-0.666]	[-1.660]	[-0.026]	[-1.824]	[-2.694]	[-0.737]	[-1.714]	[-0.983]	[-0.515]
ABSNEWS(-2)	-0.907	-0.127	-1.459	1.267	-0.469	0.834	-0.464	-0.272	0.102
	[-0.880]	[-0.141]	[-1.561]	[1.023]	[-0.422]	[0.709]	[-0.305]	[-0.193]	[0.070]
ABSNEWS(-3)	-0.404	-0.063	-0.923	-0.848	-0.833	-0.755	0.493	1.042	0.631
	[-0.393]	[-0.069]	[-0.986]	[-0.685]	[-0.748]	[-0.641]	[0.324]	[0.739]	[0.434]
ABSNEWS(-4)	-1.881*	-1.099	-0.739	0.230	1.006	-0.096	-2.137	-0.915	-2.952**
	[-1.826]	[-1.212]	[-0.789]	[0.185]	[0.905]	[-0.081]	[-1.406]	[-0.648]	[-2.025]
ABSNEWS(-5)	-2.908***	-0.994	-2.317**	1.164	0.952	-0.001	-0.025	1.520	0.211
	[-2.827]	[-1.101]	[-2.482]	[0.939]	[0.856]	[-0.001]	[-0.016]	[1.086]	[0.145]
X^2(1)Join [NEWS]	2.543	1.939	1.902	2.110	3.538	2.151	5.000	4.425	4.424
p-value	0.770	0.858	0.863	0.834	0.618	0.828	0.416	0.490	0.490
X^2(1)Join [ABSNEWS]	12.520**	5.963	10.707*	5.835	9.464*	1.502	4.979	2.878	4.515
p-value	0.028	0.310	0.058	0.323	0.092	0.913	0.418	0.719	0.478

4.5.3 Predicting index futures returns using bad news factor

One of the central questions of this research is whether the news factor is useful in predicting index futures returns and whether the impact of the news factor on index futures returns persist or reverse during later days. Theoretically, in an efficient market, asset prices immediately incorporate all available information. News from the day before is stale information and should have no impact on prices. Both the news factor as investor sentiment theory and the news factor and information theory imply that *Pessimism*, *Negative* and *Weak* news factors predict negative growth in returns in the short-term. However, the information theory predicts that the initial decrease of the return persists while the sentiment or noise traders theory see the initial decrease of returns as overreactions of investors and the returns will reverse in the longer-term.

Table 4.13 compares the coefficients, β_{2i} ,estimated for the equation (3.2) using SVAR, to test Hypothesis 2 to Hypothesis 5. The coefficients indicate the impact of one standard deviation change in the bad news factor on the index futures returns in basis points. I test Hypothesis 2 using the coefficient of the news factors' contemporaneous term and first lag individually $(i.e.\beta_{20} = 0 \; or \; \beta_{21} = 0)$; Hypothesis 3 using the joint test of the summation of the second lag to fifth lag $(i.e.\beta_{22} + \beta_{23} + \beta_{24} + \beta_{25} = 0)$; Hypothesis 4 and Hypothesis 5 by examining the summation of the β_{2i} (i.e. $\beta_{20} + \beta_{21} + \beta_{22} + \beta_{23} + \beta_{24} + \beta_{25} \neq 0$ or equals 0 respectively).

The discussion begins with Hang Seng Index Futures. The coefficients for the contemporaneous terms are -10.680, -13.160, -11.300 for *Pessimism, Negative* and *Weak*, respectively; all are significant at 1 % level. The coefficient of the first lag of *Weak* is -6.570 and significant at α =10%. This confirms Hypothesis 2 that bad news factors negatively forecast returns. The Chi-square values for the joint test of the summation of the second to fifth lags are 14.285, 1.243 and 6.032, for *Pessimism, Negative* and *Weak*; with statistical significance at 1%, 5%, and 5% respectively. The findings support Hypothesis 3, that there are returns reversals. The last three rows in Table 4.13 illustrates the summation of the coefficients, β_{2i}, i=0,1,2,3,4, and 5. The difference is 14.370 (significant at 1%), -10.300, and -3.084. A negative difference indicates that the reversals are insufficient to offset the initial negative impact of news factor on returns, while a positive difference indicates that reversals are greater than the initial negative impact of bad news factors. I confirm the Hypothesis 4, where there are partial returns reversals, as predicted by noise traders and sentiment theory. The positive difference implies that the bad news factors reflect the investor sentiment, instead of forecasting the sentiment (Tetlock, 2007). Hypothesis 5 is not supported.

The evidence illustrated in Table 4.13, on the Kuala Lumpur Composite Index Futures (KLCIF) supports Hypothesis 2 that pessimistic news negatively forecast returns. The signs of the coefficients are all negative, significant at the 1% level except for the *Weak* news factor. The *Weak* news factors do not predict the KLCIF returns on the next day, but significantly predict the returns on the day after next; the coefficient, β_{22}, is -8.960 and the t-value is 2.276. The joint chi-square test for Hypothesis 3 is not significant; indicating no reversals on the next two to five days. However, there is some evidence of reversal, considering the significance of individual coefficients. For the *Negative* equation, on the fifth day after the news, the coefficient, β_{25}, is 6.570, at the 10% level of significance. Since Hypothesis 3 is not supported, the significant summation of coefficients of contemporaneous terms and the five lags (11.757, -31.083 and -29.850) simply means that the initial negative impact of bad news factors on returns persists, up to five days after the news released. Hypothesis 4 and Hypothesis 5 are not discussed since there is no significant reversal as conclude by Hypothesis 3.

Concerning Singapore Morgan Stanley Free Index Futures (SiMSCIF), the evidence supports Hypothesis 2 that there is negative association between pessimistic news factors and returns on the same day of the news release, but not for the next day. Concerning Hypothesis 3, there is no significant returns reversal from the following days, since the summation of the coefficients (3.300, -0.589, and 0.842) from second lag to fifth lag are jointly insignificant, given the chi-square test statistics are 0.306, 0.013 and 0.877 respectively. Despite no evidence of joint significance was found, the coefficient,β_{24}, for the fourth lag, are significantly different from zero. The coefficients are 6.250 and 6.560, both are significant at the 1% level. Hypothesis 4 and Hypothesis 5 are not discussed since there is no significant reversal as conclude by Hypothesis 3.

For all the three index futures, the magnitude of the negative impact of the bad news factor on index futures returns range from around 6 to 27 basis points, for the same and next day. Only the negative impact of pessimistic media content on KLCIF returns last up to the next two days after the news is published[29]. In addition, the timing of the price reversals varies among the three index futures contracts. The HSIF returns starts to reverse after 2 days. The joint tests' evidence on KLCIF and SiMSCIF is more consistent with information theory. There is some evidence that the KLCIF returns start to reverse after three days and SiMSCI returns start to reverse after four days. The evidence partially supports Hypothesis 4.

Table 4.14 depicts the results for Equation (3.2), in which the number of contracts traded is use as a proxy for trading volume. It is noteworthy that, for the case of HSIF, the difference between the initial change and reversals is 13.950 (the p-value is 0.109), indicating the magnitude of reversals is larger than the initial decrease. The overall results are consistent with the results in Table 4.13, when the proxy for trading volume is open interest.

[29] In the context of equity index in the spot market, Tetlock (2007) finds *Pessimism* has negative impact on the next day Dow Jones' returns.

Combined with the findings in Table 4.11 and Table 4.12, I can conclude that there is a bi-directional causality between bad news factors and index futures returns. Bad news leads to lower returns, then lower returns leads to bad news. The results are consistent with the positive feedback trading theory propose by Delong et al. (1990a).

In conclusion, the findings are consistent with earlier studies (Tetlock, 2007; Verma et al., 2008) that bad news factors extracted from routine news on stock market predicts negative returns and there is some evidence of reversals on the later days. Despite weak evidence, the findings are consistent with sentiment or noise theories.

4.5.4 Predicting trading volume using bad news factors

The number of contracts traded and the open interest are generally used as a proxy for trading volume in the derivatives market. The number of contracts traded is the sum of long positions or the sum of short positions created, which reflects the strength of the market. The open interest serves as an indicator for market depth (Kyle, 1985) which represents the number of outstanding contracts that are yet to be settled or closed. The open interest increases whenever there are new buyers and sellers initiating a long and a short position, indicating capital inflows. Oppositely, when an existing long and existing short position is closed, open interest decreases, which indicates capital outflows. When a market participant sells an existing long position to a new market participant, or when someone with an existing short position sells it to a new market participant, the open interest remains unchanged. The open interest and the number of contracts traded complement each other to provide a better explanation of trading activity. Kyle's model associates a higher number of contracts traded with more informed traders; and market depth (proxy by open interest) depending on the trading activities of hedgers that is usually non-informational. However, Fung and Patterson (1999) find an asymmetrical lead-lag relationship between the number of contracts traded and the open interest, and draw the conclusion that these two variables are not endogenously determined.

Table 4.15 presents the coefficients, β_{4i} and θ_{4i} ,estimated for equation (3.7) using a structural vector autoregressive (SVAR) regression model. While the coefficients, β_{4i}, denote the impact of one-standard deviation change in the bad news factor, the coefficient, θ_{4i} , denotes the change in absolute bad news factors respectively on the percentage change in open interest of HSIF, KLCIF and SiMSCIF.

Contemporaneous *Pessimism* and *Negative* news factors, β_{40}, significantly predict HSIF's open interest as the coefficients are -0.215 and -0.269 respectively for the full sample period, from 1/1996 to 12/2008, supporting Hypothesis 6 that bad news factors are a proxy for trading costs. The absolute value of *Pessimism* significantly predicts the detrended log open interest, as the coefficient, β_{40}, is -0.343 and significant at 5%. The negative sign indicates that the investors had decided to close their position in the period of extremely high or low pessimism. However, the sign is opposite to the prediction of pessimism as a measure of sentiment and proxy for risk aversion.

For KLCIF, there is no evidence of bad news factors being a proxy of trading costs. There is some evidence although not strong, that absolute value of *Pessimism* and *Negative* predicts the detrended log open interest, the coefficients, θ_{43}, are -0.539 and -0.536, but both are significant at 10%. These negative impacts are delayed for three days.

The Straits Time's bad news factors do not predict the detrended log open interest. The absolute values of *Negative* news sentiment predict a negative relationship with detrended log open interest; the coefficient, θ_{40}, is -0.315 and significant at 5%.

I summarise two points base on Table 4.15. First, only HSIF shows evidence that bad news factors proxy for trading costs on the same day of news release, thus, no evidence is found for KLCIF and SiMSCIF. Second, the evidence is too weak to conclude that the bad news factors proxy for investor sentiment because the coefficients, θ_{4i} , are inconsistent among the three contracts and bad news factors, at different lags.

Based on Table 4.16, the bad news factors do not significantly predict the detrended log contract volume. However, there is some evidence that the absolute value of bad news factors predicts negative detrended log contract volume.

Despite the inconsistencies, the findings do shed light on the different trading behaviour of informed/ uninformed futures traders. Usually speculators are informed and thus they are more likely to be day traders, and less likely keep their positions unclosed at the end of a trading day. Thus open interest reported at the close of a trading day can proxy for the uninformed hedgers' position to meet their liquidity needs (Bessembinder & Seguin, 1993).

The well-studied volatility-volume relationship can indirectly explain the negative relationship between absolute values of news factors and trading volume. Karpoff (1987) conducted a comprehensive survey on the relationship between price changes (i.e. volatility) and trading volume and reveals that the absolute value of the price change and price change per se are popular measures for volatility. These studies conclude that volatility is positively related to volume and support the mixture of distribution hypothesis. Another hypothesis, namely the sequential arrival of information hypothesis predicts the same positive volatility-volume relationship, but outlines the roles of different types of traders and the timing of the arrival of information to these two trader types. Daigler and Wiley (1999) find a negative volatility-volume relationship for those who have early access to order flow information, i.e.

clearing members and floor traders, using the volatility estimator developed by Garman and Klass (1980).

I generate bad news factors from routine financial news articles that summarise previous day's market movements. Any extreme changes in returns will lead to extreme changes in bad news factors. Hence, returns volatility (i.e. the magnitude of price changes in either direction, up or down) is positively related to the absolute value of bad news factors. If this conjecture is true, then a positive volatility-volume relationship also implies a negative relationship between the absolute value of bad news factors and trading volume.

4.5.5 Sub-sample period analysis

The behaviour of market variables varies during different periods. Since the other factors as suggested by the extant literature already included as exogenous variables in the VAR model, I divide the sample period into two equal sub-samples to examine the consistency of the prediction over time. The behaviour of relationships among the index futures returns, bad news factors and trading volume at different is examined. The full sample period of the South China Morning Post and The New Straits Times covers 1/1996 to 12/2008. The samples are then split in half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times is from 1/1999 through 12/2008, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008.

4.5.5.1 Predicting bad news factors using regional index futures returns

Table 4.17 and Table 4.18 present the sub-sample analysis for predicting bad news factors using regional index futures returns, using detrended log open interest and detrended log number of contract trades as the proxies for trading volume respectively. Overall, the results are very similar to the analysis of the full sample. The magnitude of coefficients from sub-sample 2 are larger than sub-sample 1, especially for KLCIF, for example a 1% decrease in KLCIF prior day's returns predicts the *Pessimism* news factor decrease by 30.4% of standard deviation in sub-sample 2, while only 12.3% in sub-sample 1. The impact of returns on bad news factors lasts for 3 days in sub-sample 1, but only 2 days in sub-sample 2.

4.5.5.2 Predicting index futures returns using bad news factors

Table 4.19 and Table 4.20 present the sub-sample analysis for predicting index futures returns using bad news factors, using detrended log open interest and detrended log number of contract trades as the proxies for trading volume respectively. The results of these two tables are consistent.

All three index futures returns are negatively related to the same day's bad news factors in both sub-samples, except for SiMSCIF in sub-sample 1. The impact of bad news factors do not carry-over to the next day, except for KLCIF, which lasts for the next two days in the sub-sample 1 and next day in sub-sample 2.

For HISF and KLCIF, the magnitude of prediction reduces over time. For example, a one standard deviation increase in *Weak* news factor predicts the same day KLCIF returns decrease by 28.29 basis points in sub-sample 1, but only 5.12 basis points in sub-sample 2.

There are significant reversals for HSIF in sub-sample 1 but not in sub-sample 2. Conversely, there are significant reversals for KLCIF in sub-sample 2, but not in sub-sample 1. Sentiment is more prevalent in Hong Kong during sub-sample 1 due to the transfer of Hong Kong from the United Kingdom to China; whereas KLCIFs' investors are more sensitive to investor sentiment in 2008, possibly due to the ruling party only winning 50.27% vote in the general election, which is the lowest in Malaysia's history.

4.5.5.3 Predicting trading volume using bad news factors

Table 4.21 and Table 4.22 present the sub-sample analysis for predicting trading volume using bad news factors, using detrended log open interest and detrended log contract volume as the proxies for trading volume respectively. The results of these two tables are somewhat different.

Table 4.21 shows little evidence of the predictive power of bad news factors on detrended log open interest. In sub-sample 1, the *Pessimism* and *Negative* news factors of The Straits Times is positively related to the SiMSCIF's next day return. The results contradict the prediction of Hypothesis 6 of news factor as a measure of trading costs. The coefficients, β_{41}, are 0.301 and 0.361 respectively. This indicates that the disagreement among investors increases, resulting in an increase of open interest on the next day. In sub-sample 2, there is evidence that South China Morning Post's *Negative* and *Weak* news factors predicts lower HSIF open interest on the same day; the coefficients, β_{40}, are -0.355 and -0.241. Conversely, the same coefficients are 0.434, 0.464 and 0.417 respectively for *Pessimism, Negative and Weak* news factors for the New Straits Times.

Table 4.22 indicates that the impact of bad news factors on detrended log contract volume is very limited. In sub sample 1, only the *Negative* news factors of the New Straits Times significantly predicts the next day KLCIF contract volume, the coefficient, β_{41} , is 2.038; the *Pessimism* news factor of The Straits Times predicts the next day SiMSCIF contract volume. The impact of absolute value of news factors is stronger in sub-sample 1 than in sub-sample 2. All the news sentiment positively predicts the same day SiMSCIF contract volume; the coefficients, θ_{40}, are 5.779, 3.068, and 5.228 for *Pessimism, Negative* and *Weak* respectively. Absolute value of *Pessimism* and *Negative* news factors negatively predicts the KLCIF next day contract volume; the coefficients, θ_{41}, are -2.894 and -3.758 respectively. There is evidence that extreme high or low news sentiment negatively predicts the HSIF for the next 4 or 5 days contract volume. The coefficient, θ_{44}, are -4.424 and -2.839 for *Pessimism* and *Negative* news factors.

Table 4.17 Sub-sample Analysis: Predicting Bad News Factor Using Regional Index Futures Returns and Open Interest

This table presents the coefficients γ_{1i} estimated for equation (1) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and trading volume (in open interest).

$$N_t = \alpha_1 + \sum_{i=1}^{5} \beta_{1i} N_{t-i} + \sum_{i=0}^{5} \gamma_{1i} R_{t-i} + \sum_{i=0}^{5} \delta_{1i} V_{t-i} + \sum_{j=1}^{14} \lambda_{1j} EXOG_j + \varepsilon_{1t} \qquad (3.1)$$

I expect the γ_{1i} to be negative. The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The coefficients denote the impact of 1% decrease in HSIF, KLCIF and SiMSCIF returns on bad news factor in the unit of percentage of standard deviation. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, test for $\sum_{i=0}^{5} \gamma_{1i} = 0$; i=1,2,3,4,and 5; with degree of freedom equals to 1. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent variable	Bad News Factor			Bad News Factor			Bad News Factor		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Sub Sample 1									
Returns(-1)	-15.910***	-14.488***	-14.127***	-12.295***	-11.902***	-13.326***	-27.225***	-25.281***	-24.665***
	[-15.597]	[-13.465]	[-13.063]	[-13.348]	[-12.533]	[-14.333]	[-16.933]	[-13.667]	[-14.703]
Returns(-2)	-7.794***	-7.289***	-7.577***	-3.900***	-3.244***	-4.354***	-6.208***	-7.202***	-6.578***
	[-6.947]	[-6.278]	[-6.495]	[-3.673]	[-2.983]	[-4.037]	[-3.477]	[-3.632]	[-3.619]
Returns(-3)	-2.606**	-2.107*	-3.596***	-2.138**	-0.281	-3.020***	-2.401	-2.688	-2.116
	[-2.292]	[-1.799]	[-3.057]	[-2.027]	[-0.260]	[-2.818]	[-1.338]	[-1.348]	[-1.158]
Returns(-4)	-0.782	-0.067	-0.071	0.257	-0.122	0.460	0.867	-0.469	0.249
	[-0.688]	[-0.057]	[-0.061]	[0.242]	[-0.112]	[0.427]	[0.483]	[-0.235]	[0.136]
Returns(-5)	-1.496	-0.524	-1.350	-1.168	-0.914	-1.186	0.709	0.731	0.735
	[-1.328]	[-0.451]	[-1.156]	[-1.101]	[-0.841]	[-1.101]	[0.399]	[0.370]	[0.405]
$X^2(1)$Joint	278.780***	208.158***	203.810***	193.125***	164.154***	226.847***	300.147***	200.287***	229.199***
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.17 (Continued)

Dependent variable	Bad News Factor			Bad News Factor			Bad News Factor		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
<u>Sub Sample 2</u>									
Returns(-1)	-20.552***	-18.289***	-19.498***	-30.395***	-29.240***	-32.343***	-29.384***	-28.468***	-28.832***
	[-15.266]	[-11.784]	[-13.035]	[-15.557]	[-14.163]	[-16.058]	[-15.583]	[-14.318]	[-14.257]
Returns(-2)	-2.910*	-1.618	-3.614**	-10.976***	-11.572***	-11.597***	-5.850***	-5.092**	-5.633**
	[-1.918]	[-0.953]	[-2.186]	[-5.208]	[-5.254]	[-5.317]	[-2.769]	[-2.326]	[-2.532]
Returns(-3)	0.585	1.522	-0.532	-2.803	-3.259	-1.492	-2.708	-2.568	-4.267*
	[0.386]	[0.899]	[-0.322]	[-1.323]	[-1.473]	[-0.681]	[-1.278]	[-1.174]	[-1.913]
Returns(-4)	-0.988	-0.730	-0.147	-3.295	-3.328	-4.880**	-2.553	-1.116	-4.804**
	[-0.649]	[-0.430]	[-0.089]	[-1.559]	[-1.512]	[-2.229]	[-1.200]	[-0.509]	[-2.138]
Returns(-5)	-1.615	-1.233	0.311	1.875	1.674	1.828	-1.088	1.088	-2.821
	[-1.072]	[-0.731]	[0.190]	[0.894]	[0.768]	[0.841]	[-0.521]	[0.504]	[-1.277]
X^2(1)Joint	234.295***	140.629***	171.290***	262.245***	219.765***	281.203***	243.464***	207.386***	204.493***
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.18 Sub-sample Analysis: Predicting Bad News Factor Using Regional Index Futures Returns and Number of Contracts Traded

This table presents the coefficients γ_{1i} estimated for equation (1) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and number of contract trades.

$$N_t = \alpha_1 + \sum_{i=1}^{5} \beta_{1i} N_{t-i} + \sum_{i=0}^{5} \gamma_{1i} R_{t-i} + \sum_{i=0}^{5} \delta_{1i} V_{t-i} + \sum_{j=1}^{14} \lambda_{1j} EXOG_j + \varepsilon_{1t} \quad (3.1)$$

I expect the γ_{1i} to be negative. The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The coefficients denote the impact of 1% increase in HSIF, KLCIF and SiMSCIF returns on bad news factor in standard deviation. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, test for $\sum_{i=0}^{5} \gamma_{1i} = 0$; i=1,2,3,4,and 5; with degree of freedom equals to 1. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively

Dependent variable	Bad News Factor			Bad News Factor			Bad News Factor		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Sub Sample 1									
Returns(-1)	-15.987***	-14.519***	-14.227***	-12.187***	-11.787***	-13.130***	-27.279***	-25.271***	-24.705***
	[-15.740]	[-13.551]	[-13.205]	[-13.278]	[-12.446]	[-14.179]	[-17.063]	[-13.692]	[-14.776]
Returns(-2)	-7.759***	-7.230***	-7.521***	-4.055***	-3.353***	-4.588***	-6.286***	-7.162***	-6.648***
	[-6.930]	[-6.242]	[-6.455]	[-3.831]	[-3.093]	[-4.273]	[-3.538]	[-3.620]	[-3.670]
Returns(-3)	-2.761**	-2.211*	-3.688***	-2.437**	-0.420	-3.212***	-2.155	-2.579	-1.769
	[-2.437]	[-1.896]	[-3.145]	[-2.314]	[-0.39]	[-3.003]	[-1.207]	[-1.295]	[-0.971]
Returns(-4)	-0.654	0.151	0.023	0.178	-0.117	0.528	0.999	-0.440	0.494
	[-0.576]	[0.129]	[0.020]	[0.168]	[-0.108]	[0.491]	[0.559]	[-0.221]	[0.271]
Returns(-5)	-1.465	-0.500	-1.284	-1.279	-0.964	-1.276	0.934	0.959	0.879
	[-1.301]	[-0.431]	[-1.098]	[-1.208]	[-0.890]	[-1.188]	[0.527]	[0.485]	[0.485]
$X^2(1)$Joint	285.138***	211.438***	209.208***	193.042***	162.392***	225.512***	304.061***	200.488***	231.061***
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.18 (Continued)

Dependent variable	Bad News Factor			Bad News Factor			Bad News Factor		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Sub Sample 2									
Returns(-1)	-20.103***	-17.868***	-19.185***	-30.322***	-29.163***	-32.328***	-28.813***	-28.123***	-28.076***
	[-14.935]	[-11.503]	[-12.788]	[-15.457]	[-14.077]	[-15.979]	[-15.264]	[-14.119]	[-13.898]
Returns(-2)	-2.790*	-1.565	-3.725**	-10.800***	-11.333***	-11.285***	-5.622***	-4.725**	-5.394**
	[-1.843]	[-0.921]	[-2.247]	[-5.107]	[-5.130]	[-5.156]	[-2.667]	[-2.160]	[-2.436]
Returns(-3)	0.863	1.712	-0.283	-2.473	-2.962	-0.974	-2.287	-2.164	-3.965*
	[0.571]	[1.012]	[-0.172]	[-1.163]	[-1.335]	[-0.443]	[-1.081]	[-0.99]	[-1.785]
Returns(-4)	-0.665	-0.520	0.029	-3.261	-3.320	-4.695**	-2.824	-1.125	-5.357**
	[-0.439]	[-0.306]	[0.017]	[-1.538]	[-1.505]	[-2.139]	[-1.332]	[-0.515]	[-2.398]
Returns(-5)	-1.301	-0.992	0.492	1.949	1.701	2.109	-1.588	0.415	-3.361
	[-0.864]	[-0.588]	[0.299]	[0.926]	[0.778]	[0.967]	[-0.763]	[0.193]	[-1.528]
$X^2(1)$Joint	224.335***	134.242***	165.501***	257.971***	216.089***	277.119***	233.370***	200.639***	194.858***
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.19 Sub-sample Analysis: Predicting Regional Index Futures Returns Using Bad News Factor and Open Interests

This table presents the coefficients β_{2i} estimated for equation (2) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and trading volume (in open interest).

$$R_t = \alpha_2 + \sum_{i=0}^{5} \beta_{2i} N_{t-i} + \sum_{i=1}^{5} \gamma_{2i} R_{t-i} + \sum_{i=0}^{5} \delta_{2i} V_{t-i} + \sum_{j=1}^{14} \lambda_{2j} EXOG_j + \varepsilon_{2t} \qquad (3.2)$$

I expect the β_{2_i} to be negative when i=0 and 1, the β_{2_i} are positive when i=2,3,4 and 5. The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The coefficients denote the impact of one standard deviation change in bad news factor on the index futures return in basis points. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, with 1 degree of freedom. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent Variable	HSIF			KLCIF			SiMSCIF		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News measure	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Sub Sample 1									
NEWS	-11.460*	-15.850***	-12.390**	-27.590***	-21.620***	-28.290***	-4.830	-1.720	-2.160
	[-1.833]	[-2.675]	[-2.100]	[-3.671]	[-2.967]	[-3.800]	[-0.922]	[-0.380]	[-0.431]
NEWS(-1)	1.490	0.407	-16.600	-9.280	-12.920*	-37.480	5.500	0.326	-0.912
	[0.236]	[0.068]	[-0.707]	[-1.206]	[-1.739]	[-1.206]	[1.040]	[0.071]	[-0.180]
NEWS(-2)	14.490**	5.570	12.650**	-14.160*	-14.290*	-17.430**	-5.220	-0.747	-7.790
	[2.286]	[0.933]	[2.120]	[-1.825]	[-1.900]	[-2.270]	[-0.989]	[-0.164]	[-1.540]
NEWS(-3)	13.240**	8.380	15.860**	-8.200	-11.190	-8.290	4.140	-1.400	5.080
	[2.087]	[1.395]	[2.663]	[-1.060]	[-1.488]	[-1.085]	[0.787]	[-0.307]	[1.008]
NEWS(-4)	11.460*	5.260	4.770	1.520	0.449	5.390	3.090	1.780	4.840
	[1.837]	[0.893]	[0.810]	[0.196]	[0.060]	[0.708]	[0.588]	[0.393]	[0.963]
NEWS(-5)	-5.600	-11.040*	-4.530	0.775	1.620	0.708	-2.280	-6.370	-1.500
	[-0.949]	[-1.947]	[-0.799]	[0.108]	[0.231]	[0.100]	[-0.474]	[-1.473]	[-0.319]
sum (-2) to -(5)	48.753	14.880	43.140	-12.859	-11.350	-9.936	2.832	-4.195	4.737
$X^2(1)$	9.397***	0.786	8.461***	2.412	3.816*	2.545	0.001	0.794	0.006
p-value	0.002	0.375	0.004	0.120	0.051	0.111	0.977	0.373	0.940
Sum (0) to (-5)	23.620	-7.273	-0.240	-56.935	-57.951	-85.392	0.400	-8.131	-2.442
$X^2(1)$	3.022	0.423	1.007	12.700	15.637	14.133	0.001	0.793	0.056
p-value	0.082	0.516	0.316	0.000***	0.000***	0.000***	0.970	0.373	0.813

Table 4.19 (Continued)

Dependent Variable	HSIF			KLCIF			SiMSCIF		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News measure	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Sub Sample 2									
NEWS	-9.390*	-10.150**	-8.950**	-6.100*	-10.170***	-5.120	-14.380***	-9.620**	-15.750***
	[-1.953]	[-2.431]	[-2.069]	[-1.826]	[-3.224]	[-1.576]	[-3.023]	[-2.130]	[-3.551]
NEWS(-1)	-4.090	-15.500	-8.280*	-5.260	-10.080***	-3.130	-0.195	-14.560	-18.310
	[-0.839]	[-1.270]	[-1.888]	[-1.560]	[-3.160]	[-0.958]	[-0.040]	[-1.081]	[-0.565]
NEWS(-2)	0.894	-1.090	-0.985	3.730	3.940	-0.782	0.136	-0.039	-0.543
	[0.183]	[-0.258]	[-0.225]	[1.108]	[1.234]	[-0.240]	[0.028]	[-0.009]	[-0.119]
NEWS(-3)	6.920	4.790	4.610	-1.170	0.868	-3.250	-1.500	1.220	-7.450
	[1.419]	[1.128]	[1.055]	[-0.348]	[0.272]	[-1.000]	[-0.309]	[0.267]	[-1.629]
NEWS(-4)	4.710	1.720	-1.970	2.540	-0.362	4.090	8.780*	7.070	7.960*
	[0.966]	[0.406]	[-0.451]	[0.758]	[-0.114]	[1.263]	[1.816]	[1.553]	[1.756]
NEWS(-5)	9.560**	2.260	6.280	7.310**	9.240***	1.980	-0.665	-1.810	0.800
	[2.107]	[0.561]	[1.503]	[2.322]	[3.078]	[0.652]	[-0.148]	[-0.419]	[0.189]
sum (-2) to -(5)	22.084	10.973	14.432	22.604	34.401	5.807	10.122	9.915	6.346
$X^2(1)$	6.348**	1.293	1.204	4.574**	6.851***	0.143	0.724	0.865	0.012
p-value	0.012	0.255	0.273	0.033	0.009	0.705	0.395	0.352	0.913
Sum(0) to (-5)	8.604	-14.677	-2.798	11.244	14.151	-2.443	-4.453	-14.265	-27.714
$X^2(1)$	0.638	0.911	1.130	0.021	1.040	0.878	0.638	0.932	4.098**
p-value	0.424	0.340	0.288	0.885	0.308	0.349	0.425	0.334	0.043

Table 4.20 Sub-sample Analysis: Predicting Regional Index Futures Returns Using Bad News Factor and Number of Contracts Traded

This table presents the coefficients β_{2i} estimated for equation (2) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and number of contracts traded.

$$R_t = \alpha_2 + \sum_{i=0}^{5} \beta_{2i} N_{t-i} + \sum_{i=1}^{5} \gamma_{2i} R_{t-i} + \sum_{i=0}^{5} \delta_{2i} V_{t-i} + \sum_{j=1}^{14} \lambda_{2j} EXOG_j + \varepsilon_{2t} \qquad (3.2)$$

