MEDIA@LSE
Department of Media and Communications

MEDIA@LSE MSc Dissertation Series

Compiled by Bart Cammaerts, Nick Anstead and Ruth Garland

The Impact of Mass Media Sentiments on Returns and Volatility in Asset Markets:
Evidence from Algorithmic Content Analysis

Panu Kuuluvainen,
MSc in Media, Communication and Development

Dissertation submitted to the Department of Media and Communications, London School of Economics and Political Science, August 2014, in partial fulfilment of the requirements for the MSc in Media, Communication and Development. Supervised by Professor Robin Mansell

The Author can be contacted at: panu.kuuluvainen@gmail.com

Published by Media@LSE, London School of Economics and Political Science ("LSE"), Houghton Street, London WC2A 2AE. The LSE is a School of the University of London. It is a Charity and is incorporated in England as a company limited by guarantee under the Companies Act (Reg number 70527).

The Impact of Mass Media Sentiments on Returns and Volatility in Asset Markets

Evidence from Algorithmic Content Analysis

Panu Kuuluvainen

ABSTRACT

This dissertation investigates the effect of financial news on the decision-making and expectations of individual investors as evidenced by stock market volatility and returns. It is positioned in the juncture between investors, markets and mass media. The efficient market hypothesis is juxtaposed with the notion of performativity. It is argued that financial news, like economic theory, affects reality rather than simply describes what exists. Financial news is argued to be performative if the media disseminating information meet certain conditions regarding its material availability, circulation and authority. It is concluded that if the examined media meet these conditions, the message they transmit becomes a reality through the collective actions of individual investors. Performing an algorithmic content analysis of more than 7000 financial news articles, the study utilises the resulting daily mass media sentiment indicators to estimate realised volatility, implied volatility and returns of the S&P500 and FTSE100 stock indices. It is argued that should markets be completely efficient, using mass media sentiment indicators would not improve the capability to predict markets the following day. The empirical data demonstrates that markets are not completely efficient. Mass media sentiment indicators improve the estimate of realised volatility in the stock market when paired up with a lagged value of realised volatility. Furthermore, mass media sentiment indicators possibly predict returns and implied volatility, but due to problems in statistical inference, further analysis is needed. It is concluded that the mass media plays an important role in disseminating information, predictions and sentiments about the market, consequently affecting the collective decisions and expectations of individual investors.

Keywords: asset markets, content analysis, financial journalism, media analysis, media sentiments, sentiment analysis, stock returns, volatility

INTRODUCTION

This dissertation studies the impact of mass media on investors' behaviour in financial markets. A new type of investors emerged in the end of the 20th century and with it dawned a new era of financial journalism. The 21st century Western middle-class existence is penetrated by the barrage of information produced by 24-hour news corporations and a myriad of online publications. Since the late 1980s, the fusillade of news, reports and analysis has been appropriated by one dominant topic, reporting on financial markets. For the Western middle-class, financial markets are not a discrete realm governed by economics but an integral part of politics, welfare and general culture that has immense power in society (Preda and Cetina, 2004). This study recognises the work of Preda and Cetina (2004), MacKenzie (2003, 2006), MacKenzie and Millo (2003) and others in their efforts to study the sociology of finance. Based on the pioneering work of Gordon Clark, Nigel Thrift and Ada Tickell (2004) and MacKenzie (2003, 2006), the relationship between media and finance emerges as an essential subject of research if the sociology of contemporary financial markets is to be understood. The aforementioned scholars provide a point of departure for this study: the notion of financial markets as a mediatised environment penetrated by social structures. However, the authors have mostly focused on culture and regulation. Only Clark et al (2004) have paid significant attention to media. The role of media and communication has to be made explicit and its impact has to be empirically evaluated. The focus on mass media arises from the preliminary evidence from finance research of media having a significant impact on financial markets (e.g. Tetlock, 2007; Fang & Peres, 2009; Klibanoff, Lamont & Wizman, 1998). The efforts of MacKenzie (2003, 2006) and Clark et al (2004) have not provided conclusive empirical evidence or even verifiable hypotheses. Furthermore, their understanding of financial markets has been incomplete. Additionally, the realm of academic economics and finance remains dismissive of the notion of social systems and hierarchies affecting financial markets (MacKenzie, 2006).

It is the purpose of this study to assess the impact of mass media sentiments on financial decision-making and expectations of individuals participating in financial markets. Previous literature suggests that the role of mass media is not limited to disseminating new information. Likewise, media disseminates sentiments and speculative accounts of the future of financial markets. The impact of this extensive role and the mechanism of action are explained with concepts arising from the sociology of economics (e.g. MacKenzie, 2006). In this study, the derived hypotheses are tested with broad empirical evidence. Several thousands of financial news reports are analysed and their inferences with two stock indices are measured. The research is inspired by novel methods of data collection that allow for a

more comprehensive analysis of mass media sentiments than has previously been possible. The aspiration of this study is to reveal a 'mass media effect'; an effect that the sentiments of financials news have on the market that goes beyond disseminating new information. The study links to a broader context of media influence (e.g. Wanta, Golan & Lee, 2004) and the power of communicative networks (Castells, 2007; 2009) in the context of financial news and financial decision-making.

The contribution of this dissertation is to bridge the gap between financial literature focused on finding correlations without substantial theoretical background and sociological literature that, while providing comprehensive theoretical analysis, often does so at the cost of empirical evidence. With a focus on media, it is argued that mass media nodes in the communicative networks of investors play a significant role in impacting the investment decisions and expectations of individual investors.