I expect the β_{2i} to be negative when i=0 and 1, the β_{2i} are positive when i=2,3,4 and 5. The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The coefficients denote the impact of one standard deviation change in bad news factor on the index futures return in basis points. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, with 1 degree of freedom. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent Variable	HSIF			KLCIF			SiMSCIF		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News measure	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Sub Sample 1									
NEWS	-12.840**	-17.260***	-13.240**	-29.220***	-22.350***	-28.810***	-5.040	-1.800	-2.270
	[-2.044]	[-2.900]	[-2.235]	[-3.870]	[-3.056]	[-3.850]	[-0.957]	[-0.396]	[-0.452]
NEWS(-1)	1.750	0.613	-4.220	-9.910	-13.450*	-9.790	5.050	0.280	-1.250
	[0.276]	[0.103]	[-0.706]	[-1.281]	[-1.802]	[-1.277]	[0.950]	[0.061]	[-0.246]
NEWS(-2)	14.280**	5.790	12.390**	-13.850*	-13.980*	-16.830**	-6.080	-1.640	-8.280
	[2.246]	[0.966]	[2.069]	[-1.776]	[-1.850]	[-2.178]	[-1.148]	[-0.359]	[-1.632]
NEWS(-3)	13.220**	8.490	15.830***	-7.430	-10.010	-6.840	3.780	-1.500	4.950
	[2.077]	[1.409]	[2.647]	[-0.958]	[-1.324]	[-0.891]	[0.715]	[-0.330]	[0.980]
NEWS(-4)	10.910*	4.780	4.110	0.700	-0.504	5.310	3.000	1.900	5.040
	[1.743]	[0.808]	[0.695]	[0.090]	[-0.067]	[0.695]	[0.568]	[0.419]	[1.000]
NEWS(-5)	-6.300	-11.040*	-5.690	1.810	2.270	1.950	-1.650	-5.810	-0.875
	[-1.066]	[-1.942]	[-1.001]	[0.251]	[0.323]	[0.275]	[-0.342]	[-1.341]	[-0.187]
sum (-2) o -(5)	32.110	8.020	26.640	-18.770	-22.224	-16.410	-0.950	-7.050	0.835
$X^2(1)$	8.564***	0.755	7.258***	2.107	3.418*	1.768	0.011	0.866	0.010
p-value	0.003	0.385	0.007	0.147	0.065	0.184	0.917	0.352	0.922
Sum(0) to (-5)	21.020	-8.627	9.180	-57.900	-58.024	-55.010	-0.940	-8.570	-2.685
$X^2(1)$	2.383	0.591	0.574	13.060	15.574	13.039	0.007	0.879	0.068
p-value	0.123	0.442	0.449	0.000	0.000	0.000	0.933	0.349	0.794

Table 4.20 (Continued)

Dependent Variable	HSIF			KLCIF			SiMSCIF		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News measure	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Sub Sample 2									
NEWS	-8.550*	-9.860**	-8.370*	-6.090*	-10.160***	-4.950	-12.460***	-7.510*	-14.220***
	[-1.771]	[-2.352]	[-1.929]	[-1.830]	[-3.232]	[-1.531]	[-2.626]	[-1.668]	[-3.208]
NEWS(-1)	-4.070	-5.720	-8.550*	-5.310	-10.120***	-3.210	0.308	-5.340	-1.930
	[-0.831]	[-1.356]	[-1.951]	[-1.579]	[-3.183]	[-0.985]	[0.064]	[-1.178]	[-0.429]
NEWS(-2)	1.870	-0.025	-0.640	3.810	3.980	-0.792	-0.057	-0.282	-0.079
	[0.381]	[-0.006]	[-0.146]	[1.133]	[1.249]	[-0.244]	[-0.012]	[-0.062]	[-0.017]
NEWS(-3)	6.210	3.690	4.270	-1.040	1.140	-2.980	-2.140	-0.007	-8.170*
	[1.269]	[0.868]	[0.978]	[-0.310]	[0.358]	[-0.917]	[-0.443]	[-0.002]	[-1.797]
NEWS(-4)	3.790	0.997	-2.700	2.790	-0.133	4.390	9.610**	8.160*	9.440**
	[0.778]	[0.235]	[-0.620]	[0.834]	[-0.042]	[1.357]	[2.001]	[1.804]	[2.090]
NEWS(-5)	9.500**	2.060	5.860	7.040**	8.940***	1.730	-2.370	-3.460	-1.050
	[2.090]	[0.513]	[1.406]	[2.240]	[2.974]	[0.570]	[-0.532]	[-0.809]	[-0.249]
sum (-2) o -(5)	21.370	6.723	6.790	12.600	13.927	2.348	5.043	4.411	0.141
$X^2(1)$	5.951**	1.018	0.905	4.754**	7.108***	0.194	0.410	0.413	0.000
p-value	0.015	0.313	0.341	0.029**	0.008	0.660	0.522	0.520	0.984
Sum(0) to (-5)	8.750	-8.858	-10.130	1.200	-6.353	-5.812	-7.109	-8.439	-16.009
$X^2(1)$	0.657	1.192	1.381	0.028	0.978	0.787	0.532	1.026	3.447*
p-value	0.418	0.275	0.240	0.867	0.323	0.375	0.466	0.311	0.063

Table 4.21 Sub-sample Analysis: Predicting Open Interest Using Bad News Factor

This table presents the coefficients β_{4i} and θ_{4i} estimated for equation (13) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and trading volume (in open interest).

$$V_t = \alpha_4 + \sum_{i=0}^{5} \beta_{4i} N_{t-i} + \sum_{i=0}^{5} \theta_{4i} |N_{t-i}| + \sum_{i=0}^{5} \gamma_{4i} R_{t-i} + \sum_{i=1}^{5} \delta_{4i} V_{t-i} + \sum_{j=1}^{14} \lambda_{4j} EXOG_j + \varepsilon_{3t} \qquad (3.7)$$

I expect the β_{4i} to be negative and the θ_{4i} are positive. The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The coefficients denote the impact of one-standard deviation change in bad news factor and change in absolute bad news factors respectively on the percentage change of open interest of HSIF, KLCIF and SiMSCIF. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, with 1 degree of freedom. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively

Dependent variable	HSIF detrended log open interest			KLCIF detrended log open interest			SiMSCIF detrended log open interest		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
<u>Panel A Sub Sample 1</u>									
NEWS	-0.215	-0.157	-0.008	0.105	0.029	-0.107	0.000	-0.306	-0.015
	[-1.445]	[-1.120]	[-0.055]	[0.313]	[0.090]	[-0.322]	[-0.002]	[-1.315]	[-0.085]
NEWS(-1)	-0.015	0.016	-0.067	-0.223	-0.035	-0.245	0.301*	0.316**	0.124
	[-0.097]	[0.112]	[-0.475]	[-0.658]	[-0.105]	[-0.728]	[1.682]	[1.996]	[0.708]
NEWS(-2)	-0.125	-0.073	-0.085	-0.708**	-0.439	-0.527	-0.187	-0.199	-0.214
	[-0.830]	[-0.519]	[-0.602]	[-2.068]	[-1.321]	[-1.551]	[-1.044]	[-1.256]	[-1.226]
NEWS(-3)	0.156	0.098	0.207	0.179	0.219	0.391	-0.270	-0.205	-0.174
	[1.034]	[0.695]	[1.461]	[0.525]	[0.658]	[1.156]	[-1.510]	[-1.296]	[-1.002]
NEWS(-4)	-0.147	-0.264	0.029	-0.003	0.060	-0.121	-0.112	-0.036	-0.078
	[-0.986]	[-1.903]	[0.207]	[-0.010]	[0.181]	[-0.360]	[-0.628]	[-0.226]	[-0.451]
NEWS(-5)	-0.103	-0.066	-0.145	0.053	-0.003	0.061	0.190	0.197	0.250
	[-0.727]	[-0.493]	[-1.073]	[0.165]	[-0.008]	[0.194]	[1.173]	[1.321]	[1.553]
ABSNEWS	-0.020	-0.187	0.298	0.496	-0.117	0.218	0.054	-6.287	0.231
	[-0.089]	[-0.905]	[1.403]	[1.017]	[-0.254]	[0.450]	[0.205]	[-0.671]	[0.908]
ABSNEWS(-1)	0.151	0.028	0.182	0.307	-0.135	0.254	0.023	-0.008	0.251
	[0.662]	[0.133]	[0.846]	[0.621]	[-0.290]	[0.519]	[0.087]	[-0.036]	[0.973]
ABSNEWS(-2)	-0.154	0.038	-0.282	0.384	0.664	-0.228	0.151	0.290	0.222
	[-0.675]	[0.183]	[-1.317]	[0.775]	[1.427]	[-0.466]	[0.567]	[1.238]	[0.857]

Table 4.21 (Continued)

Dependent variable	HSIF detrended log open interest			KLCIF detrended log open interest			SiMSCIF detrended log open interest		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
ABSNEWS(-3)	0.221	0.125	0.145	-0.690	-0.562	-0.376	0.173	-0.140	-0.155
	[0.972]	[0.600]	[0.675]	[-1.391]	[-1.207]	[-0.769]	[0.649]	[-0.597]	[-0.597]
ABSNEWS(-4)	-0.258	-0.052	-0.416*	0.606	0.461	-0.049	0.195	0.356	-0.338
	[-1.133]	[-0.246]	[-1.938]	[1.224]	[0.990]	[-0.100]	[0.734]	[1.515]	[-1.296]
ABSNEWS(-5)	0.277	0.264	0.219	0.902*	0.675	0.229	-0.239	-0.098	-0.016
	[1.220]	[1.268]	[1.023]	[1.820]	[1.451]	[0.468]	[-0.904]	[-0.419]	[-0.061]
X2(1)Joint [NEWS]	3.318	5.258	3.530	5.179	2.070	4.276	7.247	8.664	5.088
p-value	0.651	0.385	0.619	0.394	0.839	0.510	0.203	0.123	0.405
X2(1)Joint [ABSNEWS]	4.271	2.093	7.298	7.610	6.754	1.261	2.150	4.142	3.725
p-value	0.511	0.836	0.199	0.179	0.240	0.939	0.828	0.529	0.590
Panel B Sub Sample 2									
NEWS	-0.207	-0.355***	-0.241*	0.434*	0.464**	0.427*	0.157	-0.014	0.147
	[-1.361]	[-2.705]	[-1.797]	[1.817]	[2.045]	[1.843]	[1.047]	[-0.095]	[1.038]
NEWS(-1)	-0.156	-0.185	-0.159	-0.070	-0.181	-0.043	-0.319**	-0.235	-0.270**
	[-1.008]	[-1.390]	[-1.174]	[-0.291]	[-0.792]	[-0.186]	[-2.082]	[-1.579]	[-1.873]
NEWS(-2)	0.183	0.158	-0.088	-0.128	-0.110	-0.227	0.127	0.079	0.102
	[1.180]	[1.177]	[-0.649]	[-0.529]	[-0.481]	[-0.972]	[0.829]	[0.531]	[0.704]
NEWS(-3)	-0.208	-0.236*	-0.279**	-0.232	-0.071	-0.405*	0.018	-0.031	0.050
	[-1.348]	[-1.761]	[-2.065]	[-0.964]	[-0.310]	[-1.744]	[0.120]	[-0.208]	[0.346]
NEWS(-4)	-0.017	-0.100	0.011	-0.174	0.042	0.031	0.026	0.034	-0.082
	[-0.111]	[-0.748]	[0.079]	[-0.726]	[0.185]	[0.134]	[0.171]	[0.233]	[-0.568]
NEWS(-5)	0.159	-0.043	0.060	-0.120	-0.299	0.033	0.021	0.062	0.082
	[1.090]	[-0.337]	[0.463]	[-0.534]	[-1.392]	[0.154]	[0.150]	[0.434]	[0.607]
ABSNEWS	-136.475*	0.152	0.055	-0.568	0.188	-0.285	-0.291	-0.392*	-0.109
	[-1.707]	[0.773]	[0.265]	[-1.559]	[0.594]	[-0.851]	[-1.309]	[-1.805]	[-0.517]
ABSNEWS(-1)	-0.146	-0.067	-0.024	0.097	-0.263	-0.045	0.046	0.142	0.284
	[-0.619]	[-0.336]	[-0.117]	[0.261]	[-0.823]	[-0.133]	[0.203]	[0.640]	[1.319]
ABSNEWS(-2)	-0.263	-0.058	0.029	-0.117	-0.088	0.011	0.134	-0.038	0.136
	[-1.111]	[-0.289]	[0.138]	[-0.315]	[-0.277]	[0.032]	[0.587]	[-0.170]	[0.635]

Table 4.21 (Continued)

Dependent variable	HSIF detrended log open interest			KLCIF detrended log open interest			SiMSCIF detrended log open interest		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
ABSNEWS(-3)	-0.410*	-0.114	-0.195	-0.099	-0.224	-0.301	0.053	0.040	0.165
	[-1.733]	[-0.571]	[-0.942]	[-0.268]	[-0.704]	[-0.888]	[0.232]	[0.180]	[0.770]
ABSNEWS(-4)	0.097	0.181	0.197	-0.708*	-0.559*	-0.166	-0.291	-0.178	-0.241
	[0.411]	[0.905]	[0.948]	[-1.914]	[-1.757]	[-0.490]	[-1.281]	[-0.800]	[-1.123]
ABSNEWS(-5)	-0.147	0.199	-0.140	-0.227	-0.085	-0.587*	-0.020	0.063	-0.220
	[-0.621]	[1.004]	[-0.676]	[-0.614]	[-0.269]	[-1.729]	[-0.087]	[0.288]	[-1.026]
X^2(1)Join [NEWS]	5.188	7.940	7.372	2.826	3.667	4.544	4.748	2.792	4.289
p-value	0.393	0.160	0.194	0.727	0.598	0.474	0.447	0.732	0.509
X^2(1)Join [ABSNEWS]	5.146	2.294	2.139	4.290	4.285	3.970	2.136	1.075	5.035
p-value	0.398	0.807	0.830	0.508	0.509	0.554	0.830	0.956	0.412

Table 4.22 Predicting Contract Volume Using Bad News Factor

This table presents the coefficients β_{4i} and θ_{4i} estimated for equation (13) using structural vector autoregressive (SVAR) regression model. The trading activity data collected from Datastream. The bad news factors (*Pessimism, Negative and Weak*) are generated by the General Inquirer program and Principal Component Analysis. The endogenous variables are close-to-close returns, bad news factors and trading volume (in open interest).

$$V_t = \alpha_4 + \sum_{i=0}^{5} \beta_{4i} N_{t-i} + \sum_{i=0}^{5} \theta_{4i} |N_{t-i}| + \sum_{i=0}^{5} \gamma_{4i} R_{t-i} + \sum_{i=1}^{5} \delta_{4i} V_{t-i} + \sum_{j=1}^{14} \lambda_{4j} EXOG_j + \varepsilon_{3t} \qquad (3.7)$$

I expect the β_{4i} to be negative and the θ_{4i} to be positive. The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The coefficients denote the impact of one-standard deviation change in bad news factor and change in absolute bad news factors respectively on the percentage change of contract volume of HSIF, KLCIF and SiMSCIF. The $X^2(1)$ is the chi-square value for the joint coefficient Wald-test, with 1 degree of freedom. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Dependent variable	HSIF detrended log contract volume			KLCIF detrended log contract volume			SiMSCIF detrended log contract volume		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
Panel A Sub Sample 1									
NEWS	0.866	0.622	0.987	-0.743	0.017	-1.614	-1.945	-1.345	-1.545
	[0.834]	[0.525]	[1.010]	[-0.646]	[0.015]	[-1.415]	[-1.533]	[-1.192]	[-1.246]
NEWS(-1)	-0.104	-0.945	0.204	1.197	2.038*	0.545	2.226*	1.291	1.090
	[-0.099]	[-0.967]	[0.207]	[1.024]	[1.809]	[0.469]	[1.730]	[1.133]	[0.866]
NEWS(-2)	0.703	0.454	-0.311	-1.124	-1.848	-0.433	-0.176	-0.385	-0.295
	[0.670]	[0.463]	[-0.315]	[-0.954]	[-1.620]	[-0.371]	[-0.137]	[-0.338]	[-0.234]
NEWS(-3)	-0.463	-0.766	0.128	-0.988	-0.524	0.097	-1.533	-1.735	-0.193
	[-0.441]	[-0.777]	[0.129]	[-0.843]	[-0.459]	[0.083]	[-1.192]	[-1.526]	[-0.154]
NEWS(-4)	2.011*	1.106	1.567	-0.463	0.348	0.400	0.941	0.566	1.014
	[1.947]	[1.144]	[1.608]	[-0.396]	[0.306]	[0.346]	[0.734]	[0.499]	[0.815]
NEWS(-5)	-0.306	-0.014	-0.520	1.267	0.914	1.106	0.482	0.706	-0.359
	[-0.311]	[-0.015]	[-0.553]	[1.163]	[0.860]	[1.030]	[0.414]	[0.657]	[-0.311]
ABSNEWS	-1.530	-0.299	-1.745	-1.657	0.140	-0.343	5.779***	3.068*	5.228***
	[-0.974]	[0.835]	[-1.181]	[-0.989]	[0.089]	[-0.207]	[3.061]	[1.840]	[2.877]
ABSNEWS(-1)	-0.450	-2.189	0.988	-2.894*	-3.758**	-2.355	-2.671	-0.335	0.210
	[-0.284]	[-1.507]	[0.660]	[-1.704]	[-2.351]	[-1.406]	[-1.391]	[-0.198]	[0.113]
ABSNEWS(-2)	-1.802	-0.806	-2.229	2.017	-0.396	0.875	-1.193	-0.757	0.852
	[-1.139]	[-0.554]	[-1.493]	[1.184]	[-0.247]	[0.522]	[-0.625]	[-0.449]	[0.458]

Table 4.22 (Continued)

Dependent variable	HSIF detrended log contract volume			KLCIF detrended log contract volume			SiMSCIF detrended log contract volume		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
ABSNEWS(-3)	-0.312	0.753	-1.765	-0.609	-0.764	0.652	-0.481	0.607	-2.157
	[-0.197]	[0.517]	[-1.181]	[-0.357]	[-0.478]	[0.389]	[-0.252]	[0.360]	[-1.160]
ABSNEWS(-4)	-4.424***	-2.839*	-2.146	2.386	1.820	1.091	-0.619	0.625	-3.250*
	[-2.793]	[-1.944]	[-1.436]	[1.402]	[1.142]	[0.650]	[-0.324]	[0.370]	[-1.741]
ABSNEWS(-5)	-3.167**	-1.238	-3.345**	0.894	2.204	0.912	1.419	1.204	3.763**
	[-1.998]	[-0.851]	[-2.243]	[0.525]	[1.382]	[0.544]	[0.749]	[0.716]	[2.032]
X^2(1)Join [NEWS]	4.564	2.889	2.857	3.745	6.287	1.809	4.999	4.230	1.561
p-value	0.471	0.717	0.722	0.587	0.279	0.875	0.416	0.517	0.906
X^2(1)Join [ABSNEWS]	13.188**	7.550	11.597**	7.033	8.687	3.075	2.996	1.049	8.372
p-value	0.022	0.183	0.041	0.218	0.122	0.689	0.701	0.959	0.137
<u>Panel B Sub Sample 2</u>									
NEWS	-0.802	-1.181*	-0.125	0.644	0.417	0.782	1.786	0.258	0.743
	[-0.972]	[-1.656]	[-0.172]	[0.570]	[0.390]	[0.714]	[1.155]	[0.170]	[0.505]
NEWS(-1)	0.172	0.669	-0.109	-0.725	-1.549	-1.340	-0.210	-0.875	0.945
	[0.204]	[0.927]	[-0.147]	[-0.635]	[-1.438]	[-1.215]	[-0.133]	[-0.572]	[0.633]
NEWS(-2)	-0.113	-0.639	-0.526	-0.040	-0.163	-0.380	-1.130	-2.122	-0.572
	[-0.134]	[-0.876]	[-0.712]	[-0.035]	[-0.151]	[-0.346]	[-0.715]	[-1.385]	[-0.382]
NEWS(-3)	-0.833	-0.626	-0.291	0.240	0.784	-0.905	-0.049	1.662	0.333
	[-0.991]	[-0.856]	[-0.396]	[0.211]	[0.728]	[-0.826]	[-0.031]	[1.088]	[0.221]
NEWS(-4)	-0.579	-0.553	-0.268	-0.115	0.484	-0.609	-0.304	-0.346	-0.946
	[-0.691]	[-0.761]	[-0.364]	[-0.101]	[0.452]	[-0.557]	[-0.194]	[-0.227]	[-0.633]
NEWS(-5)	0.403	0.157	0.471	0.056	-0.603	-0.291	1.294	1.446	1.505
	[0.507]	[0.225]	[0.667]	[0.053]	[-0.593]	[-0.283]	[0.883]	[0.990]	[1.072]
ABSNEWS	-97.534	0.726	0.618	-0.060	1.206	1.289	-2.931	-3.129	50.457
	[-0.303]	[0.682]	[0.554]	[-0.035]	[0.807]	[0.813]	[-1.279]	[-1.400]	[-0.021]
ABSNEWS(-1)	-0.784	-1.296	-0.508	-2.545	-3.263**	0.720	-2.610	-3.128	-1.128
	[-0.610]	[-1.201]	[-0.450]	[-1.455]	[-2.158]	[-0.039]	[-1.114]	[-1.369]	[-0.507]
ABSNEWS(-2)	-0.895	0.258	-1.141	0.056	-1.122	0.616	1.564	0.616	0.971
	[-0.697]	[0.238]	[-1.008]	[0.032]	[-0.743]	[0.384]	[0.664]	[0.268]	[0.437]

Table 4.22 (Continued)

Dependent variable	HSIF detrended log contract volume			KLCIF detrended log contract volume			SiMSCIF detrended log contract volume		
Source of news	South China Morning Post			New Straits Times			The Straits Times		
News sentiment	Pessimism	Negative	Weak	Pessimism	Negative	Weak	Pessimism	Negative	Weak
ABSNEWS(-3)	-0.956	-0.891	-0.740	-0.613	-0.659	-2.035	0.684	1.288	2.416
	[-0.744]	[-0.821]	[-0.656]	[-0.350]	[-0.437]	[-1.269]	[0.290]	[0.562]	[1.089]
ABSNEWS(-4)	0.988	0.610	0.354	-1.657	0.676	-0.489	-4.990**	-3.002	-3.172
	[0.769]	[0.563]	[0.313]	[-0.946]	[0.449]	[-0.304]	[-2.124]	[-1.307]	[-1.428]
ABSNEWS(-5)	-2.276*	-0.502	-1.102	2.065	0.486	0.388	-0.288	2.156	-1.648
	[-1.777]	[-0.466]	[-0.982]	[1.179]	[0.324]	[0.241]	[-0.123]	[0.955]	[-0.744]
X^2(1)Join [NEWS]	1.756	3.133	1.264	0.457	2.949	3.803	1.238	4.059	2.073
p-value	0.882	0.679	0.939	0.994	0.708	0.578	0.941	0.541	0.839
X^2(1)Join [ABSNEWS]	5.190	2.804	2.823	4.470	5.761	1.922	6.241	4.499	4.145
p-value	0.393	0.730	0.727	0.484	0.330	0.860	0.284	0.480	0.529

4.5.6 Economic significance of the findings

The earlier section of this chapter discusses the predictability of the index futures returns. Bad news factors predict lower price on the same day and the next day for the HSIF and SiMSCIF. The impact of bad news on KLCIF exhibits some momentum. The bad news predicts lower price for the same day and up to the next two days. The prices start to reverse over the next few days. Approximately, HSIF demonstrates reversal from the next two to five days; KLCIF's returns reverse on the fifth day and SiMSCIF's returns reverse on the fourth day after the news was published. The exact timing of the reversal varies with bad news factor and sample period. This could be due to the proportion on informed traders and uninformed traders being different among these three markets. Theoretically, the more informed traders in the market, the faster the mispricing is adjusted. This section discusses the possible trading strategies based on these findings.

I assume that the market participants can borrow at the risk free rate, facing no restriction of access, and I ignore the trading margin. This zero-cost trading strategy is far from the real practice of the finance world. However, the simulation results provide some insight on the feasibility of the trading strategy. I devise the trading strategies conditional upon the level of the bad news factor, whether it is high or low. There are two feasible market-timing strategies based on high or low bad news factors. Appendix B provides the examples of these strategies, considering the price of long, price of short, and price of maturity. The examples show that the overall profit is the same for the two strategies regardless the price on maturity date. It also shows total loss if the index futures price goes against the prediction based on the news factors.

The first strategy is based on high levels of bad news factors. High levels of bad news factors predict low or negative returns; the returns are expected to reverse in the near future.

This study considers the following: go long on day ***t*** of which the ranking of the bad news factors fall in the top three deciles and go short on the ***t + n*^th^** day (***n*** varies with the timing of the reversal) after the news is published, and wait for the settlement at contracts' maturity. There is another equivalent strategy: buy a long contract on day ***t*** and sell the long contract on the $\boldsymbol{t + n^{th}}$ day as the returns start to reverse. The former strategy lets the contracts stay open for one month at maximum, because the traders are assumed to roll the spot month contracts to the next month contract as the spot month contracts mature. The latter strategy lets the positions stay open for only ***n*** days. Both strategies lock the profit on the ***n***th day. However, the latter is more preferable because the investor only bears the cost of capital for ***n*** days, instead of up to one month.

The next strategy is based on low levels of bad news factors. High market returns are associated with low levels of bad news factors, followed by low or negative returns after that. Firstly, one can borrow at the risk free rate; buying a short contract on day *t* for which the ranking of the bad news factors fall in the bottom three deciles. Secondly, buying a long contract on the $t+n^{th}$ day, and then wait until the contracts mature. Alternatively, buying a short contract on day one and selling it on day *n* will bring the same result.

I apply the two strategies on the HSIF and KLCIF. First, if the bad news factor is in the prior year's top three deciles, buy a long contract on the same day *(t=0)* and sell the long contract on the $t + n^{th}$ day as the returns start to reverse. Secondly, if the bad news factors is in the prior year's bottom three deciles, buy a short contract the same day *(t=0)* and sell it on day $t + n^{th}$. I let the contract stay open for one to four days (i.e. n=1, 2, 3, and 4).

Table 4.13 and Table 4.14 shows that the HSIF and SiMSCIF returns have a contemporaneous relationship with the bad news factor, and the reversals commence from the second day to the fifth depending on sample periods and measures of news sentiment. Therefore, the returns are calculated for the transactions that are initiated on the same day as the news arrival, and let the position stay open for one to four days before locking in the profit or loss by taking the opposite position.

The strategies for the KLCIF contract will be slightly different. The news factors generated from the News Straits Times predict that the negative impact of bad news significantly persists for two days, but the negative impact lasts for up to four days disregarding statistical significance. Let day ***t*** indicates the arrival of bad news, the long contract will be initiated on day the ***t***+3, because the negative impact of the bad news on KLCIF returns last for two days. This enables the investors to buy with the lowest price after the price has fallen for two days. On the other hand, when the ranking of the bad news factors fall in the bottom third decile on day t, a short contract will be initiated on day ***t***+3. The timing leads to the highest possible shorting price. In the simulation, the contract is allowed to stay open from one to four days, i.e. close the position on ***(t***+3) +n^{th} day , where n=1,2,3 or 4.

The daily settlement prices are used to calculate returns[30]. The returns for a long HSIF and SiMSCIF contracts are defined as:

$$Return = \log\left(\frac{Settlement\ Price_{t+n}}{Settlement\ Price_{t}}\right) \tag{4.1}$$

[30] The settlement price is used because the news released in morning and the index futures prices take time to react to the bad or good news. The HSIF daily settlement price is the best matched bid and ask prices during the last two minutes of the trading session, or the last traded price if there is no matching bid and ask order during the last two minutes, or will be determined by the HKEX if former two conditions cannot be met. Hence the settlement price should reflect the investors' valuation of the index futures at the end of the day. Unless there is a very important news release during the last two minutes of the trading session and able reach the public before the market close, the closed price and settlement price will be very close. Thus it is reasonable to assume that the investors able to buy and sell at settlement price.

The returns for a long KLCIF contract is defined as:

$$Return = \log\left(\frac{Settlement\ Price_{t+3+n}}{Settlement\ Price_{t+3}}\right) \tag{4.2}$$

Another important factor that determines the feasibility of these strategies is market depth. If the index futures markets are highly liquid and investors can buy and sell without any delay, then the investors are more likely to make profit from implementing these strategies. The summary of open interest and the number of contracts traded daily is illustrated in Table 4.7 reflecting the contracts' liquidity. The numbers of contracts traded daily are 88618, 17500 and 11601 respectively for the HSIF, SiMSCI and KLCIF in 2008. I assume there are no limits to arbitrage, at least from the perspective of market liquidity.

Table 4.23 summarises the number of long and short contracts initiated as signalled by the bad news factors.

Table 4.23 The Number of Long and Short Contract Initiated as Signalled by the Bad News Factors
This table reports the number of transaction initiated based on the ranking of the Pessimism (Psm), Negative (Ngv) and Weak (Weak) bad news factors. The full sample period for HSIF and KLCIF cover 1/1996 to 12/2008, consist of approximately 3144 and 3102 observations. The SiMSCIF data range 1/1999 through 12/2008, composed of approximately 2500 observations. On the day that the bad news factors fall in the top third deciles, a long contract will be initiated. On the day that the bad news factors rank in the bottom third deciles, a short contract will be initiated.

	HSIF			KLCIF			SiMSCIF		
Decile	Psm	Ngv	Weak	Psm	Ngv	Weak	Psm	Ngv	Weak
top1	318	318	294	228	240	233	238	236	237
top2	634	637	701	467	477	469	473	470	474
top3	1077	944	1002	706	726	711	781	790	744
bottom3	948	945	972	705	700	701	713	711	711
bottom2	631	612	659	471	464	463	472	477	478
bottom1	322	333	350	221	281	271	239	243	239

Table 4.24 shows the Hang Seng Index Futures (HSIF) returns from implementing the hypothetical zero-cost trading strategy that considers the length of holding period (one to four days), the type of bad news factors (*Pessimism, Negative* and *Weak*) and the ranking of the bad news factors (top three deciles and bottom three deciles) in the return generating process. Each of the columns illustrates the impact of holding period on returns. The returns for holding the positions for three to four days outperform the one-day and two-day positions. The rows compare the performance of the three measures of news factors, and the results are mixed. *Pessimism* outperforms *Negative* and *Weak* for long positions (when the news measures are in the top third decile). However, *Negative* and *Weak* outperform *Pessimism* for short positions (when the news factors are in the bottom third decile). By treating every three rows as a block, the result shows that the transactions initiated when the bad news factor lies in the top second decile generate the highest return, ranging from 26.58 to 61.27 basis points. The returns for the going long strategy when the level of bad news factors is in top third decile is higher than going short when the level of bad news factors is in bottom third decile. Overall, the simulation shows closing the contracts four days after going long on the day when the bad news level is in the top second decile, brings optimal returns. This is consistent with the findings that the reversals occur in the next two to four days after the news is released as show in Table 4.13 and Table 4.14.

Assuming that the Hang Seng index is now at 12500 points, the round trip trading cost is \$200 as the bid-ask spread is approximately 2 to 5 index points, resulting in \$100 to \$250. The total transaction cost amounts to 4.8 to 7.2 basis points[31]. As the returns from implementing the trading strategies outweigh the cost, the assertion is that the bad news factors can be used as a market timing tool, and the returns are economically significant.

[31] This is based on news lease by the HKEX, dated 16th August, 2001. The round trip trading cost includes brokerage commissions, exchange trading fees and stamp duty. http://www.hkex.com.hk/eng/newsconsul/hkexnews/2001/010816news.htm, Accessed on 26th April 2010.

Table 4.24 Returns Generated from Trading Hang Seng Index Futures Contracts

Returns generated from trading Hang Seng Index Futures contract based on trading strategy derived from bad news factors, for the period of 1996 to 2008. The calculations are based on three bad news factor: pessimism (PSM), negative (NGV) and Weak (WEAK). The investors initiate the transaction based on the ranking of the bad news factor and let the contract open for one to four days. Investors will go long when the bad news factor is in the top three deciles or go short when the bad news factors are in the bottom three deciles. The returns are in basis points. The t values are reported in parentheses [ ] *, ** and *** denotes the average returns is significantly different from 0 at 10%, 5% and 1% respectively.