THEORETICAL FRAMEWORK AND PREVIOUS EMPIRICAL RESEARCH

The literature review begins with a discussion of the assumptions that this study and the majority of articles in the literature review make about financial markets. Next, this chapter proceeds to describe the historical and social context of the study and evaluate previous empirical evidence. Finally, the theoretical framework within which the hypotheses are positioned is presented.

The efficient market hypothesis (Samuelson, 1965; Fama, 1965) has been claimed to be the best-established theorem in social sciences (Jensen, 1978). The hypothesis argues that all information – past, present and predictions about the future – is immediately reflected in the prices of a traded asset (Fama, 1970). Fama (1965) and others (e.g. Kleidon, 1986; Marsh & Merton, 1986) have provided evidence to demonstrate that this is indeed the case. In the stock market asset prices are best modelled by a random walk. In the face of overwhelming empirical evidence, the efficient market hypothesis is accepted as a general standard in this study. However, this study agrees with Schleifer and Summers (1990) who argue that the expectations of all investors are not perfectly rational. Humans have been proven to be overconfident (Alpert & Raiffa, 1982), to extrapolate trends excessively (Andreassen & Krauss, 1987) and to overstate the importance of new information (Tversky and Kahneman, 1971). Consequently, this study assumes two types of traders: noise-traders who operate on

random beliefs of future returns and arbitrageurs[1] who operate rationally and are profit-seeking (Schleifer & Summers, 1990).

While arbitrageurs operate on rational expectations about future returns, noise-traders operate on signals that lead to non-rational conclusions, which cause noise-traders collectively to make non-random mistakes. Should markets be fully efficient, arbitrageurs would immediately take advantage of any irrational positions of noise-traders and bring prices to an equilibrium that reflects collective rational expectations (Schleifer & Summers, 1990). However, there are limits to risk-free arbitrage due to finite time horizons. Furthermore, it is assumed that both types of traders have a limited demand for assets as they either have a budget constraint or are risk-averse (Schleifer & Summers, 1990). These factors cause a delay in reaching the rational equilibrium. Furthermore, noise-traders do influence prices: a negative shock to their beliefs will result in noise-traders selling their assets to arbitrageurs, temporarily depressing prices (Schleifer & Summers, 1990). Conversely, a positive belief shock will increase valuations. As belief shocks follow a random walk, prices rebound in the next period (Schleifer & Summers, 1990). An assumption of investors operating on rational beliefs and investors operating on noise is plausible and empirically established (Shiller, 1980; Leroy & Porter, 1981; Schleifer & Summers, 1990). The line between these two types of investors is blurred, but in order to construct a theoretical framework, it is necessary to draw a distinction. The rational equilibrium is reached with a delay through arbitrageurs taking advantage of noise-traders irrational positions.

The Noise-Trader Emerges: Communicative and Regulative Changes

Modern finance theory is deeply rooted in the assumption of the efficient market. As the theorem suggests, empirical studies on the degree of temporal correlation in the time-evolution of asset prices have demonstrated the absence of memory (Mantegna, 1999). Past prices do not predict future prices on any time-series. Other empirical studies (Marshall & Cahan, 2005; Marshall, Qian, & Young, 2009) have produced similar results across a variety of fundamentals and technical factors. It can be concluded that the time-series of asset prices carry large amounts of information, but it is difficult to extract in order to make credible claims about future prices. Consequently, asset prices are best modelled in terms of a random walk (Samuelson, 1965). However, although prices are impossible to predict based on past prices, fundamentals or technical analysis, surprising correlations have been unearthed when

[1] Arbitrage is the practice of exploiting a price disparity between two markets. This study uses the term broadly to encompass any rational trading that takes advantage of irrational positions on the market.

studying the relationship between media and asset markets (e.g. Tetlock, 2007; Antweiler & Frank, 2004).

Institutional changes and the increasingly mediated experience of a growing group of individual investors have transformed financial markets, increasing the influence of mass media and social networks. Clark et al (2004) provide an excellent point of departure to the increasingly mediatised field of financial markets. The authors argue that from the late 1980s onwards, a new set of financial audiences and consumers has developed. The advances in IT and communications technology along with policy changes dramatically increased demand for mutual funds and consequently traditional assets such as bonds and stock.

The 401(k) pension funds[2] becoming commonplace in the United States made investing a part of middle-class existence while finance, popularised by films such as The Wolf of Wall Street, became a lucrative career choice for the best and brightest from the most exclusive universities (Preda & Cetina, 2004). Changes in the regulative environment made it possible for municipalities to bypass banks and raise money directly from the market by issuing bonds and for individuals to gain access to a multitude of asset classes (Clark et al, 2004). Two types of communicative changes escalated these developments (Clark et al, 2004). Advances in online banking and IT made trading available for an ever-wider group of individuals. Meanwhile, this information-poor audience demanded financial analysis and news from reporters, journalists, online gurus and institutional analysts to drive their investment decisions (Clark et al, 2004). By the end of 1990s, financial news had become entertainment and a part of the 24-hours news cycle, with a barrage of online tickers, news and analysis penetrating the *lebenswelt* of the 21st century Western middle-class (Preda & Cetina, 2004).

Clark et al (2004) argue that the media affects noise-traders and identify a number of mechanisms through which noise-traders affect markets. First of all, noise-traders base their investment decisions more on 24-hours news coverage than on financial formulas and consequently, end up temporarily depressing or appreciating prices. Secondly, rational investors take advantage of noise-based bubbles, timing their exit before a correction. Most importantly, institutional investors are human, too. Clark et al (2004) provide evidence through interviews and reviewing past literature: They argue that decisions in institutions are made by individual traders who are affected by the 24-hours news barrage from social and mass media. The rationality of institutional investors has been questioned in the past (e.g.

[2] 401(k) is a form of pension saving. Much of the investment decisions involved are the responsibility of the individual.

Hansen & Hill, 1991). However, simply implying that a proportion of institutional investors are noise-traders does not fundamentally alter the conclusions of the study. Even one investor (with no budget constraint) acting on rational principles would be sufficient to bring prices to a rational equilibrium. Nevertheless, Clark et al (2004) identify and explain the rise of a new group of consumers of financial assets and a new audience for financial news. They argue that the emergence of this group has transformed the relationship between media and finance. Financial players have been forced to acknowledge that markets can, due to the influence of mass media, move in unexpected ways. The authors claim that markets are driven less by financial fundamentals and more by media images, news and irrational exuberance, i.e. noise-traders and by the effect noise-traders have on arbitrageurs.

Clark et al (2004), while providing a good point of departure, neglect certain points. Due to the timing of the publication, the study is mostly concerned with the media feeding bull markets[3], although the nature of volatility spikes during aggressive bear markets (Hardouvelis & Theodossiou, 2002) which suggests a stronger media-effect during depressed markets. The authors give vague empirical evidence demonstrating the impact of communicative networks and mass media, but the empirical evidence and verifiability are lacking. The authors ignore the automated nature of markets and overemphasise the effect noise-traders can have on a market. As the vast majority of trades are executed either by bots or by institutions following strict financial guidelines, the impact noise-traders can have is most likely not quite as dramatic as the authors imply. There are empirical studies that might verify the authors' claim of prices driven by mass media. Next, articles investigating the impact of mass media on asset prices will be evaluated.

Measuring Noise: Empirical Evidence

Fang and Peress (2009) study the cross-sectional relation between media coverage and stock returns. Using a sample of all companies listed on the NYSE[4] and a sample of 500 on the NASDAQ[5], the authors analyse prices between January 1993 and December 2002 examining all articles published in major US newspapers. The authors find a significant negative correlation between media exposure and stock performance: stocks in oblivion earn better returns than stocks that are widely reported. This is consistent with Clark's theorem: the fact that media is reporting on a stock means generally that the information is not genuine news,

[3] Bull markets refer to a charging bull, a period when the general trend in the market is up. A bear market refers to a retreating bear, a period when the general trend is down.

[4] The New York Stock Exchange

[5] The NASDAQ Stock Market

in the sense that it would be new information for the market. The evidence provided suggests that reporting on individual stock media strengthens the impact of the original 'genuine news'. However, the study of Fang and Peress (2009) studied the period leading up to the 2001 stock market crash characterised by excessive reporting on superstar IT stock. Consequently, the results might not be generalisable across disparate periods. Furthermore, the article ignores the clustered nature of stock returns, assuming homoscedastic error terms[6], possibly exaggerating the impact of mass media. Similar studies face similar dilemmas: Klibanoff et al (1998) demonstrate that country-specific news has an effect on bond prices while Kim and Meschke (2011) found that stocks experience a strong run-up and reversal after a CEO interview on CNBC. Baber and Odean (2008) provide more evidence that verifies the article of Clark et al (2004). They argue that individuals are more likely to buy attention-grabbing stock when choosing from a large pool of potential assets. Frieder and Subrahmanyan (2005) notice a similar pattern, pointing out that individual investors prefer stock that has strong brand recognition.

All of these studies are consistent with the notion of the media having an effect on individual stock prices. However, these studies have looked at returns on very long time-scales and only accounted for the aggregate number of articles published. The phenomenon they provide evidence for is quite simple to explain with convenience alone. When choosing from a large pool of assets, people are more likely to buy assets they recognise. While convenience alone might be enough to explain the media effect, it is nonetheless possible that some part of the premium is caused by what Clark et al (2004) refer to as 'froth'; investments based on media sentiments. However, as these articles do not evaluate the sentiment of the messages, it is impossible to verify whether it is simply convenience at work or if individual investors indeed base their investment decisions on media sentiments. In order to evaluate if sentiments in communicative networks affect financial markets, different kind of evidence is essential.

Measuring Noise: Sentiment-Based Approaches

Paul Tetlock (2007, 2008) was the first to utilise algorithm-based content analysis to evaluate the impact of media sentiments on the stock market. In Tetlock (2007) the author discusses the empirics of what Clark et al (2004) largely ignored. Clark et al (2004) failed to provide a verifiable hypothesis and the aforementioned studies can only comment on the correlation between an aggregate number of news articles and the returns of individual stock. Neither is suitable for providing evidence for or against a sentiment-based media effect

[6] See Asteriou & Hall, 2008 chapter 14 p. 248 - 271

independent of fundamentals or 'genuine news'. Tetlock (2007) constructs a model to analyse the sentiment of the WSJ 'Abreast of the Market' column over a period from 1984 to 1999 and the returns of the Dow Jones[7]. Utilising the Harvard Psychosocial dictionary[8], the author counts the words and establishes a rudimentary pessimism indicator. Implementing a vector autoregressive framework[9], the author finds media pessimism to predict downward pressure on market prices. Furthermore, unusually high or low pessimism predicts higher trading volume. Finance literature suggests that time-series of asset prices tend to exhibit clustered volatility and returns. This might skew the results and impact choosing the appropriate regression models. It is possible that the author does not adequately test for such properties in the time-series. The methods and the results are consequently problematic. Using a different regression model might have brought different results, although most likely not render the results redundant.