News Ranking (Decile)	News	Returns			
		Holding Period			
		1	2	3	4
Top 1	PSM	7.43	7.91	40.08***	21.03
		[0.491]	[0.542]	[2.679]	[1.435]
	NGV	10.59	9.93	46.66***	11.25
		[0.729]	[0.686]	[3.095]	[0.826]
	WEAK	12.79	21.73*	26.91**	14.87
		[1.011]	[1.698]	[2.386]	[1.128]
Top 2	PSM	9.71	26.58**	48.17***	61.27***
		[1.021]	[2.048]	[3.176]	[3.370]
	NGV	6.30	16.64	39.06***	47.10***
		[0.646]	[1.257]	[2.626]	[2.700]
	WEAK	2.97	14.96	32.22**	40.54**
		[0.355]	[1.293]	[2.295]	[2.338]
Top 3	PSM	11.71*	22.96**	35.00***	47.87***
		[1.688]	[2.380]	[3.214]	[3.659]
	NGV	3.81	13.53	20.60**	26.92**
		[0.509]	[1.300]	[1.725]	[1.930]
	WEAK	6.61	14.91	25.79**	30.15**
		[0.943]	[1.550]	[2.255]	[2.188]
Bottom 3	PSM	2.47	5.58	15.81*	15.55
		[0.392]	[0.699]	[1.665]	[1.437]
	NGV	2.49	8.03	16.52*	21.27*
		[0.414]	[1.036]	[1.728]	[1.920]
	WEAK	0.49	5.97	19.93**	23.20**
		[0.078]	[0.724]	[2.088]	[2.166]
Bottom 2	PSM	-4.53	0.60	9.07	8.88
		[-0.622]	[0.065]	[0.810]	[0.679]
	NGV	-0.80	2.37	13.76	18.71
		[-0.111]	[0.257]	[1.220]	[1.415]
	WEAK	-7.71	-4.44	12.94	16.77
		[-0.990]	[-0.446]	[1.129]	[1.294]
Bottom 1	PSM	4.74	10.76	16.19	16.08
		[0.473]	[0.863]	[1.099]	[0.900]
	NGV	2.28	11.11	22.68	20.59
		[0.210]	[0.817]	[1.380]	[1.024]
	WEAK	-5.84	-1.35	18.00	21.74
		[-0.550]	[-0.104]	[1.186]	[1.224]

By examining the columns in Table 4.25, the returns generated from the Kuala Lumpur Composite Index Futures (KLCIF) following the suggested strategy are somewhat different from the findings for HSIF contracts. The trades initiated based on the days on which the bad news factors fall in bottom deciles bring better returns than the trades generated based on the days on which the bad news factors fall in top deciles. Specifically, the strategy of going short on day t+3 after the bad news factors falls in the bottom 1 decile generates the best returns. The returns are the highest, 44.80 basis points, based on Negative news factors and holding the position up to four days. The strategy brings positive returns even letting the positions stay open for two to four days. This reflects the fact that the impact of bad news decays gradually over several days, which is common for a developing market like Malaysia[32]. In this case, a shorter holding period is preferable over a longer one because investors face lower cost of capital and less uncertainty. The strategy based on *Negative* and *Weak* news factors outperform the one based on *Pessimism* news factor, as far as the bottom three deciles are concerned. The results support the conclusion that bad news factors lead price falls for three days, and then reverse on the fourth day and the fifth day (see Table 4.13 and Table 4.14).

The cost to initiate a RM100 000 position in a KLCIF contract is approximately 16 basis points[33] (Hassan et al., 2007). In conclusion, the bad news factors driven strategies are able to generate profits.

[32] Matteo, Aste and Dacorogna (2005) conclude that long memory is an efficient way to distinguish the degree of market development.

[33] This includes brokerage commission in futures, exchange trading fee, clearing house fee, bid ask spread.

Table 4.25 Returns Generated from Trading Kuala Lumpur Composite Index Futures Contracts
Returns generated from trading Kuala Lumpur Composite Index Futures contract based on trading strategy derived from bad news factors, for the sample period of 1996 to 2008. The calculations are based on three bad news factor: pessimism (PSM), negative (NGV) and weak (WEAK). The investors initiate the transaction based on the ranking of the bad news factor and let the contract open for one to four days. Investors will go long three days after the bad news factors fall in the top three deciles or go short on the third day following the days when bad news factors are in the bottom three deciles, due to bad news show momentum of its negative impact on KLCIF returns. The returns are in basis points. The t values are reported in parentheses [ ] *, ** and *** denotes the average returns is significantly different from 0 at 10%, 5% and 1% respectively.

		Returns			
		Holding periods			
News Ranking (decile	News	1	2	3	4
Top 1	PSM	18.68	20.33	11.21	3.44
		[1.488]	[1.064]	[0.573]	[0.160]
	NGV	-1.94	-13.00	-12.80	5.82
		[-0.164]	[-0.778]	[-0.710]	[0.260]
	WEAK	25.30*	17.40	11.31	2.49
		[1.916]	[0.951]	[0.615]	[0.126]
Top 2	PSM	2.23	-5.29	-2.26	-7.26
		[0.235]	[-0.425]	[-0.171]	[-0.491]
	NGV	5.45	3.14	2.77	0.89
		[0.615]	[0.263]	[0.214]	[0.060]
	WEAK	6.75	-1.23	0.01	-5.79
		[0.715]	[-0.097]	[0.001]	[-0.413]
Top 3	PSM	-4.86	-11.13	-10.83	-1.34
		[-0.674]	[-1.150]	[-1.018]	[-0.117]
	NGV	0.93	-5.80	-6.14	-2.81
		[0.125]	[-0.601]	[-0.585]	[-0.252]
	WEAK	2.80	-7.17	-8.59	-9.49
		[0.378]	[-0.721]	[-0.809]	[-0.843]
Bottom 3	PSM	-5.22	1.97	2.18	11.21
		[-0.694]	[0.236]	[0.206]	[1.008]
	NGV	0.49	8.86	11.17	12.19
		[0.064]	[1.072]	[1.059]	[1.094]
	WEAK	1.22	16.47**	17.48	17.95
		[0.164]	[1.962]	[1.620]	[1.636]
Bottom 2	PSM	1.43	12.14	-9.18	8.00
		[0.148]	[1.168]	[-0.725]	[0.585]
	NGV	0.01	12.51	-9.93	7.16
		[0.001]	[1.172]	[-0.765]	[0.499]
	WEAK	11.15	22.96**	-20.95	8.79
		[1.467]	[2.131]	[-1.553]	[0.733]
Bottom 1	PSM	-3.36	23.44	3.02	5.49
		[-0.201]	[1.376]	[0.148]	[0.240]
	NGV	-1.77	23.84*	32.26*	44.80**
		[-0.132]	[1.663]	[1.889]	[2.323]
		18.19*	34.36**	40.19**	25.82
	WEAK	[1.825]	[2.366]	[2.156]	[1.568]

Table 4.26 displays the simulation results for implementing the strategies on the Singapore Morgan Stanley Free Index Futures (SiMSCIF) contracts. The returns generated from implementing the suggested trading strategy are not significantly different from zero. This is due to the magnitude of the initial negative impact of bad news on the SiMSCIF ($\beta_{20} = -8.97, -4.484\ and - 7.960$) are relatively smaller as compared to the HSIF ($\beta_{20} = -10.680, -13.160\ and - 11.300$) and KLCIF ($\beta_{20} = -16.68, -15.27, and - 15.500$). However, the patterns of the findings are consistent with the simulation with HSIF contracts. The trades that are initiated based on the top three deciles of the bad news factors result in higher returns as compared to the strategy based on bottom deciles. Closing the positions after four days is more profitable as compared to holding the positions for one, two or three days. This is consistent with the findings as shown in Table 4.13 and Table 4.14, that the reversals occur after four days of the news release. The strategy that uses *Pessimism* as a benchmark to go long out-performs the one uses *Negative* or *Weak* words. The strategy performs the best when the bad news factors fall in the top two deciles and the positions are left open for four days. However, the strategy generates 17.26 basis points at best. This is consistent with earlier findings that there is no significant reversal during sub-period 1. In addition, the reversals only occur on the next four days for sub-sample 2 with a small magnitude (see Table 4.19, $\beta_{24} = 8.780\ and\ 7.960$)

Taken into account the bid-ask spread of 5.2 basis point[34], clearing fees of 4 basis points of contract value, and access trading fees of 0.75 basis points[35] of contract value, the strategy yields positive returns.

[34]Frino, A., Kruk, J., & Lepone, A (2007). *The SPI 200 in the Asia-Pacific region: Comparisons of liquidity and Transactions Costs against other stock index futures* (Edition 12), Sydney, Australia: Australia Securities Exchange.
[35] http://www.world-exchanges.org/exchanges/singapore-exchange

Table 4.26 Returns Generated from Trading Singapore Morgan Stanley Composite Index Futures Contracts

Returns generated from trading Singapore Morgan Stanley Composite Index Futures contract based on trading strategy derived from bad news factors, for the period of 1999 to 2008. The calculations are based on three bad news factor: pessimism (PSM), negative (NGV) and weak(WEAK). The investors initiate the transaction based on the ranking of the bad news factor and let the contract open for one to four days. Investors will go long when the bad news factor is in the top three deciles or go short when the bad news factors are in the bottom three deciles. The returns are in basis points. The t values are reported in parentheses [ ] *, ** and *** denotes the average returns is significantly different from 0 at 10%, 5% and 1% respectively.

News Ranking (decile)	News	Returns			
		Holding periods			
		1	2	3	4
Top 1	PSM	10.67	-1.51	-4.54	8.73
		[0.920]	[-0.130]	[-0.381]	[0.737]
	NGV	-2.33	9.72	-7.91	14.82
		[-0.198]	[0.824]	[-0.631]	[1.200]
	WEAK	-7.20	-6.74	3.62	9.96
		[-0.606]	[-0.581]	[0.316]	[0.866]
Top 2	PSM	7.83	15.54	12.16	17.26
		[0.968]	[1.491]	[0.938]	[1.186]
	NGV	4.13	7.51	2.36	11.12
		[0.512]	[0.699]	[0.175]	[0.741]
	WEAK	-1.05	0.19	2.02	6.54
		[-0.132]	[0.018]	[0.155]	[0.453]
Top 3	PSM	5.19	6.04	3.41	5.70
		[0.878]	[0.759]	[0.345]	[0.505]
	NGV	-0.90	4.23	0.17	0.13
		[-0.152]	[0.534]	[0.018]	[0.012]
	WEAK	-1.11	-0.79	-0.96	5.34
		[-0.184]	[-0.099]	[-0.095]	[0.469]
Bottom 3	PSM	-1.69	-3.02	-0.90	3.91
		[-0.305]	[-0.394]	[-0.097]	[0.366]
	NGV	-7.22	-7.27	-6.55	-2.87
		[-1.229]	[-0.921]	[-0.682]	[-0.264]
	WEAK	1.11	0.68	-3.91	1.27
		[0.197]	[0.090]	[-0.434]	[0.125]
Bottom 2	PSM	-0.15	-4.27	-9.86	-2.43
		[-0.026]	[-0.565]	[-1.090]	[-0.226]
	NGV	-6.04	-7.14	-5.76	-4.37
		[-0.852]	[-0.774]	[-0.499]	[-0.326]
	WEAK	3.74	-2.17	-7.54	2.93
		[0.529]	[-0.240]	[-0.683]	[0.226]
Bottom 1	PSM	0.96	-6.94	-10.43	-10.30
		[0.099]	[-0.604]	[-0.720]	[-0.599]
	NGV	0.16	-7.08	-9.83	-14.41
		[0.016]	[-0.543]	[-0.612]	[-0.776]
	WEAK	4.69	-18.68	-27.50*	-17.07
		[0.495]	[-1.448]	[-1.763]	[-0.940]

Table 4.27 The Effect of Holding Period, Ranking, Year and News Factors on Trading Strategy Returns
The univariate general linear model (GLM) is used to run the analysis of variance (ANOVA), examine the between subject effect (holding period, ranking, year and news factors) on the returns from implementing the proposed trading strategies. The columns show the F-test statistics for HSIF, KLCIF and SiMSCIF. The sample period for HSIF and KLCIF range 1996 to 2008. The sample period for SiMSCIF span from 1999 to 2008. The Year (1996-2008) and News (Pessimism, Negative, Weak) are fixed factor in the GLM. The holding period (one day to four days) and ranking (top 1, 2, and 3 deciles; bottom 3, 2 and 1 deciles) are covariates in the GLM. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

	Index futures contract		
Source of variance	HSIF	KLCIF	SiMSCIF
Corrected Model	9.611***	5.158***	8.487***
Intercept	3.642*	0.530	2.443
Holding period	47.503***	0.390	0.142
Ranking	26.669***	10.969***	9.200***
Year	23.065***	13.185***	23.613***
News	0.646	2.202	1.207
Year * News	1.316	1.444*	2.053***

Table 4.27 displays the *between subject effect* that determines the effectiveness of the trading strategy devised from the bad news factors. The study assumes that the returns will vary with the holding period and the ranking of the bad news factors. Because the financial time series are cyclical in nature, the factor of the time (i.e. year) is also examined. First, the holding period is significantly affecting the returns, only for the HSIF. Based on Table 4.24, the returns from holding the contracts for three to four days are higher than holding it for one to two days, while the results for KLCIF and SiMSCI are mixed. Second, the ranking of the bad news significantly affects the returns as expected for all three contracts, but the KLCIF displays a contradicting result when compared to the other two contracts. The returns for the HSIF and SiMSCI contracts are good when the trades are initiated when the bad news factors falls in the first or second deciles. The KLCIF returns are the best when the trades are initiated based on the bottom two and bottom three deciles. Third, the returns variability is significantly associated with the time factor, mainly the state of economy. Fourth, the three news factors do not significantly lead to variation in returns for all the three contracts. This suggests that the three bad news factors are indifferent indicators to initiate trades. On the other hand, the

Year*News interaction term is significant for the KLCIF and SiMSCIF contracts. This time variation in returns corroborates the theory of investors' beliefs and uncertainties in a two-state (good or bad, high-growth or low-growth) economy. The investors tend to overreact to news during bad-state because of greater uncertainty about the future leading to greater returns volatility (Veronesi, 1999). The theory of cognitive dissonance predicts that prices become more and more sensitive to bad news as a crisis exacerbates (Kaminsky & Schmukler, 1999). In contrast, the stock prices overreact to good news in the heydays of bubbles (Keijer and Prast, 2001 as cited in (Prast & de Vor, 2005)). Simon and Wiggins lll (2001) use the VIX, put-call ratio and trading index as proxies for market sentiment and conclude that these sentiment proxies are able to forecast S&P futures returns. Simulations with contrarian strategies recorded greater returns even after risk adjustment. The findings conform to the contrarian beliefs that in periods of low sentiment, assets are under-priced and the stock market will adjust to fair price again in the subsequent period. Most studies find these three sentiment measures as contrarian indicators for future price movements.

The returns generated from these hypothetical trading strategies conform to the earlier findings, with the caveat that the trading strategies are devised based on a few assumptions that make it not fully realistic.

4.6 CONCLUSION

This study begins with the proposal of possible theories to explain the association between index futures returns and bad news factors that are derived from routine newspapers' market summaries. The underlying information and investor sentiment theories lead to specific hypotheses and tests. First, the bad news proxy for unfavourable information that is already incorporated into prices. Second, the bad news proxy for negative information that is yet to be reflected into price. Third, bad news is a proxy for negative investor sentiment.

These three hypotheses imply different price behaviours and enable statistical examinations. In the first case, the price should adjust to its new equilibrium at once upon the publication of news; hence, the bad news should have no impact on the following days' index futures returns. For the second case, conservatism causes prices to react to information slowly. The bad news should have a negative impact on the index prices for a short horizon, and the impact should last permanently after that, thus no price reversal is expected. Finally, the investor sentiment theory assumes that the news contains non-information noise, predicts that bad news instills negative sentiment among investors, and thus have a negative impact on index futures returns in the short time horizon. However, the returns should reverse in the longer run, as the index futures value will adjust to its fundamental value again with the existence of arbitrage activities.

The findings are consistent with the sentiment theory and support the hypothesis that sentiment causes an initial reduction on returns and reverses later. Putting all the earlier findings together, the following conclusions are reached. First, the findings from HSIF and SiMSCIF show a similar pattern. The negative impact of bad news on returns only significantly last for one day, then the returns start to reverse two days after the news. The similarity might be attributed to the market characteristics. Both of these markets are categorised as developed market[36] by MSCI Barra and FTSE Group. In contrast, the impact of bad news on the KLCIF dies out slowly. Malaysia falls under the list of emerging markets in the MSCI Barra and FTSE Group's market classification. Second, the finding of economic significance is consistent with the findings of the impact of the bad news factor on the index futures returns. The initial negative impact of bad news on returns indicates the timing required to initiate a long or short position. The returns reversals indicate the timing required to close the position and lock in the profit.

---

[36] http://www.mscibarra.com/products/indices/tools/index_country_membership/ and http://www.ftse.com/Indices/Country_Classification/index.jsp Accessed on 12th April 2010.

I assess the validity of the bad news factor as a measure of sentiment from three aspects. First, the well-established and tested psychosocial dictionary is used to construct the bad news factors. The *Weak* and *Negative* news factor explain most of the variation contain in each news article in the samples. In addition, *Weak* and *Negative* words is combined to form a *Pessimism* news factor, with the intention to increase its explanatory power. The prediction based on these three news factors is consistent. Therefore, the bad news factors are a reliable measure of sentiment. Second, the sentiment measures are used in the SVAR estimation, and the prediction appears to be consistent among the three countries—two developed and one developing. The pattern remains similar when the samples are split in half. The slight differences can be ascribed to other market macroeconomics variables during the sample period that have impacted on returns volatility; the 1999 Asian financial crisis; the 2008 Wall Street meltdown; the January effect, the Chinese new year effect, the day-of-the-week effect and four days prior settlement are controlled for. Third, simulating returns with bad news factor driven strategies generates economically significant returns after explicit trading costs are taken into account. However, it is unclear that if implicit trading cost, for example, taxes and market depth are taken into consideration, the strategies still result in excess returns.

CHAPTER 5 : EMPIRICAL RESULTS- THE ROLE OF INVESTOR SENTIMENT ON THE INDEX FUTURES RETURNS MEAN-VARIANCE RELATIONSHIP

5.1 INTRODUCTION

The present empirical evidence on the mean-variance relationship focuses on aggregate stock markets and is mixed at best. It is argued that this is due to factors such as volatility model specification, sampling issues and conditioning variables included in the mean-variance model.

Alternatively, this study considers the role of news sentiment in determining the mean-variance relationship in the context of the Hang Seng Index Futures (HSIF), Kuala Lumpur Composite Index Futures (KLCIF) and the Singapore Morgan Stanley Free Index Futures (SiMSCIF). I find bi-directional granger causality among excess returns, volatility and news sentiment regimes. These results are consistent with the noise trader theory. During periods of high sentiment, rational arbitrageurs are reluctant to trade, thus the impact of noise traders is prevalent. Although the mean-variance relationship is not always positive, the slope coefficients of mean-variance regressions in high sentiment regimes are significantly more negative after the interaction between news sentiment and volatility is taken into account. I conclude that news sentiment attenuates the mean-variance relationship during high news sentiment periods. This confirms the risk-seeking behaviour of noise traders.

Notwithstanding, the results during low news sentiment periods are ambiguous. This could be due to the influence of the financial crisis in 2008, or asymmetry volatility.

5.2 SUMMARY STATISTICS AND PRELIMINARY TESTS

This chapter aims to examine the mean-variance relationship during different news sentiment regimes. The daily news is recoded into *Good, Bad, Newhigh* and *NewLow*, based mainly on the headlines and then by the contents. The *Newhigh* is a subset of *Good*, while *NewLow* is a subset of *Bad*. Since the news collected from South China Morning Post often quotes the term "higher benchmark" and "lower benchmark" instead of "new high" or "new low", I follow their writing style and code it as *Highbench* and *Lowbench*. The numbers of trading days are 3391, 3101 and 2359 days respectively for Hang Seng Index Futures (HSIF), KLCIF and SiMSCIF. I collect news for each of the trading days.

One of the problems of the human coding method is committing inconsistencies of coding between coders or over time (see section 3.3.3.1). I perform human-coding on the same news at two different time periods, to examine the consistencies of the resulted codes. The first coding is performed from 4th August 2010 to 19th August 2010. The Second recode is performed from 29th November 2010 to 7th December 2010. Table 5.1 provides the total count for each of the news sentiment regimes. The coding is highly consistent when the markets reach new high/low, or reach a higher/lower benchmark. The discrepancies however, mainly arise from the categories of *Good*, *Bad* and *Neutral*, when there is a mixture of good and bad news for various stock sectors in a news article for a particular day. I perform further analysis to test the hypotheses based on the first coding.

Table 5.1 Reliability of News Sentiment Coding

The daily news is recoded into *Good, Bad, NewHigh* (*Highbench* for the case of South China Morning Post) and *NewLow* (for the case of South China Morning Post), mainly based on its headline, then by its contents. The *Newhigh* is a subset of *Good*, while *NewLow* is a subset of *Bad.* The first recode is performed from 4th August 2010 to 19th August 2010. The second recode is performed from 29th November 2010 to 7th December 2010.

	South China Morning Post		New Straits Times		The Straits Times	
Code	First count	Second count	First count	Second count	First count	Second count
New High/Highbech	36	36	33	33	71	71
New Low/Lowbench	14	14	14	14	30	30
Good	229	228	109	112	267	266
Bad	172	171	57	56	208	209
Neutral	2990	2992	2935	2933	1884	1884
Number of trading days in this period	3391	3391	3101	3101	2359	2359

Table 5.2 illustrates the summary of variance measures by contract. The full sample period of HSIF and KLCIF covers from 1/1996 to 12/2008. The samples are then split in half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for SiMSCI is from 1/1999 through 12/2008, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. Panel A and B summarise the realised variances constructed using 5-minutes intraday data; Panel C and D summarise the historical daily variances; Panel E and F summarise the conditional variances. The full sample period summary statistics for the HSIF, KLCIF and SiMSCIF are presented, followed by first-half and second-half sample periods. All the returns variance series are excess leptokurtic and skewed to the right. Based on mean values, the variance of KLCIF ranks the highest, followed by HSIF and lastly SiMSCIF. This indicates that the KLCIF is the most volatile among the three index futures. KLCIF is the smallest in terms of trading volume and value (see Table 4.7), which makes it more vulnerable to change in information and sentiment. The excess returns of all the three index futures experience more turbulence during sub-period 1 than in sub-period 2. This could be due to the major stock market crash attributed to dot-com bubble of the late 1990s. The time series of the variances are highly autocorrelated and statistically significant at $\alpha = 0.01$.

Table 5.2 Summary of Variance Measures by Contract

This table summarises the returns variance measures of Hang Seng Index Futures (HSIF), Kuala Lumpur Composite Index Futures (KLCIF) and Morgan Stanley Singapore free Index Futures (SiMSCIF). The full sample period of HSIF and KLCIF covers 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for SiMSCI from 1/1999 through 12/2008, sub-period 1 from 1/1999 through 12/2003, sub-period 2 span from 1/2004 through 12/2008. Panel A and B reports the descriptive statistics for realised variance; Panel C and D reports the Historical variance and Panel E and F reports conditional variance. + all the autocorrelations are significant at α=0.01

period	Mean $\times 10^{-3}$	Maximum $\times 10^{-3}$	Minimum $\times 10^{-3}$	Std. Dev. $\times 10^{-3}$	Skewness	Kurtosis	Autocorrelation 1^{+}	2^{+}	3^{+}	4^{+}	5^{+}
Panel A : Daily Average Absolute Returns Using 5-minutes Intraday Data (*AAR*)											
Hang Seng Index Futures											
Full period	0.481	4.448	0.000	0.317	2.853	20.406	0.764	0.721	0.692	0.671	0.648
Sub-period 1	0.566	4.448	0.000	0.339	2.801	22.253	0.694	0.646	0.607	0.584	0.551
Sub-period 2	0.395	2.745	0.000	0.266	3.222	18.172	0.828	0.785	0.765	0.741	0.730
Kuala Lumpur Composite Index Futures											
Full period	0.394	8.310	0.056	0.376	7.999	130.689	0.773	0.724	0.701	0.640	0.629
Sub-period 1	0.528	8.310	0.056	0.490	6.728	86.678	0.745	0.692	0.664	0.595	0.585
Sub-period 2	0.271	1.314	0.072	0.137	2.199	11.531	0.762	0.684	0.657	0.604	0.556
Singapore Morgan Stanley Composite Index Futures											
Full period	0.021	0.518	0.001	0.032	7.554	89.868	0.732	0.682	0.571	0.543	0.532
Sub-period 1	0.022	0.160	0.001	0.016	2.772	16.307	0.605	0.464	0.396	0.399	0.371
Sub-period 2	0.020	0.518	0.001	0.042	6.526	59.807	0.752	0.717	0.599	0.566	0.557
Panel B : Daily Average Absolute Returns using 5-minutes Intraday Data (*SSR*)											
Hang Seng Index Futures											
Full period	0.031	1.799	0.000	0.066	14.078	306.244	0.459	0.409	0.421	0.342	0.308
Sub-period 1	0.038	1.799	0.000	0.079	14.440	280.631	0.365	0.320	0.358	0.245	0.206
Sub-period 2	0.023	0.776	0.000	0.049	7.545	81.083	0.687	0.620	0.561	0.571	0.549
Kuala Lumpur Composite Index Futures											
Full period	0.043	11.992	0.000	0.300	31.925	1156.270	0.543	0.350	0.319	0.216	0.236
Sub-period 1	0.075	11.992	0.000	0.430	22.324	562.819	0.539	0.344	0.312	0.208	0.228
Sub-period 2	0.013	0.354	0.001	0.019	8.007	108.715	0.569	0.457	0.448	0.362	0.316
Singapore Morgan Stanley Composite Index Futures											
Full period	0.324	1.588	0.077	0.157	1.743	8.933	0.830	0.782	0.740	0.722	0.707
Sub-period 1	0.370	1.079	0.116	0.121	1.220	6.143	0.693	0.603	0.521	0.513	0.481
Sub-period 2	0.275	1.588	0.077	0.175	2.596	12.172	0.873	0.839	0.811	0.786	0.778
Panel C: Daily Rolling Window Volatility (*RW*)											
Hang Seng Index Futures											
Full period	0.402	57.511	0.000	1.477	21.810	744.942	0.340	0.244	0.212	0.193	0.109
Sub-period 1	0.507	57.511	0.000	1.839	20.875	600.332	0.346	0.186	0.190	0.173	0.084
Sub-period 2	0.294	16.265	0.000	0.968	9.740	125.502	0.303	0.439	0.274	0.245	0.181
Kuala Lumpur Composite Index Futures											
Full period	0.413	144.919	0.000	3.293	33.811	1357.583	0.487	0.227	0.227	0.132	0.101
Sub-period 1	0.707	144.919	0.000	4.723	23.690	662.702	0.485	0.221	0.222	0.126	0.094
Sub-period 2	0.142	5.117	0.000	0.330	6.810	70.045	0.154	0.227	0.236	0.147	0.196
Singapore Morgan Stanley Composite Index Futures											
Full period	0.223	6.713	0.000	0.492	5.554	46.717	0.243	0.210	0.187	0.135	0.168
Sub-period 1	0.246	4.928	0.000	0.458	4.311	28.578	0.146	0.067	0.050	0.036	0.070
Sub-period 2	0.199	6.713	0.000	0.525	6.418	57.328	0.318	0.321	0.293	0.210	0.244

Table 5.2 (continued)

period	Mean $X10^{-3}$	Maximum $X10^{-3}$	Minimum $X10^{-3}$	Std. Dev. $X10^{-3}$	Skewness	Kurtosis	Autocorrelation 1+	2+	3+	4+	5+
Panel D: Daily Volatility (*V*)											
Hang Seng Index Futures											
Full period	0.417	52.799	0.000	1.464	18.919	572.577	0.410	0.264	0.227	0.229	0.142
Sub-period 1	0.525	52.799	0.000	1.813	18.310	471.048	0.421	0.213	0.186	0.213	0.115
Sub-period 2	0.307	13.528	0.000	0.978	8.557	94.066	0.357	0.422	0.349	0.267	0.218
Kuala Lumpur Composite Index Futures											
Full period	0.431	150.145	0.000	3.405	33.976	1369.361	0.474	0.210	0.214	0.133	0.111
Sub-period 1	0.734	150.145	0.000	4.883	23.817	668.840	0.472	0.204	0.208	0.126	0.104
Sub-period 2	0.151	5.724	0.000	0.351	6.851	72.812	0.200	0.186	0.242	0.144	0.214
Singapore Morgan Stanley Composite Index Futures											
Full period	0.233	5.973	0.000	0.510	5.270	40.434	0.238	0.239	0.195	0.188	0.178
Sub-period 1	0.255	5.030	0.000	0.483	4.477	30.838	0.157	0.067	0.056	0.043	0.065
Sub-period 2	0.209	5.973	0.000	0.537	5.885	46.931	0.304	0.381	0.309	0.307	0.271
Panel E: Daily Conditional Volatility based on GARCH											
Hang Seng Index Futures											
Full period	0.420	7.507	0.054	0.587	5.360	45.420	0.985	0.962	0.935	0.906	0.875
Sub-period 1	0.526	7.507	0.077	0.659	5.518	45.961	0.981	0.951	0.917	0.880	0.841
Sub-period 2	0.312	4.561	0.054	0.479	4.641	28.903	0.990	0.978	0.962	0.944	0.925
Kuala Lumpur Composite Index Futures											
Full period	0.443	30.470	0.029	1.406	12.828	211.389	0.966	0.907	0.845	0.780	0.716
Sub-period 1	0.742	30.470	0.032	1.979	9.121	106.393	0.964	0.903	0.839	0.771	0.704
Sub-period 2	0.167	1.069	0.029	0.157	2.458	9.683	0.966	0.936	0.904	0.866	0.830
Singapore Morgan Stanley Composite Index Futures											
Full period	0.244	2.091	0.035	0.229	3.316	19.608	0.978	0.954	0.929	0.904	0.880
Sub-period 1	0.268	1.216	0.061	0.152	1.482	6.246	0.949	0.893	0.840	0.791	0.747
Sub-period 2	0.220	2.091	0.035	0.286	3.429	16.980	0.986	0.971	0.954	0.936	0.917
Panel F: Daily Conditional Volatility based on Threshold GARCH											
Hang Seng Index Futures											
Full period	0.405	6.922	0.063	0.553	5.390	44.567	0.985	0.963	0.938	0.909	0.880
Sub-period 1	0.495	6.922	0.076	0.600	5.656	47.947	0.982	0.954	0.922	0.885	0.846
Sub-period 2	0.313	4.883	0.063	0.484	4.957	33.118	0.988	0.973	0.958	0.940	0.923
Kuala Lumpur Composite Index Futures											
Full period	0.442	31.421	0.031	1.413	13.287	227.758	0.951	0.888	0.829	0.764	0.703
Sub-period 1	0.735	31.421	0.038	1.991	9.468	114.911	0.949	0.884	0.822	0.755	0.692
Sub-period 2	0.171	1.325	0.031	0.175	2.821	12.217	0.961	0.927	0.893	0.852	0.810
Singapore Morgan Stanley Composite Index Futures											
Full period	0.246	2.186	0.037	0.240	3.573	21.840	0.978	0.956	0.932	0.909	0.885
Sub-period 1	0.266	1.329	0.064	0.156	1.764	8.322	0.949	0.898	0.847	0.799	0.754
Sub-period 2	0.225	2.186	0.037	0.303	3.549	17.966	0.985	0.972	0.955	0.939	0.920

I conduct the Granger-causality tests as the preliminary examination to discover the relationship between index futures returns, volatilities and news sentiment regimes. Table 5.3, Table 5.4 and Table 5.5 contain summaries for the pairwise Granger-causality tests run from news sentiment, volatility measures and index futures excess returns respectively.

The F-test statistics is tabulated for causality runs from variables in the rows to the variables in the columns.

Table 5.3 shows that there is a different Granger-causality pattern among the three index futures. *Bad and Lowbench* Granger-cause the volatility measures of Hang Seng Index Futures. The Singapore Morgan Stanley Index Futures displays the same pattern. However, in the case of the Kuala Lumpur Stock Index Futures, *Good* and *Newhigh* Granger-cause the volatility measures, more than the *Bad* and *Newlow*. *Newlow* (*Lowbench* for HSIF) Granger-cause excess returns for all the three index futures. This is consistent with early studies on the leverage effect (Braun, Nelson, & Sunier, 1995; French et al., 1987; Nelson, 1991), which confirm that a negative unexpected change in returns increases volatility more than a positive unexpected change in returns.

Table 5.4 tabulates the Granger-causality tests run from volatility measures to news sentiment and excess returns. Generally, the volatility measures granger-cause *Bad, Newlow* for the case of the HSIF and SiMSCIF. Volatility measures Granger-cause *Good* and *Newhigh*, for the case of the KLCIF. All the volatility measures of the three index futures returns Granger-cause their excess returns.

Table 5.5 reports the Granger-causality tests run from index futures excess returns to volatility measures and news sentiment. The excess returns significantly Granger-cause the volatility measures and news sentiment for all the three index futures.

Table 5.3 Pairwise Granger-causality Tests run from News Sentiment to Volatility Measures and Excess Returns

The F-test statistics is tabulated for causality runs from variables in the rows to the variables in the columns. *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH; *Good*=1 if the index goes down; *Highbench*=1 if the index reaches a higher benchmark; *Newhigh*=1 if the index hits a new high; *Bad* =1 if the index goes down; Lowbench=1 if the index decreases to a lower benchmark; *Newlow* if the index hits a new low; 0 otherwise. The sample period of HSIF and KLCIF covers 1/1996 to 12/2008. The sample period for SiMSCI from 1/1999 through 12/2008. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

	AAR	*SSR*	*RW*	*V*	*GARCH-M*	*TGARCH-M*	FREX
Panel A Hang Seng Index Futures (HSIF)							
GOOD	2.171*	2.011*	6.211***	0.241	1.673	0.053	0.992
HIGHBENCH	1.908	0.865	1.246	1.159	4.662***	0.466	0.855
BAD	2.292*	2.565**	8.752***	7.715***	6.319***	1.734	0.733
LOWBENH	16.183***	13.374***	24.669***	20.877***	19.396***	44.453***	4.011***
Panel B Kuala Lumpur Stock Index Futures (KLCIF)							
GOOD	1.538	2.193*	1.953*	1.804	5.847***	7.036***	1.052
NEWHIGH	0.276	0.12	3.252***	3.045***	10.697***	16.160***	0.76
BAD	2.571**	7.053***	0.717	0.697	0.484	0.315	1.6
NEWLOW	4.996***	11.708***	0.661	0.539	0.886	1.059	3.766***
Panel C Singapore Morgan Stanley Free Index Futures (SiMSCIF)							
GOOD	2.406**	2.320**	0.733	1.106	1.72	1.606	0.188
NEWHIGH	0.206	0.721	1.024	1.324	0.878	0.631	0.114
BAD	1.051	0.538	3.838***	3.667***	3.607***	3.209***	1.097
NEWLOW	3.647***	1.788	6.419***	6.580***	8.582***	7.549***	3.041***

Table 5.4 Pairwise Granger-Causality Tests Run from Volatility Measures to News Sentiment and Excess Returns

The F-test statistics is tabulated for causality runs from variables in the rows to the variables in the columns. *AAR* = Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH; *Good*=1 if the index goes down; *Highbench*=1 if the index reaches a higher benchmark; *Newhigh*=1 if the index hits a new high; *Bad* =1 if the index goes down; Lowbench=1 if the index decreases to a lower benchmark; *Newlow* if the index hits a new low; 0 otherwise. The sample period of Hang Seng Index Futures (HSIF) and Kuala Lumpur Composite Index Futures (KLCIF) covers 1/1996 to 12/2008. The sample period for Singapore Morgan Stanley Free Index Futures (SiMSCI) span from 1/1999 through 12/2008. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

	GOOD	*NEWHIGH*	*BAD*	*NEWLOW*	*FREX*
Panel A Hang Seng Index Futures					
AAR	2.082*	0.903	4.982***	9.599***	6.068***
SSR	3.160**	1.516	6.420***	18.319***	19.924***
RW	0.333	16.958***	2.758***	4.423***	2.07*
V	5.583***	15.916***	5.232***	7.711***	2.367*
GARCH-M	1.039	0.242	2.028*	3.748***	2.606**
TGARCH-M	0.842	0.968	1.917	4.787***	5.869***
Panel B Kuala Lumpur Stock Index Futures					
AAR	6.486***	5.858***	3.277***	1.901	43.694***
SSR	7.728***	15.539***	0.289	2.163	112.151***
RW	8.347***	3.886***	0.435	0.421	32.193***
V	8.305***	3.868***	0.429	0.474	34.547***
GARCH-M	8.739***	0.422	0.056	0.461	28.162***
TGARCH-M	8.568***	0.196	0.056	0.301	21.401***
Panel C Singapore Morgan Stanley Free Index Futures					
AAR	1.507	1.049	7.423***	3.547***	7.484***
SSR	0.507	1.664	6.433***	3.880***	3.424***
RW	2.158*	2.568**	3.850***	4.398***	2.950**
V	2.473**	4.405***	6.496***	6.963***	2.577**
GARCH-M	1.994	1.64	1.832	1.557	5.763***
TGARCH-M	2.120*	1.672	2.134*	2.013*	3.538***

Table 5.5 Pairwise Granger-Causality Tests Run from Index Futures Excess Returns to Volatility Measures and News Sentiment

The F-test statistics is tabulated for causality runs from variables in the rows to the variables in the columns. *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH; *Good*=1 if the index goes down; *Highbench* if the index reaches a higher benchmark; *Newhigh*=1 if the index hits a new high; *Bad* =1 if the index goes down; Lowbench=1 if the index decreases to a lower benchmark; *Newlow* if the index hits a new low; 0 otherwise. The sample period of Hang Seng Index Futures (HSIF) and Kuala Lumpur Composite Index Futures (KLCIF) covers 1/1996 to 12/2008. The sample period for Singapore Morgan Stanley Free Index Futures (SiMSCI) span from 1/1999 through 12/2008. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Panel A Granger Causality run from index futures excess returns to volatility measures

	AAR	*SSR*	*RW*	*V*	*GARCH-M*	*TGARCH-M*
HSIF Excess Returns	23.571***	28.363***	74.830***	63.257***	51.151***	326.759***
KLCIF Excess Returns	9.917***	26.999***	103.654***	110.822***	114.946***	205.998***
SiMSCIF Excess Returns	14.799***	12.981***	2.564**	2.398**	42.176***	10.964***

Panel B Granger-causality run from index futures returns to news sentiment

	GOOD	*HIGHBENCH*	*BAD*	LOWBENCH
HSIF Excess Returns	34.246***	28.413***	47.611***	30.871***
KLCIF Excess Returns	13.214***	8.886***	8.079***	3.207***
SiMSCIF Excess Returns	35.696***	19.421***	43.269***	16.268***

The above findings suggest three important points. First, the excess returns and volatility measures show bi-directional Granger-causality. Second, excess returns Granger-cause all news sentiment, except the *Newlow* results in bidirectional Granger-causality with excess returns. Wang, Keswani and Taylor (2006) include three bearish sentiment proxies (S&P 100 OEX put-call trading volume ratio, S&P OEX put-call open interest ratio and NYSE ARMS index) that yield similar results. Third, the granger causality pattern between news sentiment and volatility measures is clearly different among the three index futures. On one hand, HSIF and SiMSCIF show bi-directional Granger-causality among volatility measures and unfavourable news sentiment (i.e. *Bad, Newlow, and Lowbench*). On the other hand, KLCIF shows bi-directional Granger-causality among volatility measures and favourable news sentiment (i.e. *Good, Newhigh*). The large body literature on asymmetry returns volatility suggest that that there is asymmetry impact of good news (unexpected positive returns) and bad news (unexpected returns) on returns volatility. Bad news and forecast higher next day's volatility while good news forecast lower volatility. The leverage hypothesis ((Black, 1976; Christie, 1982; Duffee, 1995), the feedback hypothesis (Campbell & Hentschel, 1992; French et al., 1987), the prospect theory (Kahneman & Tversky, 1979), the representativeness and affect heuristics explanation by (Shefrin, 2008) are used to explain this phenomenon. However, the finding on KLCIF is not consistent with these explanations. One of the possible explanations is that 3% of the news from The Straits Times as compared to only 1% of the news from the South China Morning Post and New Straits Times are coded as *Newhigh* (see Table 5.1). Chen and Ghysels (2011) suggest that the extreme good news predict higher volatility as bad news does.

The findings imply that news sentiment does not possess direct predictive power over excess returns, but may indirectly predict excess returns through its relationship with volatility measures. This justifies the motivation to examine the mean-variance trade-off.

Next, I examine the stationarity of the data. Table 5.6 reports the t-statistics of the Augmented Dickey Fuller test. Since all the p-values are close to zero, the null hypothesis of a unit root for the series can be rejected. All the time series are stationary.

Table 5.6 Augmented Dickey Fuller Tests for Index Futures Excess Returns and Volatility Measures
This table reports the t-statistics of the Augmented Dickey Fuller test for the time series that enter the structural vector autoregressive model. The null hypothesis test for a unit root in the time series (i.e. non-stationary). *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH. The full sample period of the Hang Seng Index Futures (HSIF), and Kuala Lumpur Composite Index Futures (KLCIF) cover 1/1996 to 12/2008.The full sample period for the Singapore Morgan Stanley Free Index Futures (SiMSCIF) span 1/1999 through 12/2003. ** indicates the test statistics significant at 1% level.

Contracts	HSIF		KLCIF		SiMSCIF	
Series	t	Probability	t	Probability	t	Probability
Excess Returns	-59.890**	0.000	-59.963**	0.000	-51.505**	0.000
AAR	-5.591**	0.000	3.468**	0.000	-5.777**	0.000
SSR	-6.019**	0.000	8.110**	0.000	-6.099**	0.000
RW	-19.141**	0.000	-18.662**	0.000	-7.779**	0.000
V	-18.398**	0.000	-18.647**	0.000	-7.634**	0.000
GARCH-M	-7.498**	0.000	-9.772**	0.000	-5.078**	0.000
TGARCH-M	-7.327**	0.000	-10.832**	0.000	-5.092**	0.000

5.3 MEAN-VARIANCE RELATIONSHIP DURING HIGH SENTIMENT PERIOD

Table 5.7 illustrates the base model of the mean-variance relationship, equations 14(a) and 14(b), without the news sentiment regimes.

$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \varepsilon_{t+1}$ Base model 14(a)

$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \varepsilon_{t+1}$ Base model 14(b)

where the dependent variable is the daily excess index futures returns of the HSIF, KLCIF and SiMSCIF. Excess return equals the daily return less risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate where the Three-month treasury bill discount rate is used for HSIF and SiMSCIF while the one-month Kuala Lumpur interbank offer rate is used for KLCIF as the risk-free rate. Var_t is conditional variance while V_t is realised variance (see Section 3.3.6 for the detail specifications of the variance measures).

The Hang Seng Index Futures excess returns is negatively associated with realised volatility (*AAR* and *SSR*), but positively related to rolling window volatility (*RW*). There is no evidence of an association between HSIF excess returns and historical daily volatility (*V*), conditional volatility based on the GARCH model and Threshold GARCH model. The coefficients of determination (R^2) are very low for all the six equations. The R^2 is the highest, 0.035, when the variance measure is *RW*. The Kuala Lumpur Composite Index Futures (KLCIF) excess returns are not significantly associated with all variance measures. The Singapore Morgan Stanley Free Index Futures (SiMSCIF) excess returns are significantly related to the realised variance, *AAR* and *SSR*. The slope coefficients are -52.712 and -9.439; and the R^2 are 0.012 and 0.009 respectively. In conclusion, the mean-variance relationships implied by the base model are inconclusive. The low R^2 and the inconsistent signs of the slope coefficients suggest an omitted variable problem. This study intends to explore the role of investor sentiment in explaining the mean-variance relationship. During periods of high sentiment, noise traders who dominate the trading activity weaken the mean-variance relationship (Yu and Yuan, 2011). On the contrary, low investor sentiment interacts with variance and makes the mean-variance slope to become steeper.

I estimate Equation (14a) using realised variance (*AAR* and *SSR*) and historical variance (*RW* and *V*) while Equation (14b) uses conditional variance (*GARCH-M* and *TGARCH-M*; across three index futures and two high sentiment dummy variables (*Good* and *Newhigh/Highbench*). The equations are estimated using the ordinary least squares method. Since the sample size is large enough, the Newey-West consistent covariance estimator is used as the remedy for the heteroskedasticity and autocorrelation problems (Gujarati, 2003).

$$R_{t+1} - F_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \quad (14a)$$

$$R_{t+1} - F_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \quad (14b)$$

Hypothesis 8 (see Chapter 3) is tested based on the sign of coefficient, β_1. Capital Asset Pricing Theory implies that the coefficient, β_1 should carry a positive sign as the investors require compensation for bearing higher risks. On the other hand, the noise trader model suggests that the coefficient, β_1, can be negative if the Friedman effect is greater than the create space effect, and negative if the create space effect is greater than Friedman effect. If the coefficient, α_2, is statistically significant and positive, this implies that news sentiment is directly priced as suggested by Hypothesis 9. Based on Hypothesis 10, I expect the coefficient, β_2 , to be negative during a high sentiment regime and positive during a low sentiment regime, as the impact of noise traders is prevalent during a high sentiment period.

Table 5.8 reports the coefficients and t-test statistics, to test the Hypothesis 8, Hypothesis 9 and Hypothesis 10. Compared with the results in Table 5.7, the KLCIF equations experienced the greatest improvement in R^2 after the *Good* and *Newhigh* are included in the equations. For HSIF and KLCIF, the R^2 improves more when *Newhigh* is added to the equations (14a) and (14b) as compares to when *Good* is added to the equation. However, SiMSCIF shows the opposite. In Table 5.8, the R^2 has increased from 0.001 to 0.184 (see Panel A1) and 0.195 (see Panel A2), after the *Good* and *Newhigh* regime is added to the equations with *SSR* as measures of variance. These result in the increment of R^2 by 18.3 and 19.4 percentage points respectively. The news sentiment markedly improves the mean-variance estimation, implying that the noise traders have more prominent roles in KLCIF. In Yu and Yuan (2011), the R^2 increase by 1 to 3 percentage points. The equations of HSIF and SiMSCIF also improved, although at a smaller magnitude. The HSIF equations show improvement in R^2, ranging from 0.1 to 0.9 percentage points, while ranging from 0 to 0.9 percentage point for SiMSCIF.

The columns of Table 5.8 compare the results of the HSIF, KLCIF and SiMSCIF. Panels A1 and A2 report the cross-country results using a 5-minute intraday realised volatility measures (*AAR* and *SSR*); Panels B1 and B2 report the results using historical daily volatility measures (*RW* and *V*); Panel C1 and C2 reports the results using conditional volatility (*GARCH-M* and *TGARCH-M*) as proxy for risks. The Panels A1, B1 and C1 consist of the estimates based on *Good* Regime. The Panel A2, B2 and C2 reports the results for *Newhigh* regime.

5.3.1 Tests of Hypothesis 8 in the high-sentiment regime

I first discuss the β_1 by index futures, followed by volatility measures, and finally measures of investor sentiment.

The first two columns of Table 5.8 depict the results for the HSIF. In Panel A1, the coefficients, β_1 , are -5.42 and -34.003, and statistically significant at 5% and 1% respectively when *AAR* and *SSR* is used as a proxy for volatility. In Panel A2, the coefficients, β_1 , are -5.333 and -34.928. In Panel B1, the signs of the same coefficients turn positive, equal to 2.645 and 16.37, and significant at 1% and 5% respectively when *RW* and *V* is used as a proxy for volatility. In Panel B2, the coefficients, β_1, are 2.664 and 1.643. In panel C, the coefficients, β_1 , are positive but insignificant for all cases. Overall, the realised variance slope coefficients are positive and significant; the historical variance slope coefficients are negative and significant; while the conditional variance slope coefficients are positive but insignificant. In conclusion, there is insufficient evidence to conclude Hypothesis 8, due to the findings varying with the variance measures.

The third and fourth columns of Table 5.8 list the results for the KLCIF. The coefficients, β_1 ,are all positive (23.755 in Panel A1 and 23.417 in Panel A2), statistically significant at 1% when *SSR* is used as proxy for volatility but insignificant when *AAR* is used. Panels B1 and B2, shows estimates based on historical daily volatility measures, all the coefficients are positive and significant at 1%. In Panel C1, the β_1 are positive but insignificant for both conditional variances measures, when *Good* is used as proxy for news sentiment. In Panel C2, the β_1 are positive and significant at 10% when *Newhigh* is used together with volatility measures generated by both *GARCH-M* and *T-GARCHM*. With respect to the KLCIF, there is strong evidence to accept Hypothesis 8 that there is a positive mean-variance trade-off. This also implies that the create space effect outweighs the Friedman effect.

The fifth and sixth columns report the results for the SiMSCIF. Panel A1 shows that the coefficient, β_1 ,is -32.638 but insignificant, when *AAR* is used as measures of volatility, and

Good is used as measures of news sentiment; β_1 ,is -6.258 and significant at 5% when *SSR* is used as measure of variance. In Panel A2, coefficients, β_1 , are -50.842 and -90113 and are significant at the 1% level. The signs of coefficients, β_1, are mixed and insignificant in Panels B1 and B2. When *RW* is used as proxy of volatility, the coefficients, β_1 , are positive while the signs are negative when *V* is used. There is no evidence that the variance slope coefficients are significant in Panels C1 and C2 when the GARCH-M and T-GARCHM are used as the variance measures. There is some weak evidence to conclude that Hypothesis 8 is supported, and that the Friedman effect is greater than the create space effect.

In conclusion, the evidence found on the KLCIF data is consistently positive and statistically significant, as suggested by the Capital Asset Pricing Model. This also can be attributed to a stronger create space effect than a Friedman effect. However, the relationship between volatility measures and excess returns depends heavily on the proxy of volatility itself, for the case of the HSIF and SiMSCIF. The SiMSCIF volatility consistently results in downward revision in the excess returns, but is only significant when the realised volatility is used. The negative impact of the Friedman effect on excess returns is stronger than the positive impact on excess returns driven by noise traders who crowd out rational investors. The HSIF excess returns are inversely related to realised volatility but positively related to the historical volatility measures. The results also corroborate the Lundblad (2007) findings that there is limited relationship between conditional volatility and realised returns.

5.3.2 Tests of Hypothesis 9 in the high-sentiment regime

The noise trader theory suggests that news sentiment has an impact on required return although news sentiments are formed based on stale information. If the investors have biased beliefs, news sentiment may influence the investors' perception on risk. Consequently, the coefficient, α_2, should be positive if this perceived risk is priced.

The first two columns of Table 5.8 report the results for the HSIF. The coefficients, α_2, are insignificant, for all measures of volatilities and all measures of sentiment, except when *RW* is the measure of volatility and *Newhigh* is the measure of news sentiment (see Panel B2). The coefficient, α_2, is 0.002 and significant at 10% level. The evidence is too weak to support Hypothesis 9.

The third and fourth columns summarise the results for the KLCIF. In Panel A1, the coefficients, α_2, is 0.020 and significant in the *Good* regime when *AAR* is used as variance measure; but insignificant when *SSR* is used as a proxy for volatility. However, in Panel A2, when the news sentiment reaches a *Newhigh*, the coefficients, α_2, becomes positive and significant (0.021 and 0.004). The change in news sentiment from *Neutral* to *Newhigh* predicts a parallel shift in excess returns for the KLCIF. The coefficients, α_2, are not significant in Panels B1 and C1, but are positive and significant in Panels B2 and C2, when the high sentiment is more extreme. There is convincing evidence to support Hypothesis 9.

The fourth and fifth columns display the results for the SiMSCIF. The coefficient, α_2, are insignificant when *RW, V, GARCH-M* and *TGARCH-M* are used as measures of volatility. It is -6.258 and significant at 5% when the volatility measures is *SSR* in *Good* regime (see Panel A1); coefficient, α_2 is 0.004 and significant at 5% when the volatility measure is *AAR* in *Newhigh* regime (see Panel A2). The results are inconclusive at best, as the findings are sensitive to variance measures. Hypothesis 9 is not supported.

Overall, I do not find strong evidence suggesting that excess returns of the HSIF and SIMSCIF are affected by shifts in sentiment regime. The sentiment risk is not directly priced, which implies that the role of noise traders are limited in these two index futures markets. Since we are unable to disentangle the price pressure effect and hold more effect using the regression model, we could not rule out the possibility that the insignificance $\alpha_2 s$ are due to the two effects cancelling each other out. However, the impact of stale information prevails among investors of the KLCIF, on the next day after the market index reaches a new high level. The magnitude of increase in excess returns due to the hold more effect is larger than the decrease in excess returns due to the price pressure effect. The sentiment risk is directly priced and the roles of noise traders prevail in the KLCIF. This could be due to the fact that KLCIF is traded in a developing market and the total trading value of the index futures is smaller (see Table 1.1), thus it is more easily affected by noise traders.

5.3.3 Tests of Hypothesis 10 in the high-sentiment regime

I examine the significance of the coefficient, β_2, to confirm whether there is an indirect impact of sentiment in predicting the excess returns through its interaction with volatility; specifically attenuating the mean-variance relationship.

The estimated coefficient, β_2, are negative for all the equations estimated using the HSIF's data, but are only significant when historical daily volatility and conditional volatility are used during *Highbench* regime. In Panel B2, the coefficients are -18.810 and -14.227 with both being significant at the 1% level. In Panel C2, the coefficients are -7.184 (significant at 10%) and -9.444 (significant at 5%). The findings based on historical volatility and conditional volatility support Hypothesis 10.

The evidence for the KLCIF is strong and consistent. Ten out of twelve of the β_2 estimates are negative and significant at 1%. Two exceptions are made when conditional variance is used during the *Good* news sentiment regime. When *GARCH-M* is used, the coefficient, β_2, is -5.653 and significant at 10%. Although the coefficient is negative when TGARCH-M is used, it is insignificant. There is strong evidence to accept Hypothesis 10.

The analysis on the SiMSCIF yields consistent negative β_2, most significant when realised volatility is used. In Panel A1, *Good* regime, the coefficient, β_2, is -96.822 when *AAR* is used and -23.687 when *SSR* is used. In Panel A2 *Newhigh* regime, the coefficient, β_2, is -230.983 when *AAR* is used. All these slope estimates are significant at 1%. In addition, In Panel B1, when *V* is used in the *Good* regime, the coefficient, β_2, is -7.673 and significant at 10%. No significance evidence is found in Panels C1 and C2, when conditional variances are used. There is evidence to support Hypothesis 10, although the results vary with the variance measure.

Overall, there is clear evidence that bullish news sentiment attenuates the mean-variance relationship, despite being highly sensitive to volatility measures. Ranked by the number of significant β_2 , the impact of sentiment traders is most prevalent in the KLCIF, followed by SiMSCIF, and finally HSIF. This is consistent with the findings by Baker and Wurgler (2006), showing that assets that are low in capitalisation and high in volatility are hard to arbitrage, hence more prone to investor sentiment. In addition, Lee et al. (1991) find that closed-end fund discount, as a measure of sentiment, is highly correlates with small stocks' prices.

Table 5.7 The Base Model of Mean-variance Relationship

This table reports the slope coefficients regressions for the base model of mean-variance relationship.

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \varepsilon_{t+1}$$
$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \varepsilon_{t+1}$$

Where the dependent variable is daily excess index futures returns, namely Hong Kong, Malaysia and Singapore. Excess return equals the daily return less risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate where Three-month treasury bill discount rate is used for Hong Kong and Singapore while one-month Kuala Lumpur interbank offer rate is used for Malaysia as the risk-free rate. Var_t is conditional variance while V_t is realised variance. *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH. The sample periods begin in January 1996 for Hang Seng Stocks Index Futures (HSIF) and Kuala Lumpur Stock Index Futures (KLCIF); in April January 1999 for Singapore Morgan Stanley Composite Index Futures (SiMSCIF). The sample period ends at December 2008 for each index futures. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

HSIF	HSIF Excess Returns			KLCIF Excess Returns			SiMSCIF Excess Returns		
Variance	α_1	β_1	Adjusted R^2	α_1	β_1	Adjusted R^2	α_1	β_1	Adjusted R^2
AAR	0.003***	-5.525**	0.007	0.000	0.313	0.000	0.001***	-52.712***	0.012
	[2.578]	[-2.304]		[-0.174]	[0.083]		[2.605]	[-2.781]	
SSR	0.001**	-35.46***	0.007	0.000	2.547	0.001	0.003***	-9.439***	0.009
	[2.520]	[-2.797]		[-0.597]	[0.382]		[3.392]	[-3.014]	
RW	-0.001***	2.579***	0.035	0.000	-0.614	0.009	0.000	0.296	0.000
	[-3.114]	[4.019]		[0.159]	[-0.907]		[-0.241]	[0.148]	
V	-0.001*	1.536*	0.012	0.000	-0.764	0.015	0.001	-2.779	0.009
	[-1.743]	[1.869]		[0.432]	[-1.279]		[1.639]	[-1.415]	
GARCH-M	0.000	0.566	0.000	0.000	0.097	0.000	0.000	-1.257	0.000
	[-0.540]	[0.595]		[-0.604]	[0.246]		[0.585]	[-0.642]	
TGARCH-M	0.000	1.249	0.001	0.000	0.360	0.001	0.000	-1.065	0.000
	[-1.032]	[1.069]		[-1.015]	[1.141]		[0.493]	[-0.558]	

Table 5.8 Mean-variance Relationship during Period of High News Sentiment
This table reports the slope coefficients regressions to test the Hypothesis 8, 9 and 10:

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \quad (14a)$$
$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \quad (14b)$$

Where the dependent variable is daily excess index futures returns, namely Hong Kong, Malaysia and Singapore. Excess return equals the daily return less risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate where Three-month treasury bill discount rate is used for Hong Kong and Singapore while one-month Kuala Lumpur interbank offer rate is used for Malaysia as the risk-free rate. Var_t is conditional variance while V_t is realised variance. *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH. *Sentiment* is dummy variable refers to sentiment measures namely *Good and Newhigh*. *Good* = 1 if the daily routine news reports the market rise on the prior day; *Newhigh*=1 if the market climbed to new high; otherwise=0. These sentiment measures apply to HSIF, KLCIF and SiMSCIF with two exceptions. For the case of HSIF, *Highbench* is used instead of *Newhigh*. *Highbench* =1 if HSI rise to higher benchmark. The β_1 is expected to be positive following the capital asset pricing theory, or when create space effect is greater than Friedman effect; it is expected to be negative when Friedman effect is greater than create space effect. The α_2 is depend on the net impact of positive hold more effect and negative price pressure effect. The $\beta_2$is expected to be negative during period of high sentiment. The sample periods begin in January 1996 for Hang Seng Stocks Index Futures (HSIF) and Kuala Lumpur Stock Index Futures (KLCIF); in April January 1999 for Singapore Morgan Stanley Composite Index Futures (SiMSCIF). The sample period ends at December 2008 for each index futures. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Futures Index	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel A1 Realised returns volatility measures using 5-minutes intraday data						
Volatility measure	*AAR*	*SSR*	*AAR*	*SSR*	*AAR*	*SSR*
C	0.003**	0.001**	-0.003	-0.001***	0.001*	0.002**
	[2.530]	[2.518]	[-1.359]	[-2.854]	[1.685]	[2.45]
Volatility	-5.42**	-34.003***	8.63	23.775***	-32.638	-6.258**
	[-2.219]	[-2.679]	[1.273]	[9.869]	[-1.600]	[-2.082]
Good	0.001	0.001	0.020***	0.004	0.001	0.007***
	[-0.125]	[0.515]	[2.971]	[1.599]	[1.454]	[3.557]
Volatility**Good*	-1.391	-65.211	-42.844***	-62.397***	-96.822***	-23.687***
	[-0.215]	[-1.090]	[-2.809]	[-16.515]	[-3.174]	[-3.715]
Adjusted R^2	0.008	0.014	0.094	0.184	0.019	0.017
Panel A2 Realised returns volatility measures using 5-minutes intraday data						
C	0.003**	0.001***	-0.003	-0.001***	0.001**	0.003***
	[2.547]	[2.628]	[-1.373]	[-2.890]	[2.445]	[3.204]
Volatility	-5.333**	-34.928***	8.164	23.417***	-50.842***	-9.113***
	[-2.214]	[-2.777]	[1.288]	[9.816]	[-2.649]	[-2.861]
Newhigh	0.001	-0.001	0.021***	0.004**	0.004**	0.005
	[0.206]	[-0.572]	[6.539]	[2.239]	[2.409]	[1.515]
Volatility**Newhigh*	-6.956	-61.256	-59.426***	-65.134***	-230.98***	-19.39
	[-0.930]	[-0.839]	[-8.844]	[-27.213]	[-4.404]	[-1.472]
Adjusted R^2	0.008	0.014	0.134	0.195	0.014	0.01

Table 5.8 (Continued)

Index Futures	HSIF		KLCIF		SIMSCIF	
Dependent Variables	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel B1 Historical volatility using daily data						
Volatility measure	*RW*	*V*	*RW*	*V*	*RW*	*V*
C	-0.001***	-0.001	-0.001***	-0.001***	0.001	0.001
	[-2.892]	[-1.604]	[-3.592]	[-2.904]	[-0.306]	[1.488]
Volatility	2.645***	1.637*	3.16***	2.636***	0.943	-2.083
	[4.138]	[1.953]	[4.901]	[3.037]	[0.465]	[-1.046]
Good	0.001	0.001	0.003	0.003	0.001	0.001
	[0.042]	[0.269]	[1.379]	[1.461]	[1.095]	[0.883]
Volatility**Good*	-3.014	-3.885	-5.707***	-5.141***	-9.571	-7.673*
	[-0.586]	[-0.83]	[-7.032]	[-5.24]	[-1.614]	[-1.802]
Adjusted R^2	0.036	0.014	0.19	0.172	0.007	0.015
Panel B2 Historical volatility using daily data						
C	-0.001***	-0.001*	-0.001***	-0.001***	0.001	0.001
	[-3.003]	[-1.654]	[-3.598]	[-2.803]	[-0.276]	[1.629]
Volatility	2.664***	1.643**	3.054***	2.509***	0.33	-2.821
	[4.134]	[1.962]	[4.545]	[2.767]	[0.164]	[-1.427]
Newhigh	0.002*	0.001	0.003***	0.003**	0.001	-0.001
	[1.839]	[1.01]	[2.134]	[1.989]	[0.736]	[-0.727]
Volatility**Newhigh*	-18.81***	-14.227***	-5.745***	-5.105***	-5.728	8.694
	[-10.113]	[-9.106]	[-8.546]	[-5.627]	[-0.398]	[0.544]
Adjusted R^2	0.044	0.021	0.198	0.174	0.007	0.009
Panel C1 Conditional volatility measures using daily data						
Volatility measure	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*
C	0.001	-0.001	-0.001	0.001	0.001	0.001
	[-0.717]	[-1.082]	[-1.538]	[-1.427]	[0.200]	[0.095]
Volatility	1.108	1.836	1.119	0.988	-0.054	0.157
	[0.920]	[1.243]	[1.567]	[1.585]	[-0.027]	[0.080]
Good	0.001	0.001	0.002	0.001	0.001	0.001
	[0.511]	[0.668]	[0.753]	[0.100]	[0.569]	[0.594]
Volatility**Good*	-4.581	-5.96	-5.653*	-3.823	-6.46	-6.639
	[-1.110]	[-1.211]	[-1.789]	[-1.466]	[-1.357]	[-1.421]
Adjusted R^2	0.002	0.004	0.026	0.012	0.002	0.002
Panel C2 Conditional volatility measures using daily data						
C	0.001	-0.001	-0.001	0.001	0.001	0.001
	[-0.755]	[-1.131]	[-1.561]	[-1.471]	[0.448]	[0.373]
Volatility	1.131	1.87	1.06*	0.974*	-1.052	-0.889
	[0.969]	[1.292]	[1.764]	[1.838]	[-0.521]	[-0.453]
Newhigh	0.001	0.001	0.007***	0.009***	0.003	0.004
	[0.323]	[0.597]	[4.041]	[4.441]	[1.182]	[1.326]
Volatility**Newhigh*	-7.184*	-9.444**	-23.905***	-35.021***	-15.256	-19.502
	[-1.726]	[-2.026]	[-36.262]	[-51.146]	[-1.113]	[-1.281]
Adjusted R^2	0.004	0.006	0.12	0.119	0.001	0.001

5.4 MEAN-VARIANCE RELATIONSHIP DURING LOW SENTIMENT PERIOD

Table 5.9 reports the regressions slope coefficients to test the Hypothesis 8, 9 and 10 during the period of low sentiment.

Compared to the base models in Table 5.7, the equation (14a) and (14b) show improvement in the perspective of the coefficient of determination, R^2. Similar with the regressions during the high sentiment period in Table 5.8, the KLCIF equations experience the greatest improvement among the three index futures. The R^2 increment, ranging from 0.6 to 1.8 percentage points. However, this is lower than the regressions during the high sentiment period, which are up to 19.4 percentage points higher. The HSIF and SiMSCIF show the same range of improvement as compared to the high sentiment period. The R^2 increment of the HSIF equations range from 0 to 0.9 percentage points, while for the SiMSCIF equations range from 0.1 to 1.4 percentage points. The findings suggest that the low sentiment has greater influence on KLCIF as compared to HSIF and SiMSCIF.

First, I discuss the evidence by country, then by volatility measures. The first two columns report the results for the HSIF; the third and fourth columns report the results for the KLCIF and the last two columns reports the results for the SiMSCIF. Panels A1 and A2 reports the regression coefficients using a 5-minutes realised returns volatility (*AAR* and *SSR*); Panels B1 and B2 reports regression coefficients using historical daily returns volatility (*RW* and *V*); while Panels C1 and C2 report regressions coefficients using daily conditional volatility measures (*GARCH-M* and *TGARCH-M*). The Panels A1, B1 and C1 reports the estimates during *Bad* regime, while the Panels A2, B2 and C2 reports the estimates during *Newlow* regime.

5.4.1 Tests of Hypothesis 8 in the low-sentiment regime

As an extension to Yu and Yuan (2011) who only focus on a high sentiment regime, I test Hypothesis 8 in a low sentiment regime. During a low sentiment regime, bearish investors suffer misperceptions on risks, overreact to bad news, sell the assets at abnormally low prices, and finally suffer capital loss. This Friedman effect leads to lower expected returns. Rational investors are reluctant to trade because market risk increase as the shift of sentiment is highly uncertain. Consequently, sentiment traders enjoy higher expected returns for the risks they created. This is call create space effect. The coefficient, β_1, is the summation of Friedman effect and create space effect. I first discuss the evidence by country, then by volatility measures.