Furthermore, the pessimism index is based on a general psychosocial dictionary, which is not an appropriate tool to perform content analysis on financial news. The word capital, for instance, has a completely different meaning in financial literature and in day-to-day English. Additionally, the timeline is problematic as the results might not be generalisable to, for instance, the periods of relatively turbulent markets from 2007 to 2011 and the period of very low volatility from 2011 to 2014. The most fundamental problem with the article is the lack of lagged values of the Dow Jones in the regressions. Without this information, the article cannot evaluate to what extent the information is already absorbed in the prices the previous day. For instance, predicting the returns of day (t+1)[10] using a sentiment indicator from (t+0) might indeed give convincing results. However, if the closing price for Dow Jones (t+0) is added to the regression as an independent variable, it is possible that the sentiment indicator might lose its meaning as an explanatory variable.

Notwithstanding its problems, the article does provide evidence that confirms to an extent the theory of Clark et al (2004). If one accepts a transmission model of communication (Chandler, 1994) and assumes that it is the mass media that disseminates new information among investors, the media effect would simply be the market reacting to new information about fundamentals. The results of Tetlock (2007) suggest that mass media is not a source for new information about fundamentals but a speculative market force of its own causing

[7] Dow Jones Industrial Average

[8] The Harvard IV dictionary, that attempts to group words in to meaningful categories

[9] See Elliott & Timmermann, 2013 chapter 2, p. 158-160

[10] (t+1) refers to a period (day) in the time-series that is one period after (t+0). (t+0) can be understood as today, (t+1) the following day and (t-1) the day before. This notation will be used throughout this essay.

noise-driven trading or froth. Antweiler and Frank (2004) find similar results in relation to market volatility and online message board sentiments. Antweiler and Frank (2004) analyse the sentiment of messages posted on online message boards. They find that a high volume of pessimistic messages implies downward pressure on prices in the relevant index. The article of Clark et al (2004) argues that the media effect should be stronger during bull markets. Tetlock (2007), Antweiler and Frank (2004), Campbell et al (1992) and De Long et al (1990) disagree with Clark et al (2004) arguing that media pessimism causes downward pressure on prices and higher volatility. This article agrees with Antweiler and Frank (2004) and Tetlock (2007), as their empirical evidence is more convincing. To conclude, research on media sentiments, market volatility and returns has been conducted, but it is limited in scope, the time-series are not generalizable and the research methods require improvements. However, based on the existing literature, it can be argued that mass media sentiments affect asset returns and market volatility.

A Mechanism of Action for Noise: the Performativity of Financial News

In order to formulate verifiable hypotheses and give a theoretical framework to a media effect that goes beyond merely disseminating new information about fundamentals, sociological theory on the performativity of economics is presented. Sociological theory is necessary if the mechanism, through which mass media sentiments affect noise-traders, is to be identified. Furthermore, sociological theory is essential if the broader social context is to be understood.

MacKenzie (2006) studies the sociology of arbitrage and financial markets, providing a background theory analysing the media effect on noise-traders. He disagrees with Parsons and Smelser's (1956) notion of the economy as an independent subsystem modelled using conventional economics. Piecing together White (1981), Granovetter (1985) and Callon (1998), MacKenzie claims that the 'technical core' of economics is just as much the business of sociologists as it is the business of economists (MacKenzie & Millo, 2003). MacKenzie (2006) claims that investors are not atomized, anonymous actors making decisions based on economic fundamentals but a community of individuals aware and affected by the actions of others. Social does not surround the economy; it is at its very core. Most importantly, MacKenzie (2006) adopts the notion of performativity (Callon, 1998) arguing that economics creates the phenomena it describes rather describing something that already exists. Economics "does things, rather than simply describing an external reality" (Mackenzie, 2006:29). This notion can conveniently be expanded to encompass communicative networks and the mass media. Communicative networks are suggested as the mechanism through which individual nodes disseminate information, become aware and constitute the decisions

of other nodes, in a similar manner as Castells (2007) describes. In MacKenzie and Millo (2003) the author uses the example of Black-Scholes options pricing model as an example of economic theory being performative. The Black-Scholes formula is a tool for estimating what the price of an option should be, given a set of variables about the market and the economy. When the equation was initially introduced in the form of Black-Scholes sheets[11], to the open outcry of trading floors of the Chicago Exchange, option prices did not match with the results of the equation. However, quite soon after prices started to converge. The Black-Scholes model is one of the few models that have a persistent track record of accurately estimating option prices. However, it did not describe the world that preceded it. Rather, the introduction of the formula created the world that it attempted to describe.

In Austin (1975) the author discusses performative utterances: simply by stating 'I name this ship the Queen Elizabeth' one performs the action. However, there is an obvious difference between the Queen naming the ship Queen Elizabeth and a shipyard worker attempting to name the ship MS Stalin. This was Bourdieu's (1991) critique: there are certain social conditions that must be met for an utterance to be performative. Bourdieu called these conditions the "conditions of felicity" (Bourdieu, 1991:73). It should be noted that in the case of the Black-Scholes model the conditions were to a large extent material, which Bourdieu does not consider. Black-Scholes sheets had plenty of competition, but they were widely available, easy to carry around and easy to interpret. These were the material conditions. On the other hand, they came from a respected place in the academia i.e. from the correct location in the interrelations of language, legitimacy, cultural hierarchy and power (Bourdieu, 1991). The Black-Scholes sheets created knowledge that was confirmed by the practices the knowledge produced. This is what MacKenzie (2006) calls "Barnesian performativity", after sociologist Barry Barnes (Barnes, 1983). On a framework of performativity introduced by MacKenzie (2003, 2006), Barnesian performativity is at the highest level: general performativity refers to any knowledge that is used, effective refers to knowledge being used and having an effect. Barnesian performativity refers to knowledge that while used confirms itself, a circular logic of sorts. Barnesian performativity is a valuable theorem in terms of financial news and media power. If the nodes in a Castellian (Castells, 2007) communicative network transmit a message of tomorrow being a good day, their collective decisions will make that a reality. However, this is possible only if the nodes in question meet the conditions of felicity (Bourdieu, 1991) or have adequate networked power (Castells, 2009),

[11] The Black–Scholes is a model of a financial market for derivatives. The formula gives an estimate of the price of European-style options. The publication of the formula led to a boom in trading options on the Chicago Options Exchange.

e.g. power to influence other nodes. Financial newspapers are, in the communicative network of noise-traders, sufficiently powerful nodes.

Conceptual Framework – Main Aspects of the Theoretical Approach Used to Frame the Project

The advances in communications technology and changes in policy regulating financial markets brought about a group of information-poor noise-traders. Noise-traders affect the market through their investment decisions and are affected by the media, through their demand for 24-hour information, news and analysis of markets. The investment decisions of noise-traders are best modelled within a Castellian network (Castells, 2007). Countless nodes mutually constitute the knowledge and decisions of the network. It is the effect of one group of nodes in particular this study focuses on. This group is the mass media, as one can expect the media effect to be significant enough to be both empirically observable and of importance to social scientists and investors. Three mechanisms of action are identified in terms of noise-traders affecting the market: the investments made by noise-traders, arbitrageurs adjusting their expectations to the new environment and arbitrageurs themselves being affected by the barrage of information available.

The mechanism for action of the media effect is assumed to be the performativity of financial news. MacKenzie's (2006) notion of the Barnisian performativity of economics is extended to encompass financial news. The argument is that most news reported by the media is not genuine new information. However, if a source is in the right location in terms of the interrelations of language, legitimacy, cultural hierarchy and power conditions of felicity are met and the utterances become performative. Extending the idea of Bourdieu (1991), the source must also be materially widely available, as was the case with the Black-Scholes sheets. This can be rephrased following Castells (2007): if a node in the communicative network has sufficient networked power (Castells, 2009) and broadcasts a certain message to other nodes, the message becomes a reality. Consequently, financial news can have an effect on the market that goes beyond disseminating new information they provide (if any). The mass media effect should follow the sentiments of media, i.e. negative news days be followed by depressed markets.

Should markets be efficient and investors operate based on financial fundamentals, there should be no observable media effect for news (t+0) in period (t+1), because all information is absorbed in prices in (t+0) or before the information is printed. However, as financial news comes from sources that meet the conditions of felicity, the effect is amplified. This occurs

through noise-traders' actions and arbitrageurs adjusting their expectations because of noise-traders. Consequently, one can observe the media effect of financial news (t+0) in (t+1) as information is relatively slowly disseminated in the communicative network of investors. Therefore, sentiment analysis of mass media should be able to predict volatility and returns in period (t+1) based on media sentiments from (t+0).

The main research question of this study is whether the sentiments of financial news in period (t+0) affect stock markets in (t+1), and if so, how and to what extent. This question should be of great interest to social scientists and investors alike. For social scientists it provides empirical evidence of media power affecting financial markets. Furthermore, should the research provide evidence of a correlation between media sentiment in (t+0) and stock market events in (t+1), it would be possible to amend an existing trading strategy by hedging against fluctuations using information derived from media sentiments.

Statement of the Objectives of Research

The objective of this study is to find evidence for a sentiment-based media effect in period (t+1) during relatively recent periods of bull markets and bear markets, when explaining daily returns and volatility with media sentiments. Therefore, the study aims to test the following hypotheses derived from existing literature and the theory presented above.

> H1: There is correlation between media pessimism in (t+0) and stock market volatility in (t+1).
> H2: There is negative correlation between media optimism in (t+0) and stock market volatility in (t+1).
> H3: There is a correlation between news frequency in (t+0) and stock market volatility in (t+1).
> H4: The effect is diminished if lagged values of the index analysed are introduced as independent variables but media sentiments improve the explanatory power of the model.
> H5: The observed media effect is stronger during bear markets than during bull markets.
> H6: Media sentiments predict returns.

These hypotheses require further discussion. H1 through H3 are based on past empirical evidence (Fang & Peres, 2009; Klibanoff et al, 1998; Kim & Meschke, 2004) and the nature of volatility clustering in financial markets (Asteriou & Hall, 2008). H1 to H3 would, if accepted,

provide evidence for a media effect in the sense that Clark et al (2004) propose. H4 states that the majority of information the media provides is absorbed into prices much faster than the media or noise-traders can react. However, if using media sentiment data improves the model, this would be further evidence of a group of noise-traders investing based on media sentiments and of the performativity of financial news. H5, if accepted, contradicts the argument of Clark et al (2004) of a stronger media effect during bull markets. H6 would provide the strongest evidence for the performativity of news. If accepted, it would imply that a negative news day (t+0) predicts negative stock returns in (t+1).