The evidence on the HSIF is inconsistent across volatility measures. Based on Panel A1, the coefficients, β_1, are -6.209 and -42.340; are -5.964 and -38.143 in Panel A2; all these coefficients are significant at 5%. In Panel B1, the coefficients, β_1, are 2.769 and 1.885, both significant at 1% and 5% respectively. Panel B2 shows similar results, the coefficients are 2.668 and 1.711, significant at 1% and 5% respectively. There is no significant result in Panel C. The evidence suggests that the Friedman effect is dominant when the realised volatility is used and the create space effect is dominant when historical volatility is used. The findings are not robust across measures of volatility. In conclusion, Hypothesis 8 is not supported.

All the coefficient,β_1, are positive but insignificant in Panels A1, A2, C1 and C2 for the KLCIF; and negative but insignificant in Panels B1 and B2. Neither the Friedman effect nor the create space effect prevail, implying that the expectation toward risks is not directly priced. There is no evidence to support Hypothesis 8. However, the later evidence on Hypothesis 11 shows that the risks are indirectly priced.

Evidence on the SiMSCIF is more consistent. The coefficients, β_1, are negative except in Panel B2, the coefficient is 0.239 but insignificant. The coefficients, β_1, are significant when AAR and SSR are used as the variance measure. In Panel A1, the coefficients are -87.114 and -13.150, while they are -55.562 and -9.698 in Panel A2. All these coefficients are significant at 1%. This suggests that the Friedman effect cancels the create space effect, the investors suffer capital loss in feedback trading but did not benefit much from crowding out the rational investors. Hypothesis 8 is supported for the case of realised returns volatility, but not for historical volatility and conditional volatility.

Amid the mixed findings of the HSIF, the SiMSCIF's returns are negatively related to excess returns; the KLCIF's excess returns show no significant relationship with its volatility. Although the CAPM suggests the mean-variance relationship should be positive, Backus and Gregory (1993) state that "It can be increasing, decreasing, flat, or even non-monotonic. The shape depends on both the preferences of the representative agent and the probability structure across states". This is evident by extant studies (see Brandt & Kang, 2004; Campbell, 1987; Glosten et al., 1993; Harvey, 2001; Li, 2011; Whitelaw, 1994).

5.4.2 Tests of Hypothesis 9 in the low-sentiment regime

The hold more effect on expected returns is always negative, because bearish noise traders tend to sell the assets although the price is too low. The selling spree creates a negative price pressure effect on expected returns. Taken together the hold more effect and price pressure effect, I expect the coefficient, α_2, to be negative during low sentiment period.

The evidence on the HSIF is by no means conclusive because the coefficient, α_2 ,are negative and insignificant in Panels A1 and A2; positive and significant in Panel B1 (the α_2 is

0.003, t-value is 1.712) and B2 (α_2 is 0.009, significant at 10%; 0.012, significant at 1%); and negative and significant in Panel C1 (α_2 are -0.004 and -0.005) but insignificant in Panel C2. The mixed results arising from different measures of variance lead to an inconclusive outcome for Hypothesis 9.

The news sentiment is not priced in the KLCIF because the coefficient, α_2 , is not statistically significant in all panels. There is no evidence to support Hypothesis 9.

.

Interestingly, the evidence on the SiMSCIF is totally consistent and significant. The coefficient, α_2 , are significant, equal to -0.003 and -0.009 in Panel A1; -0.010 and -0.014 in Panel A2; -0.003 in Panel B1; -0.007 in Panel B2; -0.005 and -0.005 in Panel C1; lastly -0.014 and -0.013 in Panel C2. The negative impact of *Newlow* on excess returns is greater than the negative impact of *Bad*, for example coefficient, α_2 , is -0.005 when sentiment is *Bad* , while equals -0.014 in the *Newlow* regime. The evidence supports Hypothesis 9 that the net impact of the price pressure effect and the hold-more effect results in significantly lower excess returns during the periods of *Bad* and *Newlow* sentiment.

Only the SiMSCIF data corroborates the sentiment theory. There are two possible explanations for the HSIF and KLCIF. First, there is no panic selling in a low sentiment regime. Second, the buying of rational traders offsets the selling spree of bearish traders. Simon and Wiggins lll (2001) find that buying when the stock market is in extreme fear enhances risk adjusted profit, because of the undervaluation of assets.

5.4.3 Tests of Hypothesis 10 in the low-sentiment regime

Yu and Yuan (2011) propose that bullish traders are reluctant to go short during high sentiment periods. Corollary, I propose that sentiment traders are reluctant to go long during low sentiment period. Instead, rational traders take advantage of the low price but require higher compensation for taking the sentiment risk. This leads to a steeper mean-variance trade-off, the coefficient, β_2, is expected to be positive.

By and large, the evidence on HSIF data is insignificant. However, there are two contradicting findings. In Panel B2, the coefficient, β_2, is -9.646 in *Newlow* regime and significant at 1%. In Panel C2, coefficient, β_2, are 7.991 and 8.641, significant at 5% and 1% respectively, in the *Bad* regime. Hypothesis 10 is not supported.

The KLCIF data yield unexpected signs. In Panel A1, the coefficient, β_2, are -45.653 and -176.869 respectively when *AAR* and *SSR* is used in the *Newlow* regime; both coefficients are significant at 1%. All the coefficient, β_2, are negative and significant in Panel B1 (-7.687 and -6.511) and Panel B2 (-10.386 and -8.169). The coefficients in Panel C1 are negative and significant at 1%, which are -13.100 and -14.800. In Panel C2, it is insignificant when *TGARCH-M* is used in the *Newlow* regime. Hypothesis 10 predicts a positiveβ_2, while the findings on the KLCIF show the opposite. Given that the coefficient, β_2, is generally more negative than the coefficient, β_1, the overall mean-variance relationship turns negative ($i.e.\beta_1 + \beta_2 < 0$). The KLCIF investors are expected lower returns for bearing higher risks.

The analysis on the SiMSCIF data yield consistent positive β_2. These are significant in Panel A1, where the coefficients, β_2, are 131.090 and 26.609; while equal to 192.639 in Panel A2, except when *SSR* is used in the *Newlow* regime. In Panel B1, it is significant when *RW* is used (the β_2 is 11.121) but insignificant when *V* is used. The magnitudes of the coefficients are larger in Panel C2 than in the Panel C1. The coefficients are 30.994 and 26.629 in Panel C2 (significant at 1%), while are 17.810 and 16.812 in Panel C1 (significant at 5%). Hypothesis 10 is supported.

In conclusion, there is no evidence that noise traders profoundly affect the mean-variance trade-off of the HSIF. There is evidence that rational traders in the SiMSCIF take advantage of low prices when noise traders are bearish. The mean-variance slope coefficient becomes steeper because rational arbitrageurs need to be compensated with higher returns for taking the sentiment risk. With respect to KLCIF, taken together, the summation of β_1 and β_2 is negative, the negative mean-variance trade-off implies that traders are expecting lower returns for taking high risks. I perform analysis on these three index futures contracts using the same variance measures, mean-variance model specification, sample period and sampling frequency. Hence, the divergence of the KLCIF from a positive mean-variance relationship can be attributed to rational expectation of investors, including leverage hypothesis or volatility feedback effect. Alternatively, the behavioural explanations, including representativeness and heuristics as suggested by Shefrin (2008), and noise traders explanation by Yu and Yuan (2011) justify the negative mean-variance relationship.

Table 5.9 Mean-variance Relationship during Period of Low News Sentiment
This table reports the slope coefficients regressions to test the Hypothesis 8, 9, and 10:

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \quad (14a)$$
$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \quad (14b)$$

Where the dependent variable is daily excess index futures returns, of three markets, namely Hong Kong, Malaysia and Singapore. Excess return equals the daily return less risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate where Three-month treasury bill discount rate is used for Hong Kong and Singapore while one-month Kuala Lumpur interbank offer rate is used for Malaysia as the risk-free rate. Var_t is conditional variance while V_t is realised variance. *AAR*= Daily Average Absolute Returns using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH. *Sentiment* is dummy variable refers to sentiment measures namely *Bad and Newlow* . *Bad* = 1 if the daily routine news reports the market fell on the prior day; *Newlow*=1 if the market dipped to new low; otherwise=0. These sentiment measures apply to HSIF, KLCIF and SiMSCIF with two exceptions. For the case of HSIF, *Lowbench* is used instead of *Newlow*. *Lowbench* =1 if HSI fell to lower benchmark. The β_1 is expected to be positive following the capital asset pricing theory, or when create space effect is greater than Friedman effect; it is expected to be negative when Friedman effect is greater than create space effect. The α_2 is depend on the net impact of positive hold more effect and negative price pressure effect and it is expected to be negative during period of low sentiment. The $\beta_2$is expected to be positive during period of low sentiment. The sample periods begin in January 1996 for Hang Seng Stocks Index Futures (HSIF) and Kuala Lumpur Stock Index Futures (KLCIF); in April January 1999 for Singapore Morgan Stanley Composite Index Futures (SiMSCIF). The sample period ends at December 2008 for each index futures. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel A1 Realised returns volatility measures using 5-minutes intraday data						
Volatility measure	*AAR*	*SSR*	*AAR*	*SSR*	*AAR*	*SSR*
C	0.003***	0.001**	0.000	0.000	0.002***	0.004***
	[2.703]	[2.120]	[-0.224]	[-0.189]	[3.886]	[3.590]
Volatility	-6.209**	-42.340**	0.916	2.776	-87.114***	-13.150***
	[-2.413]	[-2.248]	[0.236]	[0.414]	[-3.858]	[-3.211]
Bad	-0.001	0.000	0.002	-0.005	-0.003**	-0.009**
	[-0.201]	[0.256]	[0.195]	[-1.573]	[-2.163]	[-2.484]
Volatility*Bad	3.535	21.36	-18.763	-52.387	131.090***	26.609**
	[0.514]	[0.706]	[-1.103]	[-1.266]	[2.578]	[2.395]
Adjusted R^2	0.008	0.015	0.006	0.007	0.026	0.018
Panel A2 Realised returns volatility measures using 5-minutes intraday data						
C	0.003**	0.001**	0.000	0.000	0.001***	0.003***
	[2.288]	[2.267]	[-0.266]	[-0.329]	[2.712]	[3.289]
Volatility	-5.964**	-38.143**	0.923	2.768	-55.562***	-9.698***
	[-2.072]	[-2.373]	[0.242]	[0.414]	[-2.653]	[-2.875]
Newlow	0.001	0.003	0.007	-0.007	-0.010***	-0.014*
	[0.099]	[0.361]	[0.633]	[-0.673]	[-3.082]	[-1.665]
Volatility*Newlow	4.471	29.683	-45.653***	-176.869***	192.639**	25.274
	[0.260]	[0.372]	[-2.691]	[-2.951]	[2.343]	[1.112]
Adjusted R^2	0.008	0.014	0.012	0.013	0.015	0.011

Table 5.9 (Continued)

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel B1 Historical returns volatility measures using daily data						
Volatility measure	*RW*	*V*	*RW*	*V*	*RW*	*V*
C	-0.001***	-0.001**	0.000	0.000	0.000	0.001*
	[-3.306]	[-2.19]	[0.475]	[0.743]	[0.322]	[1.763]
Volatility	2.769***	1.885**	-0.561	-0.703	-0.918	-3.415
	[4.239]	[2.337]	[-0.793]	[-1.127]	[-0.375]	[-1.492]
Bad	0.001	0.003*	-0.002	-0.002	-0.003***	-0.002
	[0.774]	[1.712]	[-1.116]	[-1.058]	[-2.92]	[-1.55]
Volatility*Bad	-2.753	-5.072	-7.687**	-6.511**	11.121***	9.223
	[-0.81]	[-1.400]	[-2.037]	[-2.55]	[2.633]	[1.328]
R-squared	0.037	0.021	0.021	0.028	0.013	0.015
Panel B2 Historical returns volatility measures using daily data						
C	-0.001***	-0.001**	0.000	0.000	0.000	0.001*
	[-3.248]	[-2.037]	[0.355]	[0.617]	[-0.067]	[1.787]
Volatility	2.668***	1.711**	-0.556	-0.698	0.239	-2.761
	[3.968]	[2.117]	[-0.787]	[-1.119]	[0.117]	[-1.383]
Newlow	0.009*	0.012***	-0.006	-0.006	-0.007**	-0.004
	[1.930]	[2.944]	[-1.242]	[-1.345]	[-2.553]	[-1.130]
Volatility*Newlow	-7.653	-9.646***	-10.39***	-8.169***	9.329	-4.631
	[-1.060]	[-2.637]	[-11.150]	[-9.168]	[0.766]	[-0.237]
Adjusted R^2	0.038	0.021	0.027	0.032	0.002	0.01
Panel C1 Conditional volatility measures using daily data						
Volatility measure	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*
C	0.000	0.000	0.000	0.000	0.001	0.001
	[-0.069]	[-0.492]	[-0.235]	[-0.617]	[1.586]	[1.522]
Volatility	0.052	0.575	0.165	0.425	-3.884	-3.731
	[0.053]	[0.472]	[0.396]	[1.248]	[-1.529]	[-1.480]
Bad	-0.004**	-0.005**	-0.002	-0.001	-0.005**	-0.005**
	[-2.098]	[-2.381]	[-0.784]	[-0.393]	[-2.501]	[-2.427]
Volatility*Bad	7.991**	8.641***	-13.10***	-14.80***	17.810**	16.812**
	[2.479]	[2.759]	[-2.636]	[-2.582]	[2.301]	[2.239]
R-squared	0.003	0.005	0.008	0.009	0.009	0.009
Panel C2 Conditional volatility measures using daily data						
C	0.000	0.000	0.000	0.000	0.000	0.000
	[-0.506]	[-0.971]	[-0.355]	[-0.761]	[0.885]	[0.793]
Volatility	0.497	1.162	0.147	0.416	-1.626	-1.427
	[0.492]	[0.94]	[0.359]	[1.242]	[-0.795]	[-0.711]
Newlow	0.000	-0.001	-0.011	-0.004	-0.014***	-0.013***
	[0.063]	[-0.09]	[-1.321]	[-0.509]	[-3.854]	[-3.784]
Volatility*Newlow	2.093	2.511	-17.394	-26.532*	30.994***	26.629***
	[0.240]	[0.246]	[-1.306]	[-1.924]	[3.194]	[3.031]
Adjusted R^2	0.001	0.001	0.01	0.013	0.004	0.004

5.5 SUB-SAMPLE ANALYSIS

The literature review in Chapter 2 concludes that one of the explanations for the mixed mean-variance relationship is the sampling issues. To further the investigation, I split the sample into sub-samples, with equal time lengths. The full sample period for the Hang Seng Index Futures (HSIF), and the Kuala Lumpur Composite Index Futures (KLCIF) cover from 1/1996 to 12/2008. The samples are then split into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Singapore Morgan Stanley Free Index Futures (SiMSCIF) range from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008.

5.5.1 Sub-sample analysis of mean-variance relationships during periods of high sentiment

Table 5.10 and Table 5.11 report analysis for sub-sample 1 and sub-sample 2 respectively in a high sentiment regime. Panel A reports the results on realised variance (*AAR* and *SSR*); Panel B reports the results on historical variance (RW and V); and Panel C reports the results on conditional variance (*GARCH-M* and *TGARCH-M).* Panels A1, B1 and C1 report the coefficient estimates in the *Good* Regime. PanelsA2, B2 and C2 report the results in the *Newhigh* regime.

Referring to Panel A of Table 5.10 and Table 5.11, the Hang Seng Index Futures' (HSIF) mean-variance coefficient, β_1, are significant and consistent over the two sample periods. The realised variances (*AAR* and *SSR*) are negatively related to excess returns, implying that the Freidman effect outweighs the create space effect. In Panel B, the historical variances (*RW* and *V*) are positively related to excess returns, implying that the create space effect outweighs the Friedman effect. In Panel C, the conditional variances (*GARCH-M* and *TGARCH-M*) have no significant relationship with excess returns. The variation in the direction of the mean-variance relationship can be attributed to different measures of variance, instead of the sampling period. The *Good* and *Highbench* sentiment coefficient, α_2, are insignificant in sub-samples and for all the variance measures. These indicate that there is neither a hold-more effect nor a price pressure effect, leading to the conclusion that sentiment is not directly priced in the HSIF. The extreme sentiment, *Highbench* is indirectly priced through the interaction with variance measures (*RW, V, GARCH-M* and *TGARCH-M*) during sub-sample 1. The slope coefficient for the interaction term, $\beta_{2,}$are 20.156, -15.128, -10.122, and -13.377 respectively; all these coefficients are significant at 1% level. This evidence supports Hypothesis 10, that the investors are less responsive to risk during periods of high sentiment. The same coefficients are insignificant in sub-sample 2. This could be due to the average returns in sub-period 1 being higher than sub-period 2 (see Table 4.8). The high investor sentiment is more prominent during sub-sample 1.

Based on the third and fourth columns of Table 5.10 and Table5.11, the Kuala Lumpur Composite Index Futures (KLCIF) is most responsive to the *Good* and *Newhigh* sentiment among the three index futures. The coefficients for the variance, β_1, are positive and significant during sub-sample 1, for all the six measure of variance. Conversely, in Table 5.11, when the sentiment reaches *Newhigh* in the sub-sample 2, the coefficient, β_1, equal to -14.143, -120.777, and -5.582 in the equation with *AAR, SSR* and *V* as measure of variance. The results are similar when the sentiment is *Good*. The results in sub-sample 2 dominate the results of the full sample which supports there being a positive mean-variance relationship, *ceteris paribus*. Based on Table 5.10, there is no evidence of the hold-more effect nor price-pressure effect during sub-sample 1. Table 5.11 indicates that during the sub-sample 2, the hold more effect significantly dominates the price pressure effect. The sentiment risk is priced during the sub-sample 2, while not priced during the sub-sample 1. Although the sentiment risk is positively priced, the Table 5.10 shows that during sub-period 1, the investors are less responsive to the variance in the *Good* and *Newhigh* sentiment regimes.

The Singapore Morgan Stanley Free Index Futures (SiMSCIF) excess returns show little evidence of the mean-variance relationship. Most of the coefficients are insignificant for when *RW, V, GARCH-M* and *TGARCH-M* are used as measures of variance. Overall, the evidence is too weak to be conclusive, in the high sentiment regime. In the panel A of Table 5.11, during sub-period 2, the *AAR* and *SSR* are negatively related to excess returns. In the *Good* regime the coefficients, β_1, are -43.498 and -12.300; in the *Newhigh* regime, the coefficients, β_1, are 59.855 and -15.277. The variance slope coefficients, β_1, are not significant during sub-period 1. The results in Panel B of Table 5.10 and Table 5.11 are inconsistent. In Panel A of the two tables, although not unanimous, the evidence from the *AAR* and *SSR* equations supports Hypothesis 10 that the noise traders weaken the mean-variance relationship in the high sentiment regime. The evidence is stronger during sub-period 2. Panel C of table 5.11 shows that during sub-period 2, the *GARCH-M* and *T-GARCHM* interacts with *Newhigh* sentiment and weaken the mean-variance relationship. The coefficients, β_2, are -27.119 and -28.722 respectively.

There are three implications that arise from these findings. First, the magnitude of create space effect and the Friedman effect varies over time, leading to a positive or negative mean-variance relationship, resulting from the net value of these two effects. The mean-variance relationship slope coefficient, β_1 are mostly positive during sub-sample 1 and negative during sub-sample 2. Second, the sentiment coefficient, α_2 are positive and significant for the case of the KLCIF during sub-sample 2. The evidence on the HSIF and SiMSCIF are insignificant. Third, there is weaker mean-variance relationship in the high sentiment regime. The evidence on the interaction term, β_2 is relatively strong, for the HSIF and KLCIF. The coefficients are negative and mostly significant during sub-period 2 in the *Newhigh* regime.

5.5.2 Sub-sample analysis of mean-variance relationships during periods of low sentiment

Table 5.12 and Table 5.13 report results for sub-sample 1 and sub-sample 2 during the period of low sentiment. The Panels A, B and C report the findings on realised variance (*AAR* and *SSR*), historical variance (*RW* and *V*), and conditional variance respectively (*GARCH-M* and *TGARCH-M*).

In Panel A of Table 5.12, during sub-sample 1, for the equations that *Bad* is used as a sentiment measure, all the coefficients, β_1 , α_2, and β_2 are insignificant for all measures of variance measures and for all contracts. However, there is an exception. The *Bad* sentiment has a significant negative impact on KLCIF excess returns, the coefficient, $\alpha_{2,}$ is -0.008. For the equations where *Newlow* is used as a sentiment measure, the *SSR* of HSIF positively predicts the excess returns, the coefficient, β_1, is -40.373; while the *SSR* of KLCIF positively predicts the excess returns, the coefficient, β_1 , is 24.156. The *Newlow* sentiment is negatively priced when *AAR* is used as the variance measure for the HSIF, the coefficient, α_2 is -0.024; while it is positively priced when *SSR* is used for the KLCIF, the coefficient, α_2 , is 0.007. The impact of the interaction between sentiment and variance is inconsistent among the three contracts. The coefficient β_2 is 31.319 and 125.372 for HSIF; -52.735 and -66.110 for KLCIF. All the coefficients are significant at 1%. There is no evidence of the *Newlow* sentiment weakening the mean-variance relationship of SiMSCIF.

Based on Panel B of Table 5.12, the HSIF shows the strongest evidence supporting Hypotheses 8, 9 and 10. In the equations where *Bad* is used as the sentiment measures, the coefficient, β_1, are 2.484 and 2.235; while the coefficient, β_1, equal to 11.411 and 12.933 when the sentiment is *Newlow*. All the coefficients are significant at 1%. The *Bad* sentiment is positively priced but no evidence is found for *Newlow* sentiment. The *Bad* sentiment interacting with variance brings a negative impact to excess returns; the coefficient, β_2, are -0.9080 and -10.614, both the coefficients are significant at 1%. In contrast, the *Newlow* sentiment interacts with variance, positively affecting the excess returns; the coefficient, β_2, are 11.411 and 12.933. The *Bad* and *Newlow* sentiments are negatively priced in the KLCIF. The *Newlow* has a greater negative impact on the KLCIF compared to *Bad* sentiment. The coefficient for *Bad* is -0.004 while the coefficient for *Newlow* is -0.011. There is evidence of a weaker mean-variance relationship during the period of low sentiment, the coefficients, β_2, are -9.788 and -7.604 for the interactions with RW and V. The SiMSCIF excess returns are lowered, due to interaction between *Newlow* and RW and V. The coefficient, β_2, are -33.545 and -35.255 respectively.

In Table 5.12 panel C, the measures of variance are *GARCH-M* and *TGARCH-M*. The HSIF does not show evidence of variance being directly priced, the $\beta_1$are insignificant for all cases. The coefficient, α_2, is equal to -0.004 in the *TGARCH-M* equation, implying that the price pressure effect and the hold-more effect have a negative impact on returns. The *Bad* sentiment is directly priced. The *Bad* sentiment interacts with the variances, the coefficient, β_2, are 5.858 and 7.205; while the same coefficients are equal to 10.112 and 13.740 in the *Newlow* sentiment regime. The evidence on the KLCIF is weaker. The coefficients of the interactions are inconsistent with the prediction of Hypothesis 10. The coefficient, β_2, are -12.351 and -14.015 in the *Bad* regime and are -13.643 and -23.283 in the *Newlow* regime. For the SiMSCIF, the volatility coefficient, β_1, are positive and significant at 1% in the *Newlow* equations.

Table 5.13 reports the analysis on sub-sample 2. The Panel A consists of coefficients of regressions using *AAR* and *SSR* as measures of variance. In the equations where *Bad* sentiment is used, the variance coefficients, β_1, are negative and significant for the HSIF, KLCIF and SiMSCIF. The variance has the largest impact on KLCIF excess returns, with the coefficient equal to -136.668. The HSIF excess returns is the least affected, where the coefficient, β_1, is -0.9388. In the equations where *Newlow* sentiment is used, the results are similar. The *Newlow* sentiment is positively priced in the HSIF, the coefficients,α_2, are -51.390 and -215.869. It is also positively priced in the SiMSCIF, the coefficients, α_2, are 245.072 and 54.867. The interactions *Newlow* with *AAR* and *SSR* negatively affect the HSIF excess returns. The coefficient, β_2, are -51.390 and -215.869 respectively. The *SSR* interacts with *Newlow* sentiment, increases the KLCIF excess returns, the coefficient, β_2, is 100.199. The evidence for SiMSCIF shows that the investors are more responsive to the risk in the *Bad* and *Newlow* sentiment regime. The coefficients, β_2, range from 37.969 to 245.072 and the coefficients are significant at 1%.

Panel B of Table 5.13 reports the sub-sample 2 analysis on the relationship between the index futures excess returns and historical variances, RW and V. The HSIF excess returns are positively affected by *Newlow*, the coefficients, α_2 , are 0.022 and 0.020 in the RW and V equations. The variance measures negatively predict the excess returns, the coefficients, β_2, are -19.828 and -13.771. These findings violate the prediction of Hypotheses 9 and 10. The results on the KLCIF also contradict with the prediction of Hypothesis 10 that the investors are more responsive to risk during times of low sentiment. The coefficient, β_2, is -24.757 in the *Bad* regime and are -18.783 and -15.202 in the *Newlow* regime. The *Bad* and *Newlow* sentiment is negatively priced in the SiMSCIF, as proposed by the Hypothesis 10. The coefficient, β_2, is 18.494 and significant at 1% in the *Newlow* regime.

Panel C of Table 5.13 reports the results on conditional volatility, GARCH-M and TGARCH-M during sub-sample period 2. The evidence on the SiMSCIF supports the Hypotheses 8, 9 and 10. The volatility is negatively related to excess returns, the coefficients are negative and significant in the *Bad* equations, the coefficients, β_1, are -7.140 and -6.520. The SiMSCI excess returns are significantly lower in the *Bad* (the coefficients, α_2, are-0.005 and -0.005) and *Newlow* (the coefficients, α_2, are-0.016 and -0.015) regimes. The HSIF mean-variance relationship becomes steeper in the *Bad* regime, the coefficients, β_2, are 11.567 and 10.469. However, the mean-variance relationship is weaker in the *Newlow* regime, which is inconsistent with the prediction of Hypothesis 10. The KLCIF shows no significant evidence.

In conclusion, the sub-sample analysis reveals that the low sentiment is not as consistent as high sentiment, in explaining the mean-variance relationship. This is based on the fact that the findings are inconsistent over the six measures of variance. In addition, the evidence also contradicts the predictions by the noise trader theory, as evident by failure to support Hypotheses 9 and 10.

Table 5.10 Sub-Sample 1: Mean-variance Relationship during Period of High News Sentiment
This table reports the slope coefficients regressions to test the Hypothesis 8, 9 and 10:

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \quad (14a)$$
$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \quad (14b)$$

Where the dependent variable is daily excess index futures returns, namely Hong Kong, Malaysia and Singapore. Excess return equals the daily return less risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate where Three-month treasury bill discount rate is used for Hong Kong and Singapore while one-month Kuala Lumpur interbank offer rate is used for Malaysia as the risk-free rate. Var_t is conditional variance while V_t is realised variance. *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH. Sentiment is dummy variable refers to sentiment measures namely *Good* and *Newhigh*. *Good* = 1 if the daily routine news reports the market rise on the prior day; *Newhigh*=1 if the market climbed to new high; otherwise=0. These sentiment measures apply to HSIF, KLCIF and SiMSCIF with two exceptions. For the case of HSIF, *Highbench* is used instead of *Newhigh*. *Highbench* =1 if HSI rise to higher benchmark. The β_1 is expected to be positive following the capital asset pricing theory, or when create space effect is greater than Friedman effect; it is expected to be negative when Friedman effect is greater than create space effect. The α_2 is depend on the net impact of positive hold more effect and negative price pressure effect. The β_2 is expected to be negative during period of high sentiment. The full sample period of the Hang Seng Index Futures (HSIF), and the Kuala Lumpur Composite Index Futures (KLCIF) cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Singapore Morgan Stanley Free Index Futures (SiMSCIF) from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel A1 Realised returns volatility measures using 5-minutes intraday data						
Volatility measure	*AAR*	*SSR*	*AAR*	*SSR*	*AAR*	*SSR*
C	0.002	0.001	-0.007*	-0.002***	-4.02E-04	-0.001
	[1.230]	[1.552]	[-1.860]	[-2.892]	[-0.403]	[-0.400]
Volatility	-4.345	-30.409*	12.553*	24.508***	22.144	2.311
	[-1.131]	[-1.805]	[1.709]	[11.512]	[0.446]	[0.402]
Good	0.001	0.003	0.026***	0.006	0.004	0.007
	[0.145]	[1.010]	[3.099]	[1.555]	[1.360]	[1.064]
Volatility**Good*	-3.799	-124.317	-48.052***	-63.285***	-263.874*	-23.924
	[-0.403]	[-1.404]	[-3.079]	[-18.534]	[-1.781]	[-1.219]
Adjusted R^2	0.005	0.016	0.133	0.228	0.006	0.003
Panel A2 Realised returns volatility measures using 5-minutes intraday data						
C	0.002	0.001	-0.006*	-0.002***	-1.89E-04	-3.62E-04
	[1.249]	[1.633]	[-1.866]	[-2.930]	[-0.179]	[-0.180]
Volatility	-4.282	-31.612*	11.746*	24.156***	4.219	0.685
	[-1.145]	[-1.906]	[1.716]	[11.522]	[0.079]	[0.115]
Newhigh	0.004	0.002	0.027***	0.007***	0.010	0.017
	[0.675]	[0.531]	[5.974]	[2.906]	[1.241]	[0.996]
Volatility**Newhigh*	-14.54	-186.096	-63.639***	-66.11***	-407.776	-40.468
	[-1.233]	[-1.441]	[-9.122]	[-31.316]	[-1.137]	[-0.928]
Adjusted R^2	0.007	0.016	0.179	0.241	0.001	0.001

Table 5.10 (Continued)

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel B1 Historical volatility using daily data						
Volatility measure	*RW*	*V*	*RW*	*V*	*RW*	*V*
C	-0.001**	-0.001*	-0.002***	-0.002***	-3.12E-04	-3.73E-05
	[-2.574]	[-1.882]	[-3.765]	[-3.220]	[-0.485]	[-0.058]
Volatility	2.667***	1.989**	3.318***	2.817***	1.844	0.702
	[3.590]	[2.140]	[5.728]	[3.574]	[0.602]	[0.228]
Good	-8.70E-05	0.001	0.005	0.005	0.002	0.001
	[-0.039]	[0.338]	[1.375]	[1.465]	[1.359]	[1.042]
Volatility**Good*	-3.207	-4.589	-5.872***	-5.329***	-15.522**	-13.065*
	[-0.571]	[-0.938]	[-7.981]	[-5.956]	[-1.998]	[-1.883]
R-squared	0.045	0.026	0.238	0.216	0.014	0.012
Panel B2 Historical volatility using daily data						
C	-0.001***	-0.001**	-0.002***	-0.002***	-2.66E-04	2.39E-05
	[-2.738]	[-1.965]	[-3.803]	[-3.139]	[-0.442]	[0.040]
Volatility	2.674***	1.975**	3.201***	2.676***	0.888	-0.275
	[3.569]	[2.119]	[5.179]	[3.179]	[0.301]	[-0.093]
Newhigh	0.002	0.001	0.007***	0.006***	0.004*	0.001
	[1.030]	[0.555]	[2.857]	[2.671]	[1.896]	[0.212]
Volatility**Newhigh*	-20.156***	-15.128***	-5.907***	-5.287***	-16.526	5.710
	[-15.806]	[-12.211]	[-9.552]	[-6.277]	[-0.995]	[0.230]
Adjusted R^2	0.056	0.035	0.247	0.218	0.002	2.61E-04
Panel C1 Conditional volatility measures using daily data						
Volatility measure	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*
C	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	[-1.090]	[-1.273]	[-1.583]	[-1.452]	[-1.294]	[-1.112]
Volatility	2.001	2.911	1.186	1.046	4.774	4.266
	[1.154]	[1.310]	[1.580]	[1.602]	[1.587]	[1.382]
Good	0.002	0.003	0.003	0.001	-0.004	-0.004
	[0.691]	[0.803]	[0.616]	[-0.014]	[-1.543]	[-1.458]
Volatility**Good*	-6.497	-8.301	-5.737*	-3.861	9.266	8.507
	[-1.327]	[-1.353]	[-1.706]	[-1.406]	[1.146]	[1.008]
R-squared	0.006	0.008	0.031	0.015	0.005	0.004
Panel C2 Conditional volatility measures using daily data						
C	-0.001	-0.001	-0.001	-0.001	-0.002*	-0.001*
	[-1.170]	[-1.379]	[-1.596]	[-1.485]	[-1.917]	[-1.695]
Volatility	2.042	3.025	1.117*	1.024*	6.035**	5.423*
	[1.236]	[1.405]	[1.778]	[1.852]	[2.033]	[1.763]
Newhigh	0.002	0.003	0.01***	0.011***	0.004	0.004
	[0.669]	[1.020]	[4.305]	[4.513]	[0.904]	[0.871]
Volatility**Newhigh*	-10.122***	-13.377***	-24.18***	-35.446***	-11.561	-12.994
	[-2.742]	[-3.64]	[-36.921]	[-57.665]	[-0.668]	[-0.595]
Adjusted R^2	0.009	0.013	0.146	0.145	0.003	0.003