METHODOLOGY

It is against the above theoretical background that the research outline is constructed. The methods utilised draw from the positivist research tradition and due to the nature of the research subject quantitative methods are applied. Since the objective of the research is uncovering the effect of the sentiment of financial news on the stock market, the data collected must encompass an indicator of stock market returns, volume and volatility along with data of the sentiment of the 'news day'. According to Holsti (1969), content analysis can be used for making inferences in large body of text in a systematic fashion. It is a method of generalising and summarising great amounts of data (Neuendorf, 2002). However, content analysis is often misused in social sciences. Content analysis does not meet the criteria of transparent, replicable research in many studies within the broad field of sociology. Sentiments and frames simply "emerge from analysis" (Hanson, 1995:384) and the processes which lead to specific sentiments or frames being identified are not scrutinised (e.g. Hanson, 1995; Kirstensen & Orsten, 2007; Dimitrova & Stromback, 2005). Faith is placed in the judgement of the scholar, and sentiments and themes are simply "found" (Haller & Ralph, 2001:412). Keeping these caveats in mind, content analysis is nonetheless identified as the only appropriate method for analysing a data sample large enough for observing media sentiments. The critique of Matthess and Kohring (2008) argues that sentiments should ideally not be identified *a posteriori*. Research should have an *a priori* rule set or an algorithm compute the frames from the sample. Otherwise, the research is subject to a number of biases, most importantly a confirmation bias (Matthess and Kohring, 2008).

The classical linear regression model is a way of analysing the relationship between two, or more, variables (Asteriou & Hall, 2008). A set of independent variables is used to explain values for a dependent variable by fitting a linear equation into the observed data. This method is identified as the best available method for revealing a connection between

quantified mass media sentiments and financial data based on previous studies applying similar methods (Tetlock, 2007; Fang & Peres, 2009) and econometric theory (Asteriou and Hall, 2008).

Selection of Empirical Data

Following the steps identified by Krippendorff (1980) the data analysed consists of all the articles mentioning the S&P500 or FTSE100 index from the Financial Times (FT) and the Wall Street Journal (WSJ) during the period from 4th of January 2007 to the 31st of December 2011 and subsequently, the period from 1st of January 2013 to the 15th of July 2014. The periods are chosen, because the first period is relatively turbulent – as measured by volatility – with both bear and bull markets. Characteristic for the second period is a bull-run paired with record low volatility across the entire time-series. This choice makes it possible to study media sentiment and volatility/returns in very different market environments. The newspapers are chosen because of their position as global leaders in terms of opinion and circulation. They have a broad audience both in their respective countries as well as internationally. They can be expected to meet the conditions of felicity (Bourdieu, 1991) or have sufficient networked power (Castells, 2009) to impact noise-traders. The unit of analysis is the 'news day' in a given newspaper, i.e. all the articles published on a given day mentioning the relevant indices.

The quantitative time-series comprises of data from the S&P500[12] and the FTSE100[13]. These indices are chosen, because they are the most liquid and broadly followed stock indices among the readers of the FT and the WSJ. As the hypotheses defined earlier require data on returns and volatility, the time-series will include data on the daily returns of both indices along with data of the volatility, as measured by absolute returns and implied volatility. Absolute returns is an indicator of volatility that shows how much the market moved regardless of the direction. Implied volatility is commonly regarded as the 'fear gauge' of the stock market as it measures the 30-day volatility expectations (Whaley, 2000). It has a high correlation with absolute returns, but both are included in the study. As implied volatility is calculated based on option prices, it is a collective prediction of future volatility. Therefore, using both implied and realised volatility allows one to study whether the expectations of investeros are more predictable than their actions. Daily absolute returns of the S&P500 or FTSE100, on the other hand, indicate the realised volatility or what occured on a given day.

[12] Standard and Poor's 500 Index is a capitalization-weighted index of 500 stocks.

[13] The FTSE100 is an index of the 100 companies listed on the London Stock Exchange with the highest market capitalization.

Similar data is available on the FTSE100, although the implied volatility of the FTSE100 is not public. The analysed dependent variables will therefore be the daily realised volatility (or absolute returns) and the returns of the FTSE100 and S&P500 along with the corresponding value of implied volatility of the S&P500.