Table 5.11 Sub-Sample 2: Mean-variance Relationship during Period of High News Sentiment
This table reports the slope coefficients regressions to test the Hypothesis 8, 9 and 10:

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \quad (14a)$$
$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \quad (14b)$$

Where the dependent variable is daily excess index futures returns, namely Hong Kong, Malaysia and Singapore. Excess return equals the daily return less risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate where Three-month treasury bill discount rate is used for Hong Kong and Singapore while one-month Kuala Lumpur interbank offer rate is used for Malaysia as the risk-free rate. Var_t is conditional variance while V_t is realised variance. *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH. *Sentiment* is dummy variable refers to sentiment measures namely *Good and Newhigh* . *Good* = 1 if the daily routine news reports the market rise on the prior day; *Newhigh*=1 if the market climbed to new high; otherwise=0. These sentiment measures apply to HSIF, KLCIF and SiMSCIF with two exceptions. For the case of HSIF, *Highbench* is used instead of *Newhigh*. *Highbench* =1 if HSI rise to higher benchmark. The β_1 is expected to be positive following the capital asset pricing theory, or when create space effect is greater than Friedman effect; it is expected to be negative when Friedman effect is greater than create space effect. The α_2 is depend on the net impact of positive hold more effect and negative price pressure effect. The $\beta_2$is expected to be negative during period of high sentiment. The full sample period of the Hang Seng Index Futures (HSIF), and the Kuala Lumpur Composite Index Futures (KLCIF) cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Singapore Morgan Stanley Free Index Futures (SiMSCIF) from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel A1 Realised returns volatility measures using 5-minutes intraday data						
Volatility measure	*AAR*	*SSR*	*AAR*	*SSR*	*AAR*	*SSR*
C	0.003***	0.001***	0.004***	0.002***	0.001*	0.003***
	[2.952]	[2.583]	[4.013]	[4.050]	[1.845]	[3.074]
Volatility	-8.058**	-45.182**	-14.548***	-122.382***	-43.498*	-12.300***
	[-2.454]	[-2.122]	[-3.627]	[-4.134]	[-1.786]	[-2.622]
Good	-0.002	-0.001	-0.004	0.000	0.002	0.006***
	[-0.897]	[-0.697]	[-0.633]	[-0.073]	[1.530]	[2.815]
Volatility*Good*	3.997	20.284	16.415	85.473	-80.514**	-20.187***
	[0.594]	[0.491]	[0.699]	[0.399]	[-2.474]	[-2.598]
Adjusted R^2	0.014	0.015	0.026	0.034	0.041	0.045
Panel A2 Realised returns volatility measures using 5-minutes intraday data						
C	0.003***	0.001***	0.004***	0.002***	0.001***	0.004***
	[2.959]	[2.633]	[4.123]	[4.115]	[2.727]	[4.208]
Volatility	-7.976**	-45.327**	-14.143***	-120.777***	-59.855***	-15.277***
	[-2.444]	[-2.140]	[-3.757]	[-4.338]	[-2.818]	[-3.618]
Newhigh	-0.003	-0.002	-0.006	-0.002	0.002	0.007***
	[-1.061]	[-1.264]	[-1.540]	[-0.749]	[1.484]	[2.27]
Volatility*Newhigh*	4.643	46.736	22.831*	177.098	-257.449***	-34.826***
	[0.608]	[1.184]	[1.851]	[1.263]	[-6.713]	[-2.626]
Adjusted R^2	0.014	0.016	0.025	0.034	0.036	0.039

Table 5.11 (Continued)

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel B1 Historical volatility using daily data						
Volatility measure	*RW*	*V*	*RW*	*V*	*RW*	*V*
C	-0.001	1.84E-05	0.001	0.001**	-2.39E-05	0.001**
	[-1.393]	[0.038]	[1.246]	[2.249]	[-0.045]	[2.056]
Volatility	2.632*	0.477	-3.816	-5.885*	0.218	-4.530
	[1.884]	[0.267]	[-1.001]	[-1.820]	[0.074]	[-1.613]
Good	0.000	-0.001	-0.002	-0.002	0.000	0.000
	[-0.091]	[-0.555]	[-1.166]	[-1.285]	[0.325]	[0.307]
Volatility**Good*	0.899	2.644	18.219	18.389	-4.753	-3.599
	[0.076]	[0.231]	[1.337]	[1.188]	[-0.504]	[-0.583]
R-squared	0.021	0.001	0.014	0.031	0.002	0.035
Panel B2 Historical volatility using daily data						
C	-0.001	1.86E-05	0.001	0.001**	-1.24E-05	0.001**
	[-1.410]	[0.040]	[1.168]	[2.178]	[-0.023]	[2.160]
Volatility	2.708*	0.545	-3.463	-5.582*	-0.114	-4.987*
	[1.954]	[0.306]	[-0.933]	[-1.769]	[-0.039]	[-1.883]
Newhigh	0.001	0.000	0.001	-0.002	0.000	-0.002
	[1.137]	[0.299]	[0.421]	[-1.023]	[-0.181]	[-1.066]
Volatility**Newhigh*	-13.871	-9.144	-6.223	24.536	0.875	11.118
	[-1.579]	[-0.877]	[-0.189]	[0.819]	[0.045]	[0.554]
Adjusted R^2	0.024	0.002	0.009	0.026	0.000	0.034
Panel C1 Conditional volatility measures using daily data						
Volatility measure	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*
C	2.69E-04	3.69E-05	3.85E-05	1.87E-05	3.82E-04	2.76E-04
	[0.444]	[0.056]	[0.091]	[0.047]	[0.671]	[0.486]
Volatility	-0.317	0.420	-0.088	0.031	-1.735	-1.192
	[-0.148]	[0.183]	[-0.033]	[0.013]	[-0.554]	[-0.397]
Good	-0.001	0.000	-0.002	-0.003	0.002	0.002
	[-0.557]	[-0.140]	[-0.620]	[-0.954]	[1.371]	[1.367]
Volatility**Good*	1.287	-0.553	12.949	16.880	-8.193	-8.189
	[0.388]	[-0.151]	[0.609]	[0.894]	[-1.455]	[-1.500]
R-squared	0.000	0.000	0.001	0.002	0.009	0.008
Panel C2 Conditional volatility measures using daily data						
C	2.98E-04	8.51E-05	-1.87E-05	-5.94E-05	6.65E-04	5.71E-04
	[0.500]	[0.131]	[-0.047]	[-0.154]	[1.323]	[1.137]
Volatility	-0.353	0.329	0.319	0.552	-3.163	-2.667
	[-0.165]	[0.144]	[0.136]	[0.252]	[-1.229]	[-1.075]
Newhigh	-0.002	-0.002	-0.001	-0.001	0.004	0.004
	[-1.187]	[-0.930]	[-0.325]	[-0.395]	[1.632]	[1.607]
Volatility**Newhigh*	2.719	1.471	7.011	8.930	-27.119**	-28.722**
	[1.018]	[0.445]	[0.886]	[1.142]	[-2.168]	[-2.181]
Adjusted R^2	4.45E-04	3.43E-04	1.80E-04	2.82E-04	0.007	0.006

Table 5.12 Sub-sample 1: Mean-variance Relationship during Period of Low News Sentiment
This table reports the slope coefficients regressions to test the Hypothesis 8, 9, and 10:

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \quad (14a)$$
$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \quad (14b)$$

Where the dependent variable is daily excess index futures returns, of three markets, namely Hong Kong, Malaysia and Singapore. Excess return equals the daily return less risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate where Three-month treasury bill discount rate is used for Hong Kong and Singapore while one-month Kuala Lumpur interbank offer rate is used for Malaysia as the risk-free rate. Var_t is conditional variance while V_t is realised variance. *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH. *Sentiment* is dummy variable refers to sentiment measures namely *Bad and Newlow*. *Bad* = 1 if the daily routine news reports the market fell on the prior day; *Newlow*=1 if the market dipped to new low; otherwise=0. These sentiment measures apply to HSIF, KLCIF and SiMSCIF with two exceptions. For the case of HSIF, *Lowbench* is used instead of *Newlow*. *Lowbench* =1 if HSI fell to lower benchmark. The β_1 is expected to be positive following the capital asset pricing theory, or when create space effect is greater than Friedman effect; it is expected to be negative when Friedman effect is greater than create space effect. The α_2 is depend on the net impact of positive hold more effect and negative price pressure effect and it is expected to be negative during period of low sentiment. The $\beta_2$is expected to be positive during period of low sentiment. The full sample period of the Hang Seng Index Futures (HSIF), and the Kuala Lumpur Composite Index Futures (KLCIF) cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Singapore Morgan Stanley Free Index Futures (SiMSCIF) from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The t values are reported in parentheses [ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel A1 Realised returns volatility measures using 5-minutes intraday data						
Volatility measure	*AAR*	*SSR*	*AAR*	*SSR*	*AAR*	*SSR*
C	0.002	0.001	-0.001	-2.32E-04	8.54E-05	1.28E-04
	[1.263]	[1.226]	[-0.501]	[-0.347]	[0.077]	[0.060]
Volatility	-4.752	-35.509	2.176	3.000	-2.964	-0.348
	[-1.210]	[-1.483]	[0.453]	[0.418]	[-0.052]	[-0.055]
Bad	0.001	0.001	-0.001	-0.008**	-0.003	-0.005
	[0.269]	[0.497]	[-0.165]	[-2.193]	[-1.325]	[-1.297]
Volatility*Bad	0.176	7.368	-17.909	-41.212	75.261	11.214
	[0.030]	[0.297]	[-0.861]	[-0.824]	[0.891]	[1.011]
Adjusted R^2	0.005	0.013	0.008	0.008	0.001	0.001
Panel A2 Realised returns volatility measures using 5-minutes intraday data						
C	0.003	0.001	-0.001	-0.002**	-1.30E-04	-3.66E-04
	[1.457]	[1.574]	[-0.577]	[-2.930]	[-0.123]	[-0.180]
Volatility	-6.229	-40.373*	2.344	24.156***	6.836	1.016
	[-1.415]	[-1.933]	[0.494]	[11.522]	[0.128]	[0.168]
Newlow	-0.024***	-0.005	0.006	0.007***	-0.003	0.000
	[-2.903]	[-1.324]	[0.407]	[2.906]	[-0.562]	[-0.025]
Volatility*Newlow	31.319***	125.732***	-52.735***	-66.110***	-130.289	-15.658
	[3.617]	[4.603]	[-3.262]	[-31.316]	[-0.813]	[-0.815]
Adjusted R^2	0.013	0.024	0.017	0.241	0.003	0.003

Table 5.12 (Continued)

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel B1 Historical volatility using daily data						
Volatility measure	*RW*	*V*	*RW*	*V*	*RW*	*V*
C	-0.002***	-0.001**	2.12E-04	3.66E-04	1.31E-05	1.99E-04
	[-3.103]	[-2.501]	[0.282]	[0.519]	[0.021]	[0.324]
Volatility	2.848***	2.235***	-0.551	-0.683	0.230	-0.502
	[3.969]	[2.594]	[-0.734]	[-1.026]	[0.076]	[-0.164]
Bad	0.005***	0.007***	-0.004*	-0.004	-0.003**	-0.002
	[3.215]	[4.028]	[-1.737]	[-1.611]	[-2.020]	[-1.340]
Volatility*Bad	-9.080***	-10.614***	-7.090*	-6.118**	8.751	4.870
	[-3.283]	[-5.084]	[-1.816]	[-2.29]	[1.326]	[0.575]
R-squared	0.058	0.046	0.023	0.030	0.005	0.001
Panel B2 Historical volatility using daily data						
C	-0.001***	-0.001**	1.37E-04	2.90E-04	-1.43E-04	1.05E-04
	[-2.808]	[-2.093]	[0.184]	[0.415]	[-0.242]	[0.176]
Volatility	2.516***	1.801**	-0.543	-0.675	0.863	-0.139
	[3.284]	[1.924]	[-0.722]	[-1.011]	[0.293]	[-0.047]
Newlow	-0.003	-0.001	-0.011***	-0.011***	0.000	0.002
	[-0.842]	[-0.334]	[-2.881]	[-2.998]	[0.126]	[0.687]
Volatility*Newlow	11.411***	12.933***	-9.788***	-7.604***	-33.545***	-35.225***
	[8.232]	[2.982]	[-10.653]	[-8.975]	[-2.613]	[-5.922]
Adjusted R^2	0.049	0.026	0.031	0.036	0.005	0.007
Panel C1 Conditional volatility measures using daily data						
Volatility measure	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*
C	-0.001	-0.001	-2.45E-04	-4.41E-04	-0.001	-0.001
	[-0.731]	[-1.035]	[-0.389]	[-0.755]	[-1.371]	[-1.179]
Volatility	0.724	1.479	0.178	0.445	4.836	4.274
	[0.568]	[0.883]	[0.391]	[1.220]	[1.624]	[1.388]
Bad	-0.003	-0.004*	-0.004	-0.003	-0.005	-0.005
	[-1.323]	[-1.764]	[-1.196]	[-1.003]	[-1.611]	[-1.596]
Volatility*Bad	5.858*	7.205**	-12.351**	-14.015**	14.863	13.007
	[1.683]	[2.168]	[-2.575]	[-2.403]	[1.263]	[1.205]
Adjusted R^2	0.003	0.005	0.009	0.010	0.005	0.004
Panel C2 Conditional volatility measures using daily data						
C	-0.001	-0.001	-3.02E-04	-0.001	-0.001*	-0.001
	[-0.835]	[-1.101]	[-0.487]	[-0.876]	[-1.684]	[-1.517]
Volatility	0.754	1.444	0.164	0.442	5.709*	5.261*
	[0.612]	[0.890]	[0.365]	[1.223]	[1.948]	[1.736]
Newlow	-0.003	-0.008	-0.022*	-0.012	-0.007	-0.007
	[-0.698]	[-1.564]	[-1.795]	[-1.201]	[-1.343]	[-1.393]
Volatility*Newlow	10.112**	13.740**	-13.643	-23.283*	0.859	0.183
	[2.679]	[2.458]	[-1.058]	[-1.667]	[0.060]	[0.015]
Adjusted R^2	0.004	0.007	0.012	0.015	0.005	0.005

Table 5.13 Sub-sample 2: Mean-variance Relationship during Period of Low News Sentiment
This table reports the slope coefficients regressions to test the Hypothesis 8, 9, and 10:

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \quad (14a)$$
$$R_{t+1} - F_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \quad (14b)$$

Where the dependent variable is daily excess index futures returns, of three markets, namely Hong Kong, Malaysia and Singapore. Excess return equals the daily return less risk free rate. R_t is the daily return of index futures returns. RF_t is the risk-free rate where Three-month treasury bill discount rate is used for Hong Kong and Singapore while one-month Kuala Lumpur interbank offer rate is used for Malaysia as the risk-free rate. Var_t is conditional variance while V_t is realised variance. *AAR*= Daily Average Absolute Returns Using 5-minutes Intraday Data ; *SSR*= Daily Average Absolute Returns using 5-minutes Intraday Data; *RW*= Daily Rolling Window Volatility; *V*= Daily Volatility; *GARCH-M*= Daily Conditional Volatility based on GARCH, *TGARCH-M*= Daily Conditional Volatility based on Threshold GARCH. *Sentiment* is dummy variable refers to sentiment measures namely *Bad and Newlow* . *Bad* = 1 if the daily routine news reports the market fell on the prior day; *Newlow*=1 if the market dipped to new low; otherwise=0. These sentiment measures apply to HSIF, KLCIF and SiMSCIF with two exceptions. For the case of HSIF, *Lowbench* is used instead of *Newlow*. *Lowbench* =1 if HSI fell to lower benchmark. The β_1 is expected to be positive following the capital asset pricing theory, or when create space effect is greater than Friedman effect; it is expected to be negative when Friedman effect is greater than create space effect. The α_2 is depend on the net impact of positive hold more effect and negative price pressure effect and it is expected to be negative during period of low sentiment. The $\beta_2$is expected to be positive during period of low sentiment The full sample period of the Hang Seng Index Futures (HSIF), and the Kuala Lumpur Composite Index Futures (KLCIF) cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Singapore Morgan Stanley Free Index Futures (SiMSCIF) from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008. The t values are reported in parentheses[ ]. *, ** and*** denotes the test statistics is significant at 10%, 5% and 1% respectively.

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel A1 Realised returns volatility measures using 5-minutes intraday data						
Volatility measure	*AAR*	*SSR*	*AAR*	*SSR*	*AAR*	*SSR*
C	0.004***	0.001***	0.004***	0.002***	0.002***	0.006***
	[3.158]	[2.624]	[3.609]	[3.412]	[4.397]	[5.137]
Volatility	-9.388***	-58.371**	-13.056***	-136.668***	-105.861***	-23.642***
	[-2.720]	[-2.416]	[-3.213]	[-3.072]	[-4.724]	[-4.51]
Bad	-0.005	-0.002	0.005	0.000	-0.002	-0.010**
	[-0.569]	[-0.448]	[0.699]	[0.004]	[-1.388]	[-2.476]
Volatility*Bad	12.131	76.781	-11.009	41.878	146.054***	37.969***
	[0.657]	[0.712]	[-0.803]	[0.816]	[2.835]	[2.699]
Adjusted R^2	0.018	0.022	0.026	0.035	0.070	0.070
Panel A2 Realised returns volatility measures using 5-minutes intraday data						
C	0.003**	0.001**	0.004***	0.002***	0.001***	0.005***
	[2.313]	[1.985]	[3.837]	[4.115]	[2.960]	[4.257]
Volatility	-6.312*	-32.646	-13.843***	-125.270***	-65.817***	-16.601***
	[-1.873]	[-1.519]	[-3.466]	[-4.140]	[-2.833]	[-3.656]
Newlow	0.045***	0.023***	-0.007	-0.008	-0.011**	-0.022**
	[2.872]	[2.590]	[-0.571]	[-0.650]	[-2.500]	[-2.434]
Volatility*Newlow	-51.390***	-215.869***	5.504	100.199*	245.072***	54.867***
	[-2.963]	[-4.173]	[0.449]	[1.698]	[4.317]	[2.654]
Adjusted R^2	0.030	0.033	0.025	0.035	0.040	0.043

Table 5.13 (Continued)

Index Futures	HSIF		KLCIF		SiMSCIF	
Dependent variable	Excess return	Excess return	Excess return	Excess return	Excess return	Excess return
Panel B1 Historical volatility using daily data						
Volatility measure	*RW*	*V*	*RW*	*V*	*RW*	*V*
C	-0.001	-2.07E-05	4.53E-04	0.001*	1.89E-04	0.001**
	[-1.209]	[-0.039]	[0.978]	[1.852]	[0.297]	[2.003]
Volatility	2.538	0.527	-2.578	-4.827	-1.960	-6.044*
	[1.428]	[0.246]	[-0.662]	[-1.388]	[-0.500]	[-1.878]
Bad	-0.001	0.000	0.006**	0.003	-0.002**	-0.002
	[-0.413]	[0.116]	[2.056]	[0.936]	[-1.966]	[-0.894]
Volatility*Bad	0.520	-0.143	-24.757***	-10.421	12.465**	12.266
	[0.157]	[-0.032]	[-4.166]	[-1.327]	[2.141]	[1.293]
Adjusted R^2	0.021	0.001	0.024	0.031	0.026	0.051
Panel B2 Historical volatility using daily data						
C	-0.001*	-2.50E-04	0.001	0.001**	3.35E-05	0.001**
	[-1.822]	[-0.549]	[1.094]	[1.985]	[0.062]	[2.231]
Volatility	3.369**	1.469	-3.106	-5.118	-0.269	-4.993*
	[2.515]	[0.828]	[-0.800]	[-1.531]	[-0.091]	[-1.871]
Newlow	0.022**	0.020**	0.004	0.005	-0.007**	-0.007**
	[2.101]	[2.126]	[0.517]	[0.670]	[-2.040]	[-2.110]
Volatility*Newlow	-19.828***	-13.771***	-18.783*	-15.202***	18.494***	21.931
	[-6.233]	[-5.633]	[-1.810]	[-3.450]	[3.047]	[1.234]
Adjusted R^2	0.060	0.037	0.014	0.031	0.004	0.036
Panel C1 Conditional volatility measures using daily data						
Volatility measure	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*	*GARCH-M*	*TGARCH-M*
C	4.81E-04	3.67E-04	-1.43E-04	-1.91E-04	0.001**	0.001**
	[0.877]	[0.631]	[-0.358]	[-0.495]	[2.401]	[2.196]
Volatility	-1.133	-0.765	1.456	1.720	-7.140**	-6.520**
	[-0.575]	[-0.361]	[0.640]	[0.807]	[-2.230]	[-2.069]
Bad	-0.005	-0.005*	0.007	0.005	-0.005*	-0.005*
	[-1.608]	[-1.654]	[0.921]	[0.766]	[-1.954]	[-1.862]
Volatility*Bad	11.567**	10.496**	-33.769	-26.414	20.134**	18.846**
	[2.048]	[2.185]	[-1.325]	[-1.281]	[2.164]	[2.077]
Adjusted R^2	0.007	0.008	0.005	0.005	0.028	0.027
Panel C2 Conditional volatility measures using daily data						
C	1.12E-04	-1.12E-04	-7.95E-05	-1.21E-04	0.001*	0.001
	[0.190]	[-0.177]	[-0.203]	[-0.321]	[1.677]	[1.482]
Volatility	0.233	0.956	0.893	1.122	-3.876	-3.345
	[0.110]	[0.421]	[0.395]	[0.530]	[-1.474]	[-1.314]
Newlow	0.024	0.024	0.007	0.008	-0.016***	-0.015***
	[1.440]	[1.447]	[0.557]	[0.595]	[-4.400]	[-4.333]
Volatility*Newlow	-37.584*	-31.749*	-48.025	-42.758	41.621***	36.019***
	[-1.867]	[-1.644]	[-1.058]	[-1.129]	[6.449]	[6.707]
Adjusted R^2	0.011	0.012	0.004	4701.741	0.013	0.012

5.6 CONCLUSION

This chapter shows that the investor sentiment improves the estimation of the mean-variance relationship. Specifically, the high news sentiment is proven to be better explaining the mean-variance trade-off than the low news sentiment, across six variance measures and across two sub-samples equal in time length. In addition, inclusion of the high sentiment regime to the base model has improve the coefficient of determination, R^2, while the equations for the low sentiment regime do not show improvement in R^2.

The high sentiment influences the index futures returns in two ways: directly priced as a form of sentiment risk (α_2) or indirectly affecting the excess returns through its interaction with risk measures (β_2) . The evidence shows that these two impacts do not necessary take effect simultaneously, that is, the α_2 *and* β_2 are significant at the same time. However, the finding suggests that if it does, the direct effect is partially offset by the indirect effect. In the high sentiment regime, the positive net impact on excess returns of the hold more effect and the price pressure effect are partially offset by the negative impact due to investors being less responsive to risks.

The sub-sample analysis reveals that the coefficients, β_1 and α_2 , carry opposite signs during different sample periods. This suggests that the Friedman effect, create space effect, hold-more effect and price pressure effect do affect the index futures returns differently over the sampling period. This could be due to macroeconomics factors that are not included in the model. However, the coefficient, β_2, are consistently negative in the high sentiment regime.

CHAPTER 6 : CONCLUSION

6.1 INTRODUCTION

I try to answer two questions in this thesis. First, what is the role of routine media content in explaining the daily index futures market behaviour? Second, what is the role of sentiment in explaining the index futures mean-variance relationship? These questions are motivated by the persistent mispricing of index futures and corroborative evidence on irrationality of market participants. Section 6.2 set out the review of research aims, hypotheses and major findings. Sections 6.3 and 6.4 discuss the contributions and limitations of the thesis. Lastly, I propose directions for future studies in section 6.5.

6.2 REVIEW OF RESEARCH AIMS, HYPOTHESES AND MAJOR FINDINGS

Noise traders are assumed to dominate the trading activity during periods of extreme sentiment. Based on this assumption, I argue that bad news factors generated from routine financial news predicts lower returns on the same and the next day of news release. The returns reverse on later days due to under-pricing. I derive trading strategies based on the findings to examine the profitability of the strategies. I also argue that if a sentiment regime affects the mean-variance relationship, the news sentiment is priced, therefore can mitigate mispricing if it can be included in the pricing model.

6.2.1 The role of media content in explaining the index futures returns and trading volume

This thesis aims to uncover the role of media content in explaining the index futures returns and trading volume. Table 6.1 summarises the hypothesis and findings.

Table 6.1 Summary of Hypotheses and Findings for Research Question One
This table illustrates the summary of Hypothesis 1 to Hypothesis 6. The findings are consistent among HSIF, KLCIF and SiMSCIF; and over sample periods.

	Hypothesis	HSIF	KLCIF	SiMSCIF
H1	Lag returns negatively forecast bad news factors	Supported	Supported	Supported
H2	Bad news factors negatively forecast returns	Supported	Supported	Supported
H3	There are returns reversals as predicted by noise or sentiment or stale information theory	Partially Supported	Partially Supported	Partially Supported
H4	There are partial returns reversals as predicted by noise and sentiment theory	Partially Supported	Partially Supported	Partially Supported
H5	There are full returns reversals as predicted by stale information theory	Not Supported	Not Supported	Not Supported
H6	Bad news factors negatively forecast trading volume if bad news factors proxy for trading cost	Partially Supported	Not Supported	Not Supported
H7	High absolute values of bad news factors forecast high trading volume if bad news factors proxy for investor sentiment	Not supported	Not supported	Not supported

Hypothesis 1 proposes that lag returns negatively forecast bad news factors. The evidence is strong and consistent among the HSIF, KLCIF and SiMSCIF; and over the split-half sample periods. For sub-sample 1, the lag returns are able to forecast bad news factors for up to three days. However, the forecast ability only lasts for two days in sub-sample 2. The findings of the full sample period are dominated by sub-sample 1. The impacts are the greatest on the next day (15 to 28 basis points) and gradually reduce to about 5 to 6 basis points, finally around 2 basis points on the third day. In conclusion, Hypothesis 1 is supported. Since I derived bad news factors from daily news that summarise the prior day stock market performance, this implies that the bad news factors consistently capture the market activities, and justifies its relevance in this study.

Hypothesis 2 proposes that bad news factors forecast returns. There are a few possible explanations. First, the bad news factors capture real information that is yet to be incorporated into stock prices. This is unlikely because the news is merely a summary of the prior days' trading activities. Second, the bad news factors capture the sentiment, noise or stale information. Hypotheses 3 to Hypothesis 5 confirm this. Evidence from the HSIF, KLCIF, and SiMSCIF support Hypothesis 2. The results are consistent over both sample periods. The negative impact of bad news on returns starts on the same day for all the three index futures contracts. The impact dies out on the same day for the HSIF and SiMSCIF, while the effect on KLCIF's returns dies out on the next day. One standard deviation increase in bad news factors predicts returns to drop by 4 to 16 basis points on the same day. The impact of bad news factors on KLCIF's returns is the greatest, up to 18 basis points (the same day plus the next day).

Taken together the conclusion for Hypothesis 1 and Hypothesis 2 is that there is a bidirectional causality between bad news factors and index futures returns. This is consistent with the positive feedback theory by Delong et al. (1990a). The low returns trigger high bad news factors, or vice versa.

Hypothesis 3 tests whether there are returns reversals, as predicted by noise, sentiment or stale information theory. This also rules out the possibility that bad news factors proxy for pure information that is yet to be incorporated into stock prices. According to the joint coefficient tests for the next two days to the next five days (second lag to fifth lag), only the HSIF demonstrates significant reversals. However, based on individual coefficients for the second lag to fifth lag, all the three futures contracts shows some evidence of reversal at different times. For the HSIF, the returns reversals begin from day two, and last up to the fourth day. The KLCIF's returns reversals only start to take effect on the fifth day. The SiMSCIF experience returns reversal only on the fourth day. The HSIF and SiMSCIF returns reversals is significant when *Pessimism* and *Weak* news factors are used in the SVAR, while the KLCIF returns reversal is significant when *Negative* news factor is used. Since there are inconsistent results over the timing of reversals and bad news factors, the evidence partially supports Hypothesis 3.

Despite the fact that the evidence does not unanimously support Hypothesis 3, Hypothesis 4 and Hypothesis 5 together, help to define the role of bad news factors. Hypothesis 4 proposes that bad news factors proxy for noise or investor sentiment. Hypothesis 5 suggests that bad news factors proxy for stale information. The HSIF's returns decrease on the bad news publication day. The reversals on later days exceed that initial negative change of returns. According to Tetlock (2007), this is evidence of bad news factors reflecting sentiment, instead of forecasting sentiment. There is only weak evidence on the KLCIF and SiMSCIF returns reversals, and the magnitudes of these reversals do not offset the initial negative change. There is some evidence to conclude that the bad news factors extracted from the New Straits Times and The Straits Times proxy for noise or sentiment. The conclusion made for Hypothesis 4 completely precludes Hypothesis 5.

Hypothesis 6 proposes that bad news factors negatively forecast trading volume if bad news factors proxy for trading cost. Evidence from the HSIF support Hypothesis 6 in that a one standard deviation increase in *Pessimism* and *Negative* news factors leads to around 0.2% decrease in open interest. This implies that investors have decided to close their positions based upon bad news and the market faces money outflow. There is no significant evidence from the KLCIF and SiMSCIF. There is no evidence to conclude that the bad news factors are able to predict contract volume.

Hypothesis 7 suggests that high absolute values of bad news factors forecast high trading volume if bad news factors proxy for investor sentiment. Although there is some evidence supporting this hypothesis, it is too weak to be conclusive. As for HSIF, a one standard deviation increase in the absolute value of *Pessimism* leads to a 0.343% decrease in contemporaneous open interest. There is a similar impact of absolute value of *Negative* on SiMSCIF's open interest. The absolute values of bad news factors significantly predict the next day's contract volume. Although there is some evidence supporting this hypothesis, it is too weak to be conclusive. In addition, the direction is the opposite of the prediction of sentiment theory.

Taken together, the negative predictability of bad news factors on index futures returns is consistent over contracts of different markets and sample periods. However, the bad news factors do not convincingly predict open interests and contract volumes.

The hypothetical trading strategies devised from the above findings generate returns up to 61.17, 44.17 and 17.2 basis points from trading the HSIF, KLCIF and SiMSCIF contracts. The estimated transaction cost is up to 7.2 basis points (for HSIF) and 16 basis points (for KLCIF). This strategy is profitable given the investors only pay margin rather than full contract value. The analysis of variance F tests reveal that the holding period and the ranking (intensity) of bad news significantly affect the returns generated from all three contracts. Shorting contracts when the bad news factors are ranked in the top two deciles generates the greatest returns. There is no significant difference in returns generated by following *Pessimism*, *Negative* or *Weak*. However, the interaction between news factors and year does affect the KLCIF and SiMSCIF returns.

In conclusion, the bad news factors reasonably capture the investor sentiment, over the contracts traded in both developed and emerging markets.

6.2.2 The role of investor sentiment in the returns mean-variance relationship

This study investigates whether investor sentiment is priced and has a role in the mean-variance relationship. Evidence of persistent mispricing of index futures and mixed findings on the mean-variance relationship motivate this study. I assume that investors are prone to cognitive biases instead of being fully rational. The cognitive biases lead to a biased expectation for risk and return among sentiment investors. There are more sentiment traders in the market and trade actively when the sentiment is relatively high or low.

Hypothesis 8 tests the relationship between the returns and volatility measures. The Capital Asset Pricing Theory implies a positive mean-variance relationship regardless of sentiment (Merton, 1973). DeLong et al. (1990b) noise traders theory suggests that the mean-variance relationship depends on the net impact of the create space effect (always positive regardless of sentiment) and the Friedman effect (always negative regardless of sentiment). There is a positive mean-variance relationship if the create space effect is greater than the Friedman effect, and vice versa.