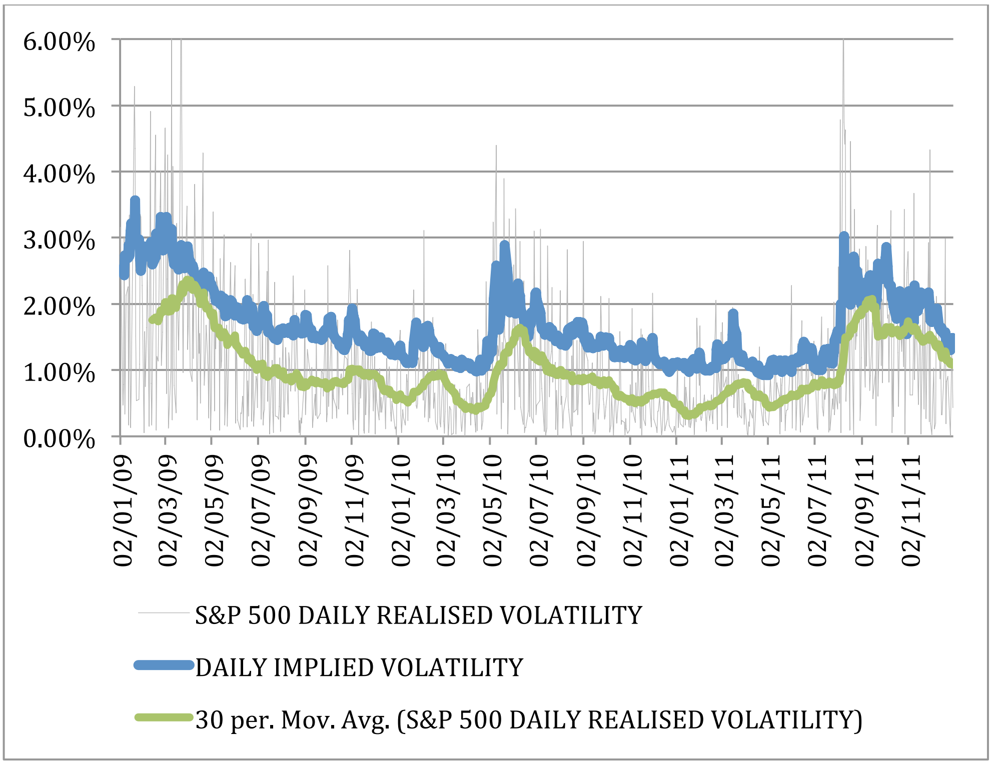

Figure 1. S&P500 Implied Volatility and Realised Volatility

Design of Research Tools

The critique of Matthess and Kohring (2008) is applied and the sentiment analysis will be conducted using the QDAMiner software to count the occurrences of words and phrases indicating pessimism, optimism and uncertainty. The indicator is computed dividing the amount of negative, positive or uncertain words with the number of total words. The Loughran & McDonald Financial Sentiment Dictionary is utilized for categorising the words as using a general sentiment dictionary has proven problematic as discussed in the literature review (Loughran & McDonald, 2011; c.f. Tetlock, 2007). The Loughran & McDonald Financial Sentiment Dictionary is specifically formulated to identify sentiments in financial news and has been used successfully in Loughran & McDonald (2011).

Each news day comprises of all the articles from one newspaper mentioning either the S&P500 or the FTSE100 index. All the articles from a given day are pooled, the total word count computed after which a normalized indicator for pessimism, optimism and uncertainty are computed. A preliminary sentiment analysis was conducted on 50 articles from the sample. The researcher read the articles and marked down if the article indicated pessimism, optimism or uncertainty. These manual results were compared to the results given by the The Loughran & McDonald Financial Sentiment Dictionary. Comparing the manual results with the algorithmic results gave an agreement rate of 71%. This was deemed insufficient for subsequent analysis. Therefore, the QDAMiner phrase finder along with a manual scan of the preliminary sample was used to amend the The Loughran & McDonald Financial Sentiment Dictionary. A similar test was conducted after the amendments with an agreement rate of 85.5%. In this study, the two coders were the researcher and QDA Miner Software. It has been established in literature on content analysis that the agreement rate is not a sufficient measure of intercoder reliability (Carletta, 1996). A more robust measure was also calculated: the Cohen's Kappa. Cohen's Kappa improves the estimate of intercoder reliability by taking into account that the two coders might agree by chance alone.

$$K = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}$$

Pr(a) indicates the observed agreement among the two coders. Pr(e) is the probability that the coders would agree by chance, calculated from the observed data. The initial analysis returned K = 0.34, which, as established in Fleiss et al (1981) and Landis & Koch (1977), is not sufficient. However, after improvements to the sentiment dictionary Cohen's Kappa of 73% was reached. A Kappa above 61% is considered substantial by Landis and Koch (1977) and excellent by Fleiss et al (1981). Although both articles can be critiqued for quite arbitrary conclusions, it was decided that paired with the high agreement rate this was a sufficient indicator of an acceptable level of intercoder reliability.

Normalizing the sentiment data is done using feature scaling in order to bring each media indicator for each day into the range $0 \leq x \leq 1$[14]. The result of this are pessimism, optimism and uncertainty indicators that vary between 0 and 1 for each newspaper and news day. The total number of articles published is also included in the time-series, in order to study H3;

[14] Feature scaling is conducted with the following formula:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

the correlation between the stock market indicators and the aggregate number of articles published on a given day.

The sentiment data and financial data will be combined into one main time-series (for both periods) for subsequent analysis using linear regression models. In this analysis, the independent variables are the mass media sentiment indicators (MMSI: positivity, negativity, uncertainty and aggregate number of articles) while the dependent variables are the absolute returns (realised volatility), implied volatility (as measured by the VIX) and daily returns. The model for each dependent variable can be written as:

$$Y_t = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + u_t$$

For instance in the case of the Financial Times and the S&P500 implied volatility (Y_t)

Table 1. The Dependent Variables of the Regression Models

$$X_{1t} = normalized\ positivity\ indicator$$
$$X_{2t} = normalized\ negativity\ indicator$$
$$X_{3t} = normalized\ uncertainty\ indicator$$
$$X_{4t} = aggregate\ number\ of\ articles$$
$$X_{5t} = Lagged\ value\ (t-1)\ of\ the\ dependent\ variable\ Y_t$$

$$VIX_{(t+1)} = \alpha + \beta_1 NPI_{(t+0)} + \beta_2 NNI_{(t+0)} + \beta_3 NUI_{(t+0)} + \beta_4 ANA_{(t+0)} + u_t$$

$\beta_1 - \beta_4$ are respective coefficients of X_n. The error term is denoted u_t. A similar model will be constructed for each newspaper and returns, implied volatility and realised volatility. A standard method for estimating the coefficients of a linear regression model is that of the least sum of squares method[15]. Based on Asteriou & Hall (2008) and the application of Tetlock (2007) it is argued that the method is appropriate for this analysis.

This chapter has outlined the theoretical foundations of the chosen methodology along with the practicalities of conducting the research. Algorithmic content analysis with an *a priori* defined codebook is identified as the appropriate method to assess media sentiments in the

[15] Least squares implies that the solution minimizes the sum of the squares of the errors. See Asteriou and Hall (2008).

sample. Two periods of reporting from the FT and the WSJ are chosen based on the standing of the publications in the financial world and the divergent nature of the periods in terms of volatility and returns. Linear regression models are used to identify inferences using the OLS method. The returns, the implied volatility and realized volatility are studied in order to establish whether mass media sentiments can predict returns and volatility or expectations of volatility.