Hypothesis 9 tests whether sentiment is priced. When sentiment is high, the net impact of the positive hold-more effect and negative price pressure effect determine the return-sentiment relationship (DeLong et al., 1990a). During times of low sentiment, the hold-more effect and the price-pressure effect are always negative.

Hypothesis 10 tests the sensitivity of return to risk as suggested by Yu and Yuan (2011). Noise traders are less responsive to risks during periods of high sentiment implying a flatter mean-variance relationship. Inversely, they are overly cautious during periods of low sentiment, implying a steeper mean-variance relationship.

Table 6.2 summarises the results of the high sentiment regime. The Panels A, B and C summarise the Hypotheses 8, 9 and 10 respectively. Table 6.3 summarises the findings on the low sentiment regime. The findings vary with the contract, measure of volatility and sample period. The KLCIF demonstrates the most consistent results, and the HSIF is the least consistent. The HSIF is trading in a more developed market with higher trading volume, there are less nonsynchronous trading problems, information travels faster, and therefore the noise traders are less likely to dominate the market activities. The results are more significant when the news regime is relatively extreme, i.e. *Newhigh* and *Newlow*. This is consistent with a study on agricultural futures contracts, that the extreme sentiment (top 20 decile) have stronger correlation with returns as opposed to above-median sentiment (Wang, 2001). Simon and Wiggins lll (2001) also find that when the sentiment is more extreme (top 10 decile), the impact of sentiment is greater over the longer period, that is, 30-day returns as compared to 10-day returns.

Referring to Table 6.2, the evidence from favourable news regimes on Hypothesis 8 and Hypothesis 9 are too weak to be conclusive, for the HSIF and SiMSCIF. The HSIF's findings on Hypothesis 8 are consistent over the sample periods, but irregular across measures of volatility. However, the KLCIF returns illustrate a consistent relationship with variance measures and news regimes. The KLCIF returns are positively correlated with variance in the *Good* and *Newhigh* regimes during sub-period 1 and negatively correlated with the variance in sub-period 2. The *Newhigh* sentiment is positively priced during sub-period 2, but not significantly priced during sub-period 1. In Panel C, the KLICF's mean-variance relationships are weakening in *Good* and *Newhigh* news regimes, dominating sub-sample 1. Although the evidences are not as strong, I conclude the same for the HSIF and SiMSCIF.

Panel A of Table 6.3 shows a similar pattern as in Panel A of Table 6.2, except for the equations from the KLCIF. This implies that both high and low news regimes reasonably capture the variation in the mean-variance relationship. Evidence from the HSIF and KLCIF are too weak to form meaningful conclusions for Hypothesis 8 and Hypothesis 9. However, the evidence from the SiMSCIF is somewhat convincing. Generally, there is a negative mean-variance relationship meaning that the Friedman effect outweighs the create space effect. The price pressure effect and hold-more effect negatively affect the excess returns in *Bad* and *Newlow* news regimes. In Panel C, evidence from the SiMSCI supports Hypothesis 10 that the investors are more responsive to risk in a low sentiment regime. This is also true for the HSIF in sub-period 1, but the results indicate that the mean-variance relationship is weaker during sub-period two, which is counter intuitive. Similarly, the findings from the KCLIF consistently show that investors are less responsive to risks in a low sentiment regime.

The mean-variance relationship remains inconclusive despite attempts to rule out the possible conditioning information. The different proxies used in the literature are mainly responsible for the mixed findings and need further investigation. Although there is mixed evidence on Hypothesis 8 and Hypothesis 9, I can conclude that the findings are in line with Hypothesis 10 that there is a weaker mean-variance relationship during periods of high-sentiment. The mixed evidence on Hypothesis 10 during periods of low-sentiment remains unexplained.

Table 6.2 Summary of Hypotheses and Findings for Research Question Two during the High Sentiment Regime

This table summarises the findings for Hypothesis 8 to Hypothesis 10. The rows compare the variation between sub-samples. The columns compare the variation among six measures of volatility. Which *AAR* = Average Absolute Returns; *SSR* = Sum of Squares Returns in 5-minutes interval ; *RW* = 30-day Moving Average Rolling Window Volatility ; *V*= Daily historical volatility; GM = Conditional volatility using *GARCH-M* model , and TGM = Conditional volatility using *TGARCH-M* model. The full sample period of the South China Morning Post, and The New Straits Times cover 1/1996 to 12/2008. The samples are then split half into sub-period 1 (1/1996 to 6/2002) and sub-period 2 (7/2002 to 12/2008). The full sample period for the Straits Times from 1/1999 through 12/2003, sub-period 1 from 1/1999 through 12/2003, sub-period 2 from 1/2004 through 12/2008.

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \quad (14a)$$
$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \quad (14b)$$

The table only includes coefficients that are at least significant at $\alpha = 0.10$. A positive (+) sign indicates the coefficient is greater than zero. A negative (-) sign indicates the coefficient is less than zero.

News Regime	Sample	HSIF						KLCIF						SiMSCIF					
		AAR	*SSR*	*RW*	*V*	GM	TGM	*AAR*	*SSR*	*RW*	*V*	GM	TGM	*AAR*	*SSR*	*RW*	*V*	GM	TGM
Panel A																			
Hypothesis 8																			
$\beta_1 > 0$ (positive mean-variance trade-off or create space effect outweigh the Friedman effect)																			
$\beta_1 < 0$ (Friedman effect outweigh the create space effect)																			
Good	Full	-	-	+	+				+	+	+				-				
Good	Sub1		-	+	+			+	+	+	+								
Good	Sub2	-	-	+				-	-		-			-	-				
Newhigh	Full	-	-	+	+			+	+	+	+	+	+	-	-				
Newhigh	Sub1		-	+	+			+	+	+	+	+	+					+	+
Newhigh	Sub2	-	-	+				-	-		-			-	-		-		
Panel B																			
Hypothesis 9																			
$\alpha_2 > 0$ (hold-more effect dominate the price pressure effect)																			
$\alpha_2 < 0$ (price pressure effect dominate the hold-more effect																			
Good	Full							+							+				
Good	Sub1							+								-	-		
Good	Sub2														+				
Newhigh	Full			+				-	+	+	+	+	+	+					
Newhigh	Sub1														+				
Newhigh	Sub2							+	+	+	+	+	+			-			
Panel C																			
Hypothesis 10																			
$\beta_2 < 0$ (investors are less responsive to risk in the high sentiment regime)																			
Good	Full							-	-	-	-	-		-	-		-		
Good	Sub1							-	-	-	-	-		-					
Good	Sub2													-	-				
Newhigh	Full			-	-	-	-		-	-	-	-	-	-					
Newhigh	Sub1			-	-	-	-	-	-	-	-	-	-						
Newhigh	Sub2							+						-	-			-	-

Table 6.3 Summary of Hypotheses and Findings for Research Question Two during the Low Sentiment Regime

This table summarises the findings for Hypothesis 8 to Hypothesis 10. The rows compare the inconsistencies between sub-samples. The columns compare the inconsistencies among six measures of volatility. Which *AAR* = Average Absolute Returns ; *SSR* = Sum of Squares Returns in 5-minutes interval ; *RW* = 30-day Moving Average Rolling Window Volatility ; *V*= Daily historical volatility; GM = Conditional volatility using *GARCH-M* model , and TGM = Conditional volatility using *TGARCH-M* model. The sample periods begin in January 1996 for Hang Seng Stocks Index Futures and Kuala Lumpur Stock Index Futures; in April January 1999 for Singapore Morgan Stanley Composite Index Futures. The sample period ends at December 2008 for each index futures.

$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 Var_t(R_{t+1}) + \alpha_2 Sentiment_t + \beta_2 Sentiment_t Var_t(R_{t+1}) + \varepsilon_{t+1} \qquad (14a)$$
$$R_{t+1} - RF_{t+1} = \alpha_1 + \beta_1 V_t + \alpha_2 Sentiment_t + \beta_2 Sentiment_t V_t + \varepsilon_{t+1} \qquad (14b)$$

The table only includes coefficients that are at least significant at $\alpha = 0.10$. A positive (+) sign indicates the coefficient is greater than zero. A negative (-) sign indicates the coefficient is less than zero.

News Regime	Sample	HSIF						KLCIF						SiMSCIF					
		AAR	*SSR*	*RW*	*V*	GM	TGM	*AAR*	*SSR*	*RW*	*V*	GM	TGM	*AAR*	*SSR*	*RW*	*V*	GM	TGM
Panel A Hypothesis 8 $\beta_1 > 0$ (positive mean-variance trade-off or create space effect outweigh the Friedman effect) $\beta_1 < 0$ (Friedman effect outweigh the create space effect)																			
Bad	Full	-	+	+	+									-	-				
Bad	Sub1			+	+														
Bad	Sub2	-	-					-	-					-	-		-	-	-
Newlow	Full	-	-	+	+									-	-				
Newlow	Sub1		-	+	+				+									+	+
Newlow	Sub2	-		+				-	-					-	-		-		
Panel B Hypothesis 9 $\alpha_2 < 0$ (The price pressure effect and the hold-more effect have negative impact on returns during low sentiment period)																			
Bad	Full				+	-	-							-	-	-		-	
Bad	Sub1			+	+		-		-	-						-			
Bad	Sub2						-			+					-	-		-	-
Newlow	Full			+	+									-	-	-		-	-
Newlow	Sub1	-							+	-	-	-							
Newlow	Sub2	+	+	+	+									-	-	-	-	-	-
Panel C Hypothesis 10 $\beta_2 > 0$ (investors are more responsive to risk in the low sentiment regime)																			
Bad	Full					+	+			-	-	-	-	+	+	+		+	
Bad	Sub1			-	-	+	+			-	-	-	-						
Bad	Sub2					+	+			-				+	+	+		+	+
Newlow	Full				-			-	-	-	-		-	+				+	+
Newlow	Sub1	+	+	+	+	+	+	-	-	-	-	-	-			-	-		
Newlow	Sub2	-	-	-	-	-	-		+	-	-			+	+	+		+	+

6.3 CONTRIBUTIONS OF THE THESIS

This thesis contributes to the derivatives asset pricing and investor sentiment literature. The findings have implications on index futures trading strategies and index futures pricing models.

Built on the method developed by Tetlock (2007), I include the contemporaneous term to examine the impact of news on the same trading day and expand the study to index futures contracts of different size and structure. This study finds that the bad news factors extracted from daily market news using General Inquirer and Harvard IV Psychosocial Dictionary consistently predicts the index futures returns across the HSIF, KLCIF and SiMSCIF. As evident in the tests of Hypothesis 4, this method is able to capture investor sentiment pertaining to local context on daily basis. This is more efficient than a sentiment survey, which is technically difficult and costly to conduct on a daily basis. This is particularly practical for emerging markets, where there are limited resources.

This thesis adds to literature of the index futures pricing formation process, by confirming the role of news content, as a proxy for investor sentiment. This rules out the role of news content, as proxy for information, stale information, or trading costs. The investor sentiment theory explains the findings of negative predictability of bad news factors over returns and the returns reversal phenomenon. The trading strategies devised, based on the findings, are proven to yield positive returns after taking into account trading costs. Since investors only pay margins instead of full settlement when trading index futures, the capital cost is much lower than trading in spot markets.

The literature review reveals that the investigation of the role of investor sentiment on derivatives is relatively limited as compared to the spot market. This study adds to the literature of derivatives market that Investors overreact to bad news, under-price the assets and cause returns reversals. The role of investor sentiment in the price discovery process is the same but the timing is different in cash and futures markets. Tetlock (2007) finds that the negative sentiment negatively predicts the next day's Dow Jones Index returns. I find that the bad news factors negatively predict the HSIF (same day), SiMSCIF (same day) and KLCIF (up to next three day) returns. These findings have implications for optimal timing strategies. On the day when the bad news factors are high, investors of HSIF and SiMSCIF initiate long contracts, while investors of KLCIF wait for another three days.

This thesis also provides some insight on the returns mean-variance relationship. A previous study investigates the NYSE-Amex stock index, using monthly data and three variance measures (Yu and Yuan, 2011). This study uses daily data, including daily realised variance estimated from 5-minutes intraday data, historical daily variance and conditional variance. Instead of using the sentiment index proposed by Baker and Wurgler (2006) to identify the high and low sentiment regime, this study uses the general stock market news. Since the news reports the prior day's market performance, it is less likely to contain new information. This makes it an appropriate sentiment measure. Inclusion of investor sentiment is insufficient to attain a robust positive mean-variance trade-off as suggested by rational expectations theory. I find that the sentiment is not directly priced, but it indirectly alters the mean-variance relationship.

This study establishes relationships among returns, sentiment and volatility. During periods of high sentiment, the noise traders are more confident and less responsive to risk; therefore weakening the mean-variance relationship. This implies more risk-taking behaviour. Investors are more responsive to risks during periods of low sentiment. They become risk-averse and demand for higher returns for bearing risks. These suggest that investor sentiment has a significant role in the index futures pricing models, consistent with Fung and Lam (2004) that pricing error is related to investor sentiment. The findings also have implications on risk-hedging practices. The hedgers who are more risk-averse should avoid executing trades during periods of high sentiment, because the compensation for bearing risk is lower.

6.4 LIMITATIONS OF THE THESIS

Some caveats on the measurement of sentiment must precede the interpretation of the results. First, I generate bad news factors using General Inquirer based on Harvard lV Psychosocial Dictionary. The dictionary is well developed and tested over time. Although General Inquirer is able to handle words according to syntax and lexicons, it cannot distinguish certain combinations of words. Second, I use principal component analysis to generate the *Pessimism* news factors. The choice of extraction method will affect the calculation of factor scores. Third, I use the human-score method to generate high and low sentiment regimes. Although I strictly follow the predetermined coding rules, there are slight discrepancies between the first coding and second coding. Despite our best efforts to minimise the error, other content analysis methods may generate different results.

I only include three index futures contracts in this study, because well established and tested word count software developed in English warrant common used of English in the sample market, to ensure the reliability of measurement of the media content. This permits meaningful comparison between markets. Therefore, I exclude Korea, Japan and China, although these markets are among the top in the region by contract volume.

Due to the limited predictive power of conditional volatility over realised returns, Lundblad (2007) suggests that analysis of the mean-variance relationship requires a very large sample, in order to produce meaningful inferences. There three selected index futures contracts are relatively new, thus limiting the sampling period. The data for the Hang Seng Index Futures and Kuala Lumpur Composite Index Futures spans 13 years, from 3rd January 1996 to 31st December 2008. The data for the Singapore Morgan Stanley Free Index Futures spans 11 years, from 3rd January 1998 to 31st December 2008. Although the mean-variance analysis in this study is plagued by small sample problems and produces mixed results, the results of the SVAR analyses are robust across sample periods, contracts and measures of sentiment.

6.5 DIRECTIONS FOR FUTURE RESEARCH

This study examines the impact of general investor sentiment on index futures' return mean and variance. A lack of a unanimous measure of investor sentiment remains a long-standing challenge in this area of research. Future studies can emphasise on the development of an investor sentiment classification algorithm, in order to improve the objectivity, reliability and validity of the sentiment measures.

In 2013, the New York Stock Exchange was the world's top stock exchange and CME Group was ranked first for index futures exchanges[37]. The past literature reports information spill over from US markets to the Asia-Pacific markets. The information disseminates faster in English because translation of news into local language delays the process. Researchers can conduct comparative studies on markets where the English language is not widely used, to examine whether there is a difference in the market-timing ability of news sentiment. This will shed light on sentiment measurement methods, specifically whether local language outperforms foreign language. The findings can be used to improve the validity of the content analysis.

[37] http://www.world-exchanges.org/files/statistics/pdf/2013_WFE_Market_Highlights.pdf

There are no assumptions made on the specific types of cognitive bias underlying the tests of the hypotheses in this study. We can expand the knowledge by researching how specific cognitive biases lead to overreaction and subsequent returns reversion as found in Chapter 4 of this thesis, and attenuated mean-variance relationships as found in chapter 5.

In Chapter 4, I find that the bad news factors negatively affect the KLCIF returns for up to three days, but only affects the HSIF and SiMSCIF returns on the same day. In Chapter 5, during low sentiment periods, I find that investors on the SiMSCIF are more responsive to risk, but the KLCIF investors are less responsive to risk. Since these contracts are different in size and liquidity levels affect price-discovery [38], future studies can extend this study by incorporating market size.

Another possible explanation is the role of hedgers and speculators. Wang (2001) finds hedger and speculator sentiment correlate with returns in different directions. Unfortunately, the trading data by type of traders was not available at the time of study. In 2009, 89% of those trading the HSIF are institutional traders [39]. In the same year, institutional traders only account for 29% of the KLCIF's contracts. Future studies can explore the role of investor sentiment by type of traders.

The fundamental information that will affect financial trading activities includes macroeconomics announcement and company specific news releases. Investors interpret information differently and not all investors adjust their beliefs accordingly, as suggested by the Theory of Subjective Expected Utility. Trading based investor sentiment measures are considered lagging measures because investors receive information and form their beliefs

[38] For example, the liquidity affects order imbalance, which in turn predicts the yield curve. See Brandt and Kavajecz (2004).
[39] https://www.hkex.com.hk/eng/stat/statrpt/factbook/factbook2009/Documents/31.pdf

before they trade. Hence, the future direction of investor sentiment research should emphasise on the automated text analysis because information is delivered in words, then interpreted by investors and finally reflected in their trading activities. The ability to quantify information contained in news has lead-time advantage over a sentiment measure proxy by trading data.

The main argument of the mean-variance relationship lies in the measurement of conditional variance as a proxy for the risk measure. The various ARCH class models, realised volatility measures and historical volatility measures are heavily investigated. These measures take into account both down-side and up-side of the deviations from expected returns. This is intuitively inappropriate because the risk measure should be measuring unfavourable outcomes. Further studies based on semi-variance may be able to give an answer to this issue.

REFERENCES

Amisano, G., & Giannini, C. (1997). *Topics in structural VAR econometrics* (2nd ed.). New York: Springer.

Andersen, T. G., & Bollerslev, T. (1998a). Deutsche Mark–Dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *The Journal of Finance, 53*(1), 219-265. doi: 10.1111/0022-1082.85732

Andersen, T. G., & Bollerslev, T. (1998b). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review, 39*(4), 885-905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics, 61*(1), 43-76.

Andersen, T. G., Bollerslev, T., & Lange, S. (1999). Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon. *Journal of Empirical Finance, 6*(5), 457-477. doi: http://dx.doi.org/10.1016/S0927-5398(99)00013-4

Andreassen, P. B., & Kraus, S. J. (1990). Judgmental extrapolation and the salience of change. *Journal of Forecasting, 9*(4), 347-372.

Antweiler, W., & Frank, M. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance, 59*(3), 1259-1294.

Backus, D. K., & Gregory, A. W. (1993). Theoretical relations between risk premiums and conditional variances. *Journal of Business & Economic Statistics, 11*(2), 177-185. doi: 10.1080/07350015.1993.10509946

Baillie, R. T., & DeGennaro, R. P. (1990). Stock returns and volatility. *The Journal of Financial and Quantitative Analysis, 25*(2), 203-214. doi: 10.2307/2330824

Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets, 7*(3), 271-299.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance, 61*, 1645-1680.

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

Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics, 104*(2), 272-287. doi: http://dx.doi.org/10.1016/j.jfineco.2011.11.002

Bali, T. G., Cakici, N., Yan, X., & Zhang, Z. H. E. (2005). Does idiosyncratic risk really matter? *The Journal of Finance, 60*(2), 905-929. doi: 10.1111/j.1540-6261.2005.00750.x

Bali, T. G., Demirtas, K. O., & Levy, H. (2009). Is there an intertemporal relation between downside risk and expected returns? *Journal of Financial & Quantitative Analysis, 44*(4), 883-909.

Bali, T. G., & Peng, L. (2006). Is there a risk–return trade-off? Evidence from high-frequency data. *Journal of Applied Econometrics, 21*(8), 1169-1198. doi: 10.1002/jae.911

Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *The Quarterly Journal of Economics, 116*(1), 1-53.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307-343.

Barberis, N., Shleifer, A., & Wurgler, J. (2005). Comovement. *Journal of Financial Economics, 75*(2), 283-317.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. In G. M. Constantinides, M. Harris & R. Stulz (Eds.), *Handbook of the Economics of Finance* (Vol. 1, pp. 1053-1128): Elsevier.

Beaumont, R., Daele, M., Frijns, B., Lehnert, T., & Muller, A. (2008). Investor sentiment, mutual fund flows and its impact on returns and volatility. *Managerial Finance, 34*(11), 772-784.

Bessembinder, H., & Seguin, P. J. (1992). Futures trading activity and stock price volatility. *The Journal of Finance, 47*(5), 2015-2034.

Bessembinder, H., & Seguin, P. J. (1993). Price volatility, trading volume, and market depth: Evidence from futures markets. *The Journal of Financial and Quantitative Analysis, 28*(1), 21-39.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177-181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307-327. doi: Doi: 10.1016/0304-4076(86)90063-1

Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, 116-131.

Bollerslev, T., Gibson, M., & Zhou, H. (2011). Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Journal of Econometrics, 160*(1), 235-245. doi: 10.1016/j.jeconom.2010.03.033

Bollerslev, T., & Zhou, H. (2006). Volatility puzzles: A simple framework for gauging return-volatility regressions. *Journal of Econometrics, 131*(1–2), 123-150. doi: 10.1016/j.jeconom.2005.01.006

Brailsford, T. J., & Cusack, A. J. (1997). A comparison of futures pricing models in a new market: The case of individual share futures. *Journal of Futures Markets, 17*(5), 515-541. doi: 10.1002/(sici)1096-9934(199708)17:5<515::aid-fut2>3.0.co;2-i

Brandt, M. W., & Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach. *Journal of Financial Economics, 72*(2), 217-257. doi: http://dx.doi.org/10.1016/j.jfineco.2002.06.001

Brandt, M. W., & Kavajecz, K. A. (2004). Price discovery in the U.S. Treasury market: The impact of orderflow and liquidity on the yield curve. *The Journal of Finance, 59*(6), 2623-2654. doi: 10.1111/j.1540-6261.2004.00711.x

Braun, P. A., Nelson, D. B., & Sunier, A. M. (1995). Good news, bad news, volatility, and betas. *The Journal of Finance, 50*(5), 1575-1603.

Brav, A., & Heaton, J. B. (2002). Competing theories of financial anomalies. *The Review of Financial Studies, 15*(2), 575-606.

Brooks, C. (1998). Predicting stock index volatility: Can market volume help? *Journal of Forecasting, 17*(1), 59-80. doi: 10.1002/(sici)1099-131x(199801)17:1<59::aid-for676>3.0.co;2-h

Brooks, C. (2008). *Introductory econometrics for finance* (2nd ed.). New York: Cambridge University Press.

Brown, G. W. (1999). Volatility, sentiment, and noise traders. *Financial Analysts Journal, 55*(2), 82-90.

Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance, 11*(1), 1-27.

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business, 78*(2), 405-440. doi: 10.1086/427633

Brown, S. J., Goetzmann, W. N., Hiraki, T., Shiraishi, N., & Watanabe, M. (2003). Investor sentiment in japanese and U.S. daily mutual fund flows. from SSRN http://ssrn.com/paper=375322

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics, 18*(2), 373-399. doi: http://dx.doi.org/10.1016/0304-405X(87)90045-6

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. *The Quarterly Journal of Economics, 108*(4), 905-939.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics, 31*(3), 281-318. doi: http://dx.doi.org/10.1016/0304-405X(92)90037-X

Campbell, J. Y., & Kyle, A. S. (1993). Smart money, noise trading and stock price behaviour. *The Review of Economic Studies, 60*(1), 1-34.

Carchano, Ó., & Pardo, Á. (2009). Rolling over stock index futures contracts. *Journal of Futures Markets, 29*(7), 684-694.

Chan, K. C., Fung, H. G., & Leung, W. K. (2004). Daily volatility behavior in Chinese futures markets. *Journal of International Financial Markets, Institutions and Money, 14*(5), 491-505.

Chan, K. C., Karolyi, G. A., & Stulz, R. (1992). Global financial markets and the risk premium on U.S. equity. *Journal of Financial Economics, 32*(2), 137-167. doi: http://dx.doi.org/10.1016/0304-405X(92)90016-Q

Chen, N. F., Kan, R., & Miller, M. H. (1993). Are the discounts on closed-end funds a sentiment index? *The Journal of Finance, 48*(2), 795-800.

Chen, X., & Ghysels, E. (2011). News—Good or Bad—and Its Impact on Volatility Predictions over Multiple Horizons. *Review of Financial Studies, 24*(1), 46.

Chopra, N., Lee, C. M. C., Shleifer, A., & Thaler, R. H. (1993). Yes, discounts on closed-end funds are a sentiment index. *The Journal of Finance, 48*(2), 801-808.

Chou, R. Y. (1988). Volatility persistence and stock valuations: Some empirical evidence using garch. *Journal of Applied Econometrics, 3*(4), 279-294. doi: 10.1002/jae.3950030404

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics, 10*(4), 407-432. doi: http://dx.doi.org/10.1016/0304-405X(82)90018-6

Chua, H. H. (2008, October 17). 1.44 million read ST. *The Straits Times*, from http://www.straitstimes.com/Breaking%2BNews/Singapore/Story/STIStory_291524.html

Ciner, C. (2006). Hedging or speculation in derivative markets: the case of energy futures contracts. *Applied Financial Economics Letters, 2*(3), 189 - 192.

Conrad, J., & Kaul, G. (1993). Long-term market overreaction or biases in computed returns? *The Journal of Finance, 48*(1), 39-63.

Cornell, B. (1985). Taxes and the pricing of stock index futures: Empirical results. *Journal of Futures Markets, 5*(1), 89-101. doi: 10.1002/fut.3990050109

Cornelli, F., Goldreich, D., & Ljungqvist, A. (2006). Investor sentiment and pre-IPO markets. *The Journal of Finance, 61*(3), 1187-1216.

Corsi, F., Mittnik, S., Pigorsch, C., & Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews, 27*(1-3), 46-78. doi: 10.1080/07474930701853616

Coval, J. D., & Shumway, T. (2001). Is sound just noise? *The Journal of Finance, 56*(5), 1887-1910.

Crystal, D. (1995). *The Cambridge encycopedia of the English language* (2nd ed.). Cambridge, UK: Cambridge University Press.

Daigler, R. T., & Wiley, M. K. (1999). The impact of trader type on the futures volatility-volume relation. *The Journal of Finance, 54*(6), 2297-2316.

Daníelsson, J., & Love, R. (2006). Feedback trading. *International Journal of Finance & Economics, 11*(1), 35-53.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance, 53*(6), 1839-1885.

Darrat, A. F., Gilley, O. W., Li, B., & Wu, Y. (2011). Revisiting the risk/return relations in the Asian Pacific markets: New evidence from alternative models. *Journal of Business Research, 64*(2), 199-206. doi: http://dx.doi.org/10.1016/j.jbusres.2010.02.008

Davis, K. (2005, May 12). More Malaysian reading the NST, *New Straits Times,* p. 4.

Day, T. E., & Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics, 52*(1–2), 267-287. doi: http://dx.doi.org/10.1016/0304-4076(92)90073-Z

De Bondt, W. F. M., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance, 40*(3), 793-805.

De Bondt, W. P. M. (1993). Betting on trends: Intuitive forecasts of financial risk and return. *International Journal of Forecasting, 9*(3), 355-371.

De Santis, G., & Imrohoroğlu, S. (1997). Stock returns and volatility in emerging financial markets. *Journal of International Money and Finance, 16*(4), 561-579. doi: http://dx.doi.org/10.1016/S0261-5606(97)00020-X

Deffner, G. (1986). Microcomputers as aids in Gottschalk-Gleser rating. *Psychiatry Research, 18*(2), 151-159.

DeLong, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990a). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance, 45*(2), 379-395.

DeLong, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990b). Noise trader risk in financial markets. *The Journal of Political Economy, 98*(4), 703-738.

Dewally, M. (2000). Internet investment advice: investing with a rock of salt. Retrieved from http://ssrn.com/paper=206089 doi:10.2139/ssrn.206089

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance, 1*(1), 83-106. doi: http://dx.doi.org/10.1016/0927-5398(93)90006-D

Donaldson, R. G., & Kim, H. Y. (1993). Price barriers in the Dow Jones Industrial Average. *The Journal of Financial and Quantitative Analysis, 28*(3), 313-330.

Duffee, G. R. (1995). Stock returns and volatility a firm-level analysis. *Journal of Financial Economics, 37*(3), 399-420. doi: http://dx.doi.org/10.1016/0304-405X(94)00801-7

Dumas, B., Kurshev, A., & Uppal, R. (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *The Journal of Finance, 64*(2), 579-629. doi: 10.1111/j.1540-6261.2009.01444.x

Durant, K. T. (2008). *Sentiment drift and its effect on the classification of web log posts.* (Ph.D.), Harvard University, United States, Massachusetts. Retrieved from http://proquest.umi.com/pqdweb?did=1459933561&Fmt=7&clientId=56581&RQT=309&VName=PQD

Ederington, L. H., & Guan, W. (2005). Forecasting volatility. *Journal of Futures Markets, 25*(5), 465-490. doi: 10.1002/fut.20146

Elton, E. J., Gruber, M. J., & Busse, J. A. (1998). Do investors care about sentiment? *The Journal of Business, 71*(4), 477-500.

Engle, R. F., & Lee, G. (1999). A long-run and short-run component model of stock return volatilit. In R. F. Engle & H. White (Eds.), *Cointegration, Causality, and Forecasting-A Festschrift in Honour of Clive W.J. Granger* (pp. 475-497). Oxford: Oxford University Press.

Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica: Journal of the Econometric Society*, 391-407.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56. doi: Doi: 10.1016/0304-405x(93)90023-5

Figlewski, S. (1984). Hedging performance and basis risk in stock index futures. *The Journal of Finance, 39*(3), 657-669.

Figlewski, S. (1997). Forecasting volatility. *Financial Markets, Institutions & Instruments, 6*(1), 1-88. doi: 10.1111/1468-0416.00009

Fisher, K. L., & Statman, M. (2000). Investor sentiment and stock returns. *Financial Analysts Journal, 56*(2), 16-23.

Frazzini, A., & Lamont, O. (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics, 88*(2), 299-322.

French, S., Schwert, G., & Stambaugh, R. (1987). Expected stock returns and volatility. *Journal of Financial Economics, 19*(1), 3-29.

Friedman, M. (Ed.). (1953). *Essays in positive economics* Chicago: University of Chicago Press. .

Frino, A., Kruk, J., & Lepone, A. (2007) The SPI 200 in the Asia-Pacific region: Comparisons of liquidity and Transactions Costs against other stock index futures. *Market Insights* (12 ed.). Sydney, Australia: Australian Securities Exchange.

Fung, A. K. W., & Lam, K. (2004). Overreaction of index futures in Hong Kong. *Journal of Empirical Finance, 11*(3), 331-351.

Fung, A. K. W., Mok, D. M. Y., & Lam, K. (2000). Intraday price reversals for index futures in the US and Hong Kong. *Journal of Banking & Finance, 24*(7), 1179-1201.

Fung, H.-G., & Patterson, G. A. (1999). The dynamic relationship of volatility, volume, and market depth in currency futures markets. *Journal of International Financial Markets, Institutions and Money, 9*(1), 33-59.

Fung, J., & Draper, P. (1999). Mispricing of index futures contracts and short sales constraints. *Journal of Futures Markets, 19*(6), 695-715.

Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *The Journal of Business, 53*(1), 67-78.

Gay, G., & Jung, D. (1999). A further look at transaction costs, short sale restrictions, and futures market efficiency: the case of Korean stock index futures. *Journal of Futures Markets, 19*(2), 153-174.

Gemmill, G., & Thomas, D. C. (2002). Noise trading, costly arbitrage, and asset prices: Evidence from closed-end funds. *The Journal of Finance, 57*(6), 2571-2594.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics, 76*(3), 509-548. doi: http://dx.doi.org/10.1016/j.jfineco.2004.03.008

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance, 48*(5), 1779-1801.

Glushkov, D. (2006). *Sentiment beta*. SSRN eLibrary. University of Texas. Austin. Retrieved from http://ssrn.com/paper=862444

Goetzmann, W. N., Massa, M., & Rouwenhorst, K. G. (2000). Behavioral factors in mutual fund flows. *Yale School of Management Working Papers*.

Gottschalk, L., Winget, C., & Gleser, G. (1969). *Manual of instructions for using the Gottschalk-Gleser content analysis scales: Anxiety, hostility, and social alienation--personal disorganization*. Berkeley: University of California Press.

Goyal, A., & Santa-Clara, P. (2003). Idiosyncratic risk matters! *The Journal of Finance, 58*(3), 975-1008. doi: 10.1111/1540-6261.00555

Grootveld, H., & Hallerbach, W. (1999). Variance vs downside risk: Is there really that much difference? *European Journal of Operational Research, 114*(2), 304-319. doi: http://dx.doi.org/10.1016/S0377-2217(98)00258-6

Gujarati, D. (2003). *Basic econometrics* (6th ed.). New York: McGraw-Hill.

Guo, H., & Neely, C. J. (2008). Investigating the intertemporal risk–return relation in international stock markets with the component GARCH model. *Economics Letters, 99*(2), 371-374. doi: http://dx.doi.org/10.1016/j.econlet.2007.09.001

Guo, H., & Whitelaw, R. F. (2006). Uncovering the risk–return relation in the stock market. *The Journal of Finance, 61*(3), 1433-1463. doi: 10.1111/j.1540-6261.2006.00877.x

Harlow, W. V., & Rao, R. K. (1989). Asset pricing in a generalized mean-lower partial moment framework: Theory and evidence. *Journal of Financial and Quantitative Analysis, 24*(3), 285-311.

Harrison, P., & Zhang, H. H. (1999). An Investigation of the Risk and Return Relation at Long Horizons. *Review of Economics and Statistics, 81*(3), 399-408. doi: 10.1162/003465399558337

Harvey, C. R. (2001). The specification of conditional expectations. *Journal of Empirical Finance, 8*(5), 573-637. doi: http://dx.doi.org/10.1016/S0927-5398(01)00036-6

Hassan, T., Mohamad, S., Ariff, M., & Nassir, A. M. (2007). Stock index futures prices and the Asian financial crisis. *International Review of Finance, 7*, 119-141.

Hemler, M. L., & Longstaff, F. A. (1991). General equilibrium stock index futures prices: Theory and empirical evidence. *Journal of Financial and Quantitative Analysis, 26*(03), 287-308. doi: doi:10.2307/2331208

Hibbert, A. M., Daigler, R. T., & Dupoyet, B. (2008). A behavioral explanation for the negative asymmetric return–volatility relation. *Journal of Banking & Finance, 32*(10), 2254-2266. doi: http://dx.doi.org/10.1016/j.jbankfin.2007.12.046

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance, 56*(4), 1533-1597.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance, 54*(6), 2143-2184.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65-91.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-291.

Kaminsky, G. L., & Schmukler, S. L. (1999). What triggers market jitters?: A chronicle of the Asian crisis. *Journal of International Money and Finance, 18*(4), 537-560.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *The Journal of Financial and Quantitative Analysis, 22*(1), 109-126.

Klüppelberg, C., Lindner, A., & Maller, R. (2004). A continuous-time GARCH process driven by a Lévy process: Stationarity and second-order behaviour. *Journal of Applied Probability, 41*(3), 601-622.

Kumar, A., & Lee, C. M. C. (2006). Retail investor sentiment and return comovements. *The Journal of Finance, 61*(5), 2451-2486.

Kurov, A. (2008). Investor sentiment, trading behavior and informational efficiency in index futures markets. *Financial Review, 43*(1), 107-127.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica, 53*(6), 1315-1335.

Lean, H. H., Smyth, R., & Wong, W. K. (2007). Revisiting calendar anomalies in Asian stock markets using a stochastic dominance approach. *Journal of Multinational Financial Management, 17*(2), 125-141.

Lee, C., Shleifer, A., & Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *The Journal of Finance, 46*(1), 75-109.

Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking and Finance, 26*(12), 2277-2299.

Lei, Q., & Wu, G. (2005). Time-varying informed and uninformed trading activities. *Journal of Financial Markets, 8*(2), 153-181.

Lei, Y. C. (2005). *The trading volume trend, investor sentiment, and stock returns.* (Ph.D.), Louisiana State University and Agricultural & Mechanical College, United States -- Louisiana. Retrieved from http://proquest.umi.com/pqdweb?did=953994361&Fmt=7&clientId=56581&RQT=309&VName=PQD

Lemmon, M., & Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *Review of Financial Studies, 19*(4), 1499-1529. doi: 10.1093/rfs/hhj038

Li, J. (2011). Volatility components, leverage effects, and the return–volatility relations. *Journal of Banking & Finance, 35*(6), 1530-1540. doi: http://dx.doi.org/10.1016/j.jbankfin.2010.10.026

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*(1), 13-37. doi: 10.2307/1924119

Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *The Review of Financial Studies, 3*(2), 175-205.

Ludvigson, S. C., & Ng, S. (2007). The empirical risk–return relation: A factor analysis approach. *Journal of Financial Economics, 83*(1), 171-222. doi: http://dx.doi.org/10.1016/j.jfineco.2005.12.002

Lundblad, C. (2007). The risk return tradeoff in the long run: 1836-2003. *Journal of Financial Economics, 85*(1), 123-150.

Luo, J. S., & Li, C. A. (2008). Futures market sentiment and institutional investor behavior in the spot market: The emerging market in Taiwan. *Emerging Markets Finance and Trade, 44*(2), 70-86.

Müller, G., Durand, R. B., & Maller, R. A. (2011). The risk–return tradeoff: A COGARCH analysis of Merton's hypothesis. *Journal of Empirical Finance, 18*(2), 306-320. doi: 10.1016/j.jempfin.2010.11.003

Matteo, T. D., Aste, T., & Dacorogna, M. M. (2005). Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development. *Journal of Banking & Finance, 29*(4), 827-851.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica, 41*(5), 867-887.

Mitra, L., Mitra, G., & Dibartolomeo, D. (2009). Equity portfolio risk estimation using market information and sentiment. *Quantitative Finance, 9*(8), 887-895.

Morris, R. (1994). Computerized content analysis in management research: A demonstration of advantages & limitations. *Journal of Management, 20*(4), 903-931.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*(4), 768-783. doi: 10.2307/1910098

Nair-Venugopal, S. (2001). The sociolinguistics of choice in Malaysian business settings. *International Journal of the Sociology of Language, 2001*(152), 21-52. doi: doi:10.1515/ijsl.2001.056

Neal, R., & Wheatley, S., M. (1998). Do measures of investor sentiment predict returns? *The Journal of Financial and Quantitative Analysis, 33*(4), 523-547.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

Oberlechner, T., & Hocking, S. (2004). Information sources, news, and rumors in financial markets: Insights into the foreign exchange market. *Journal of Economic Psychology, 25*(3), 407-424.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *The Journal of Business, 53*(1), 61-65. doi: 10.2307/2352357

Pontiff, J. (1997). Excess volatility and closed-end funds. *The American Economic Review, 87*(1), 155-169.

Poon, S. H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature, 41*(2), 478-539. doi: 10.1257/002205103765762743

Prast, H. M., & de Vor, M. P. H. (2005). Investor reactions to news: a cognitive dissonance analysis of the euro-dollar exchange rate. *European Journal of Political Economy, 21*(1), 115-141.

Purnanandam, A. K., & Swaminathan, B. (2004). Are IPOs really underpriced? *The Review of Financial Studies, 17*(3), 811-848.

Qiu, L. X., & Welch, I. (2006). Investor sentiment measures. Retrieved from http://ssrn.com/paper=589641

Ritter, J. R. (1991). The long-run performance of initial public offerings. *The Journal of Finance, 46*(1), 3-27.

Ritter, J. R. (2003). Behavioral finance. *Pacific-Basin Finance Journal, 11*(4), 429-437.

Ritter, J. R., & Welch, I. (2002). A review of IPO activity, pricing, and allocations. *The Journal of Finance, 57*(4), 1795-1828.

Rosenberg, S. D., Schnurr, P. P., & Oxman, T. E. (1990). Content analysis: A comparison of manual and computerized systems. *Journal of Personality Assessment, 54*(1), 298.

Rossi, A., & Timmermann, A. (2010). *What is the Shape of the Risk-Return Relation?* Paper presented at the AFA 2010 Atlanta Meetings Paper.

Samsell, D. P. (2007). *The effect of survey-based sentiment measures on the predictability and volatility of stock returns conditioned on the payout yield and issue yield.* (Ph.D.), Old Dominion University, United States -- Virginia. Retrieved from http://proquest.umi.com/pqdweb?did=1428847961&Fmt=7&clientId=56581&RQT=309&VName=PQD

Sanders, D. R., Irwin, S. H., & Leuthold, R. M. (2003). The theory of contrary opinion: A test using sentiment indices in futures markets. *Journal of Agribusiness, 21*(1), 39-64.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance, 44*(5), 1115-1153. doi: 10.1111/j.1540-6261.1989.tb02647.x

Scruggs, J. T. (1998). Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: A two-factor approach. *The Journal of Finance, 53*(2), 575-603. doi: 10.1111/0022-1082.235793

Sentana, E. (1995). Quadratic ARCH models. *The Review of Economic Studies, 62*(4), 639-661. doi: 10.2307/2298081

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425-442. doi: 10.1111/j.1540-6261.1964.tb02865.x

Shefrin, H. (2000). *Beyond greed and fear*: Harvard Business School Press Cambridge, MA.

Shefrin, H. (2008). *A behavioral approach to asset pricing* (2nd ed.). Burlington, USA: Academic Press.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review, 71*(3), 421-436.

Shiller, R. J. (1989). *Market volatility*. Cambridge, Mass. :: MIT Press.

Shiller, R. J. (2002). From efficient market theory to behavioral finance. Retrieved from http://ssrn.com/paper=349660

Shiller, R. J. (2005). *Irrational exuberance*. New Jersey: Princeton University Press.

Shiller, R. J., Fischer, S., & Friedman, B. M. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity, 1984*(2), 457-510.

Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *The Journal of Economic Perspectives, 4*(2), 19-33.

Sias, R. W., Starks, L. T., & Tinic, S. M. (2001). Is noise trader risk priced? *Journal of Financial Research, 24*(3), 311.

Siegel, J. J. (1992). Equity risk premia, corporate profit forecasts, and investor sentiment around the stock crash of October 1987. *The Journal of Business, 65*(4), 557-570.

Simon, D., & Wiggins III, R. (2001). S&P futures returns and contrary sentiment indicators. *Journal of Futures Markets, 21*(5), 447-462.

Solt, M., & Statman, M. (1988). How useful is the sentiment index? *Financial Analysts Journal*(September/October), 45-55.

Sortino, F. A., & Van Der Meer, R. (1991). Downside risk. *The Journal of Portfolio Management, 17*(4), 27-31.

Statman, M., Thorley, S., & Vorkink, K. (2006). Investor overconfidence and trading volume. *Review of Financial Studies, 19*(4), 1531-1565. doi: 10.1093/rfs/hhj032

Sutcliffe, C. (2006). *Stock index futures*. Aldershot: Ashgate Publishing, Ltd.

Swaminathan, B. (1996). Time-varying expected small firm returns and closed-end fund discounts. *The Review of Financial Studies, 9*(3), 845-887.

Tang, G. Y. N., & Shum, W. C. (2004). The risk–return relations in the Singapore stock market. *Pacific-Basin Finance Journal, 12*(2), 179-195. doi: http://dx.doi.org/10.1016/j.pacfin.2003.07.001

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance, 62*(3), 1139-1168.

Theodossiou, P., & Lee, U. (1995). Relationship between volatility and expected returns across international stock markets. *Journal of Business Finance & Accounting, 22*(2), 289-300. doi: 10.1111/j.1468-5957.1995.tb00685.x

Thomascrampton. (2008). Hong Kong Standard claims free strategy brought victory over South China Morning Post Retrieved November 15, 2008, from http://www.thomascrampton.com/media/hong-kong-standard-claims-free-strategy-brought-victory-over-south-china-morning-post/

Trochim, W. M. K. (2001). *Research methods knowledge base* (2nd ed.). Cincinnati, Ohio Atomic Dog Pub.

Tumarkin, R., & Whitelaw, R. (2001). News or noise? Internet postings and stock prices. *Financial Analysts Journal, 57*(3), 41-51.

Turner, C. M., Startz, R., & Nelson, C. R. (1989). A Markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics, 25*(1), 3-22. doi: http://dx.doi.org/10.1016/0304-405X(89)90094-9

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Verma, R., Baklaci, H., & Soydemir, G. (2008). The impact of rational and irrational sentiments of individual and institutional investors on DJIA and S&P500 index returns. *Applied Financial Economics, 18*, 1303-1317.

Veronesi, P. (1999). Stock market overreactions to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies, 12*(5), 975.

Wang, C. (2001). Investor sentiment and return predictability in agricultural futures markets. *Journal of Futures Markets, 21*(10), 929-952.

Wang, C. (2003a). The behavior and performance of major types of futures traders. *Journal of Futures Markets, 23*(1), 1-31.

Wang, C. (2003b). Investor sentiment, market timing, and futures returns. *Applied Financial Economics, 13*(12), 871-878.

Wang, C. (2004). Futures trading activity and predictable foreign exchange market movements. *Journal of Banking & Finance, 28*(5), 1023-1041.

Wang, J. (2007). Testing the general equilibrium model of stock index futures: Evidence from the Asian crisiss. *International Research Journal of Finance and Economics, 10*, 107-116.

Wang, J. (2009). Stock market volatility and the forecasting performance of stock index futures. *Journal of Forecasting, 28*(4), 277-292.

Wang, J. c. (2008). Degree of market imperfections: Evidence from four Asian index futures markets. *Applied Financial Economics, 18*(15), 1233 - 1246.

Wang, Y. H., Keswani, A., & Taylor, S. J. (2006). The relationships between sentiment, returns and volatility. *International Journal of Forecasting, 22*(1), 109-123.

Warther, V. A. (1995). Aggregate mutual fund flows and security returns. *Journal of Financial Economics, 39*(2-3), 209-235.

Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Newbury Park, CA.: Sage.

Weinstein, N. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39*(5), 806-820.

West, K. D., & Cho, D. (1995). The predictive ability of several models of exchange rate volatility. *Journal of Econometrics, 69*(2), 367-391. doi: http://dx.doi.org/10.1016/0304-4076(94)01654-I

Whitelaw, R. F. (1994). Time variations and covariations in the expectation and volatility of stock market returns. *The Journal of Finance, 49*(2), 515-541. doi: 10.1111/j.1540-6261.1994.tb05150.x

Wongchoti, U., Wu, F., & Young, M. (2009). Buy and sell dynamics following high market returns: Evidence from China. *International Review of Financial Analysis, 18*(1-2), 12-20.

Wysocki, P. (1999). Cheap talk on the Web: The determinants of postings on internet stock message boards. *University of Michigan Unpublished Manuscript*.

Yadav, P. K., & Pope, P. F. (1994). Stock index futures mispricing: profit opportunities or risk premia? *Journal of Banking & Finance, 18*(5), 921-953. doi: Doi: 10.1016/0378-4266(94)00026-3

Yang, A. S., & Wu, M. L. (2010). Exploring the relationship between investor sentiment and price volatility. *Quantitative Finance, 11*(6), 955-965.

Yen, G., Lee, C. F., Chen, C. L., & Lin, W. C. (2001). On the Chinese Lunar New Year effect in six Asian stock markets: An empirical analysis *Review of Pacific Basin Financial Markets and Policies, 4*(4), 463-478.

Yu, J., & Yuan, Y. (2011). Investor sentiment and the mean-variance relation. *Journal of Financial Economics, 100*(2), 367-381. doi: 10.1016/j.jfineco.2010.10.011

Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control, 18*(5), 931-955. doi: http://dx.doi.org/10.1016/0165-1889(94)90039-6

Zweig, M. E. (1973). An investor expectations stock price predictive model using closed-end fund premiums. *The Journal of Finance, 28*(1), 67-78.

APPENDICES

APPENDIX A: INDEX FUTURES CONTRACT SPECIFICATIONS

	Hong Kong	Malaysia	Singapore
Date started	May 1986	December 1995	First published in May 1993 and with historical data going back to January 1988
Underlying instrument	Hang Seng Index	Kuala Lumpur Composite Index	Morgan Stanley Singapore Free (SiMSCI) Index
Contract code	HSI	FKLI	SG
Contract multiplier	HK$50 per index point	RM50 per index point	S$200 per index point
Minimum fluctuation	One index point (HK$50)	0.5 index point (RM25)	0.1 index point (S$20)
Margin	Initial HK$72500/lot Maintenance HK$58000/lot	Initial RM3500/lot Maintenance RM3500/lot	Initial SGD2640/lot Maintenance SGD 2400/lot
Contract months	Spot, next calendar month and next two calendar quarter months	Spot month, next calendar month, next two calendar quarterly months. The calendar quarterly months are March, June, September and December.	2 nearest serial months and four quarterly contract months on March, June, September and December cycle
Trading hours for underlying instrument	10.00 am – 12.30 pm and 2.30 pm – 4.00 pm	9.00 am – 12.30 pm and 2.30 pm – 5.00 pm	9.00 am – 12.30 pm and 2.00 pm – 5.00 pm
Contract Trading hours	Pre-opening 8:45 a.m. – 9:15 a.m. and 12:30 p.m. – 1:00 p.m. 9.15 am – 12.00 pm and 1.00 pm – 4.15 pm 5.00 pm – 11.00pm The trading hours on eves of Christmas, New year and Chinese New Year shall be 9.45 am – 1.00 pm.	Pre-opening 8.15am-8.45am and 2.00pm-2.30pm 8.45 am – 12.45 pm and 2.30 pm – 5.15 pm	Pre-opening 8.15am -8.28 am non-cancel period 8.28 am-8.30 am 8.45 am – 12.35 pm and 2.00 pm – 5.15 pm
Trading hours on last trading day	Expiring contract month closes at 4.00 pm on the last trading day	8.45 am – 12.45 pm and 2.30 pm – 5.15 pm	8.45 am – 12.35 pm and 2.00 pm – 5.15 pm
Last trading day	The business day immediately preceding the last business day of the contract month	The last business day of the contract month.	The second last business day of the contract month
Final settlement price	Average of quotations of the Hang Seng Index taken at five minute intervals during the last trading day	The final settlement value, shall be the average value, rounded to the nearest 0.5 of an index point (values of 0.25 and 0.75 and above being rounded upwards) of the KLCI for the last half hour of trading on Bursa Malaysia Securities Berhad on the final trading day excepting	The final settlement price shall be the value of the SiMSCI computed based on the Special quotation methodology applied on each component stock of the SiMSCI on the day following the last trading day.

		the highest and lolst values.	
Daily price limit	5% price limit up/down from the last traded price.	20% per trading session for the respective contract months except the spot month contract. There will be no price limit for the second month contract for the final five business days before expiration.	Whenever the price moves by 15% in either direction from the previous day's settlement price, trading at or within the price limit of +/- 15% is allowed for the next 10 minutes. After this "cooling off" period had elapsed, there shall be no price limit for the remainder of the trading day. There shall be no price limit on the last trading day of the expiring contract month.
Position limit	10,000 position delta combined in all Contract Months.	10000 contracts, net gross open position	A person shall not own or control more than 10,000 contracts net long or net short in all contract months combined.

Source:

1. https://www.hkex.com.hk/eng/prod/drprod/hkifo/fut.htm
2. http://www.sgx.com/wps/portal/sgxweb/home/products/derivatives/equity/sgx_simsci
3. http://www.bursamalaysia.com/market/derivatives/products/equity-derivatives/ftse-bursa-malaysia-klci-futures-fkli/

APPENDIX B: MEASURES OF INVESTOR SENTIMENT IN SPOT MARKET

B1 Subjective measures of investor sentiment for the spot market based on surveys

Survey is an important indicator of investor sentiment for market practitioners and academic scholars. Yale School of Management constructs a Stock Market Confidence Index, led by Robert Shiller since 1989. The research team constructs four confidence indices from the responses of high-income individuals and institutional investors. The one-year Index indicates the percentage of the population who expects Dow to increase in the subsequent year. The Buy on Dips Index describes the percentage of the population believing that the market will recover after dropping by 3% the previous day. Crash Index measures the percentage of the population thinking that the market has little chance to crash in the coming six months. Finally, Valuation Index shows the percentage of the population who perceive that market value is not too high. The team ensures the reliability of the indexes through a rigorous sampling design. Individual sample is drawn from a list of “High-Grade Multi-Investors” from 1989 to 1999. From 1999 onwards, a random sample of high-income Americans is purchased from Survey Sampling Inc. A random sample of institutional investors is drawn from Money Market Directory and their fund managers. They maintain the sample size at about 100 for the survey in each period for the two respondent categories.

Michigan Consumer Confidence Index (MCCI) was constructed in 1946 through a research conducted by Survey Research Centre, University of Michigan. The survey produces a proven indicator of the US future economic conditions. The Leading Indicator Composite Index is published by the US Department of Commerce includes MCCI. The survey questions are designed to reflect consumers’ views of their future personal financial prospect, near term economic conditions and long-term economic forecasts. At least 500 respondents are selected each month and the survey is deemed to be representative of the US population. The survey adopts the rotating panel design, where, 40% of the respondents are reselected six months later. This enables changes in responses to be measured. Qiu and Welch (2006) propose MCCI as an alternative proxy of sentiment: "…consumer confidence seems to be a concept similar to investor sentiment. Many investors are likely to be bullish about the economy when they are bullish about the stock market and vice-versa."

The UBS and the GALLUP Organisation compile the UBS Index of Investor Optimism. The survey covers the United States and Europe. Each survey consists of one thousand head of households, who are holding a portfolio of more than ten thousand dollars. The survey is carried out in the first two weeks of each month, and the report is published on the last Monday of the month. Similar to the MCCI, the survey questions cover expectation on personal and macro economy, for example:

"Overall, how optimistic are you that you will be able to achieve your investment targets over the next twelve months?"

"Thinking about your own household and the things that impact on your ability to invest over the next twelve months, how would you rate your ability to maintain or increase your current income over the next twelve months?"

The American Association of Individual Investor (AAII) has been conducting individual investor sentiment surveys since July 1987. The AAII has 150,000 members as at 2009. All members can respond to the weekly survey that is published on the AAII website every Thursday. However, the response rate of the survey is unknown. Members are asked for their opinion about the direction of the stock market over the next six month. The questions are closed-ended. The respondents are subject to three options: 1) Up, indicating bullish sentiment, 2) No change, indicating neutral point of view and 3) Down, indicating bearish sentiment. De Bondt (De Bondt, 1993) uses the AAII survey as a proxy for sentiment and finds that individual investors assume continuation in past price movements. Fisher and Statman (2000) study the impact of sentiment to stock returns. They use the monthly percentage of bullish investors as the proxy sentiment to examine the relationship between large-capitalisation stocks (proxy by S&P Index) and small capitalisation stocks (proxy by CRSP 9-10 Index). Brown and Cliff (2004) use the bull-bear spread from the AAII survey as one of the proxies for amateurs' sentiment in the study. Brown and Cliff (2004) and Verma, Baklaci, and Soydemir (2008) conclude that the AAII survey measures individual sentiment and the Investor Intelligence (see section 2.3.6) measures institutional sentiment.

B2 Objective measures of investor sentiment for the spot market based on trading data

Hardy (1939), Zweig (1973) and Malkiel (1997) (as cited in Samsell, 2007) are among the earliest to attempt to use trading data as a proxy for investor sentiment. Zweig (1973) assumes that there are two categories of investors. First, the non-professionals with high marginal cost, who are less clear about the future direction of their investments. Second, the professionals who enjoy low marginal cost and are clear about the fair value of their investment. The closed-end fund market is dominated by non-professionals or small investors, as professionals are not motivated to trade in this market segment. While fair value of most equities is hard to measure, closed-end funds are an exception. Fair price of a close-end fund is simply the net asset value of the securities held by the fund. When non-professionals are bullish (bearish) and bid the closed-end fund share price over (under) its net assets value, the fund is traded at premium (discount). The share prices of the closed-end funds are solely determined by non-professionals, implying that closed-end fund discounts (premiums) are a proxy for investor sentiment for non-professionals i.e. small investors (Zweig, 1973). Lee, Shleifer and Thaler (1991) reach the same conclusion but this is proven to be controversial in later studies (see Swaminathan (1996) and Neal and Wheatley (1998)). The paper concludes that closed-end funds are traded at discounts most of the time as unpredictable investor sentiment introduces risk for holding closed-end fund shares, and the risk is priced by imposing discounts to closed-end fund prices. These two papers confirm the role of closed-end fund discounts as proxy for small investor sentiment. Neal and Wheatley (1998) are convinced that closed-end fund discounts can predict the size premium and the Small-minus-Big (SMB) returns. Gemmill and Thomas (2002) find that discount fluctuation is related to small trader sentiment. However, the size of discount is related to managerial cost and cost of arbitrage in the long run. Arbitrage becomes expensive due to limits to arbitrage and high fees charged by fund managers is channelled to investor in the form of discounts.

Not all studies conclude the same. Chen, Kan and Miller (1993) claim that the findings in Lee, Shleifer and Thaler (1991) are highly sensitive to the particular sampling period, thus they reject the above conclusion by challenging the validity of the sample. Chopra, Lee, Shleifer and Thaler (1993) using the same data set as Chen et al. (1993), once again find consistency with Lee et al. (1991). However, Elton, Gruber and Busse (1998) find the opposite. The study concludes that the investor sentiment proxy given by changes in closed-end fund discounts is not an important factor in the returns generating process, and the risk of hard-to-predict sentiment is not priced in common stocks. Swaminathan (1996) argues that closed-end fund discounts is not a pure sentiment measure because it reflects rational expectations of future market fundamentals as it forecasts future earnings growth rates and inflation. However, the author does not completely reject the possibility of sentiment related explanations.

Kumar and Lee (2006) add another possible proxy for retail traders' sentiment. They use buy-sell imbalance (BSI) to proxy retail investor sentiment. The BSI is positive when there is more buying than selling. The increase in retail demand implies that investors are bullish. The result confirms the relationship between small trader sentiment and the market excess returns. This conclusion is consistent with Lee, Shleifer and Thaler (1991), despite the use of a different proxy.

Mutual fund flows are believed to be a sentiment measure that affects resource allocation to mutual fund and becomes a factor in the price discovery process. Warther (1995) divides mutual fund flows into expected fund flows and unexpected fund flows. Only unexpected inflow to mutual funds is positively related to future returns. The return-inflow relationship is category specific i.e. bond returns and inflow to bond funds. Brown, Goetzmann, Hiraki, Shiraishi and Watanabe (2003) assert that if investors are bullish about the future market prospect, assets are supposed to be allocated to funds that perform well in a rising market; if the sentiment is bearish more wealth will flow into funds that make a profit out of falling markets. Evidence from Japan documents flows to “Bull” or “Bear” fund is negatively related. Investors in the US view domestic funds and foreign funds as substitutes. When the US market is bad, investors move their investment to foreign funds. It is natural to conclude that allocation of funds to different fund categories reflects overall market sentiment. Frazzini and Lamont (2008) derive investor sentiment from mutual fund flows with a precise definition: “actual ownership by mutual funds minus the ownership that would have occurred if every fund had received identical proportional inflows, every fund manager chooses the same portfolio weights in different stocks as he actually did, and stock prices were the same as they actually were”. The evidence suggests that high sentiment leads to misallocation of assets to those funds already overpriced, leading to low subsequent returns. The conclusion of high equity issuance follows demand induced by high sentiment is consistent with Baker and Stein (2004).

Baker and Stein (2004) posit that market liquidity can be a sentiment indicator if short sales constraints exist. The model assumes irrational investors over-weight the private signals they received and under-weight the decision made by other investors. Positive market signals invoke positive sentiment among irrational traders. Short-sales constraints keep them away from the market unless they value an asset higher than its fundamental value. Increasing liquidity measures is an indication of irrational investors swayed by positive sentiment present in the market. The study proves that liquidity measures, i.e. equity issuance and share turnover predict future returns of a CRSP equal-weighted portfolio. Lei (2005) uses trading volume trends as a measure of investor sentiment as a solution to the non-stationary problem of trading volume series. High trading volume tends to predict negative expected stock return.

Anomalies in the initial public offering (IPO) market motivate researchers to study the IPO puzzle. Ritter (1991) observes that the IPO offer price is not too high, but that the after-market price is way above its fair value. In the long-run, share prices reverse to fundamental values and under-perform a comparable portfolio in the long-run. Firms that issue IPOs in heavy-volume years face a worse situation. Ritter and Welch (2002) advocate non-rational and agency cost explanations for these phenomena. Purnanandam and Swaminathan (2004) document that IPOs are over-priced at offer, the price is exceptionally high in after-market and reverts to fundamental value in the long-run. Despite disparity in opinion about the IPOs offer prices, the findings are consistent with Ritter (1991). The finding of overvalued IPOs outperforming undervalued IPOs in the first day going public is inconsistent with asymmetric information theories. Small traders actively trade in grey market (pre-IPO) to speculate on stocks that are about to go public. Cornelli, Goldreich and Ljungqvist (2006) explain anomalies in IPO pricing in relation to the small trader sentiment. Overly optimistic small traders bid the price above fair value in grey market, and this becomes a good predictor of price in the after-market. Underwriters and book building investors take advantage of irrational traders by selling overpriced shares to the latter. Price reversals follow and consequently the long-term returns are low. The combination of these finding relates IPO pricing, volume, IPOs after-market returns and long run returns to investor sentiment. Baker and Wurgler (2007) include IPO volume and IPO first day returns in the creation of a composite sentiment index.

APPENDIX C: TRADING STRATEGIES

I assume that the market participants can borrow at the risk free rate, facing no restriction of access, and ignore the trading margin. This zero-cost trading strategy is far from the real practice of the finance world. However, the simulation result provides some insights on the feasibility of the trading strategy. I devise the trading strategies conditional upon the level of bad news factor, whether it is high or low. There are two feasible market-timing strategies. I assume the negative impact on of the bad news lasts up to four days. Appendix B provides examples of these strategies, considering the price of a long position, the price of a short position and price at maturity. The examples show that the overall profit is the same for the two strategies regardless of the price on the maturity date. It also shows the total loss if the index futures price goes against the prediction based on the news factors.

Example 1 : Trading strategies based on high bad news factors

When the level of the bad news factors are high, the price level decreases on the same day, and the price is expected to reverse (become higher) on the fourth day. Assuming the price at maturity is higher than the price of the long position:

		Strategy 1: buy a long contract on day zero, buy a short contract on day 4 and wait until the contracts mature.		Strategy 2: buy a long contract on day one and sell it on day 4
Day	Price	Long	Short	Long contract
0	100	100		100
4	110		110	110
Maturity	120	120	120	
profit		20	-10	10
overall profit		10		10

All else equal, now consider if the maturity price is lower than the price of the long position:

Day	Price	Strategy 1: buy a long contract on day zero, buy a short contract on day 4 and wait until the contracts mature.		Strategy 2: buy a long contract on day one and sell it on day 4
		Long	Short	Long contract
0	100	100		100
4	110		110	110
Maturity	90	90	90	
profit		-10	20	10
Overall profit		10		10

These two tables show that the overall profit is the same for the two strategies regardless of the price on the maturity date.

Example 2: Trading strategies based on low bad news factors

When the level of the bad news factors are low, the price level increases on the same day, and it is expected to be reversed (become lower) on the fourth day.

Day	Price	Strategy 1: buy a short contract on day zero, buy a long contract on day 4 and wait until the contracts mature.		Strategy 2: buy a short contract on day one and sell it on day 4
		short	long	Long contract
0	110	110		100
4	100		100	110
Maturity	90	90	90	
profit		20	-10	10
overall profit		10		10

All else equal, now consider if the maturity price is lower than the price of the long position:

Day	Price	strategy 1: buy a short contract on day zero , buy a long contract on day 4 and wait until the contracts mature.		Strategy 2: buy a short contract on day one and sell it on day 4
		short	long	Long contract
0	110	110		100
4	100		100	110
Maturity	120	120	120	
profit		-10	20	10
Overall profit		10		10

These two tables show that the overall profit is the same for the two strategies regardless the price on the maturity date.

Example 3: Trading strategies based on high bad news factors, but the price does not reverse to a higher level

When the level of the bad news factors are high, the price level decreases on the same day, and the price is expected to reverse (become higher) on the fourth day. However, price goes against the expectation. Assuming the maturity price is higher than the price of a long position:

		Strategy 1: buy a long contract on day zero, buy a short contract on day 4 and wait until the contracts mature.		Strategy 2: buy a long contract on day one and sell it on day 4
Day	Price	Long	Short	Long contract
0	100	100		100
4	90		90	90
Maturity	120	120	120	
profit		20	-30	-10
Overall profit		-10		-10

All else equal, now consider if the maturity price is lower than the price of a long position:

		strategy 1: buy a long contract on day zero, buy a short contract on day 4 and wait until the contracts mature.		Strategy 2: buy a long contract on day one and sell it on day 4
Day	Price	Long	Short	Long contract
0	100	100		100
4	90		90	90
Maturity	90	90	90	
profit		-10	0	-10
Overall profit		-10		-10