Table 2. The Dependent Variables

SP&500 (Y_t)	FTSE100 (Y_t)
Implied Volatility	Implied Volatility
Realised Volatility	Realised Volatility
Daily Returns	Daily Returns

For both time-series, the dependent variables from table 2 are explained using the mass media sentiment indicators X_1 to X_4 as independent variables.

Furthermore, for all dependent variables an additional test is conducted. X_5 will be a (t-1) lagged value of the index analysed, as it can be assumed that a time-series of returns or volatility exhibits significant autocorrelation. This test is compared to a regression where only X_5 will be used to explain any Y_t in order to establish to what extent X_1-X_4 improve the capability to predict Y_t. In an efficient market the best forecast of Y_t is $Y_{(t-1)}$. If the market is not efficient, the forecast of Y_t can be improved by adding information about noise in the form of the mass media sentiment indicators. Any difference in the R-squared of these regressions will be interpreted as a mass media effect, although it is expected that X_1 to X_4 will exhibit significantly higher p-values when the regression is amended with the lagged values of the index analysed. The initial OLS-regressions are follows:

For both periods. 1st of January 2007 – 31st of December 2011 and the 1st of January 2013 – 15th of July 201.

$$SP500\ Returns_{(t+1)} = \alpha + \beta_1 NPI_{(t+0)} + \beta_2 NNI_{(t+0)} + \beta_3 NUI_{(t+0)} + \beta_4 ANA_{(t+0)} + u_t$$

$$SP500\ Implied\ Volatility_{(t+1)} = \alpha + \beta_1 NPI_{(t+0)} + \beta_2 NNI_{(t+0)} + \beta_3 NUI_{(t+0)} + \beta_4 ANA_{(t+0)} + u_t$$

$$SP500\ Realised\ Volatility_{(t+1)} = \alpha + \beta_1 NPI_{(t+0)} + \beta_2 NNI_{(t+0)} + \beta_3 NUI_{(t+0)} + \beta_4 ANA_{(t+0)} + u_t$$

$$FTSE100\ Returns_{(t+1)} = \alpha + \beta_1 NPI_{(t+0)} + \beta_2 NNI_{(t+0)} + \beta_3 NUI_{(t+0)} + \beta_4 ANA_{(t+0)} + u_t$$

$$FTSE100\ Realised\ Volatility_{(t+1)} = \alpha + \beta_1 NPI_{(t+0)} + \beta_2 NNI_{(t+0)} + \beta_3 NUI_{(t+0)} + \beta_4 ANA_{(t+0)} + u_t$$

Subsequently, all tests will be conducted again amended with X_5. These results will then be compared with results explaining the six dependent variables using only X_5 to compare the R-squared of these regressions.
For example:

$$SP500\ Returns_{(t+1)} = \alpha + \beta_1 NPI_{(t+0)} + \beta_2 NNI_{(t+0)} + \beta_3 NUI_{(t+0)} + \beta_4 ANA_{(t+0)} + SP500\ Returns_{(t+0)} + u_t$$

RESULTS

The results section is divided in two subsections. The first section summarises the regressions, statitics and significance tests. The second section states whether the hypotheses have corroborated or not and discusses the implications of the results.

The first time-series analysed was the Financial Times and the Wall Street Journal articles 2009-2011. However, as preliminary results demonstrated that adding the WSJ to the sample provided no improvements in reliability or significance, but greatly increased the need for processing power and disk space, the WSJ articles were removed from subsequent analysis. After the removal, the number of time-series observations, n, was 753 and the information of 4971 articles were analysed. The articles were scraped from the respective websites using a Python-script created by the researcher. The words were categorised into negative, positive and uncertain categories. Articles were grouped by day and the mass media sentiment indicators were calculated. Ten observations were removed from the sample due to errors in data collection[16]. These steps were repeated for the 2013-2014 sample (n=387, 3063 articles analysed).

Thereafter, the data was paired with the relevant financial data. The daily returns, daily realised volatility and implied volatility were added to the time-series. The implied volatility data for FTSE100 was not available despite numerous attempts to receive said data from the

[16] The mean word count of the total sample was 606. News days where total word count was smaller than 50 on a given day were removed

FTSE Group. Newspapers publish on most days of the year, but there are only 252 trading days the redundant sentiment observations were removed.

Table 3. The Independent Variables (X) and Dependent Variables (Y)

NPI (X_1) Normalised Positivity	Normalised positivity indicator
NNI (X_2) Normalised Negativity	Normalised negativity indicator
NUI (X_3) Normalised Uncertainty	Normalised uncertainty indicator
ANA (X_4) Number of Articles	Aggregate number of articles
SP_RETURN_DAILY (Y)	The returns of the S&P500 on the day (%)
IMPLIED_VOL (Y)	The 30-day implied volatility the VIX predicted that day (%)
SP_RETURN_ABS (Y)	The realised volatility of the S&P500

A brief look at the nature of the two periods is necessary before moving on to the main results. Figures 2 and 3 indicated that the period from 2009 to 2011 was a relatively turbulent one. Although the general trend was a bull market, there were some steep up and downs. Market volatility was considerably above its long-term trend. The period from 2013 to 2014, as illustrated by figure 3 and 4, was a straightforward bull market with little volatility to speak of. The average daily volatility during the 2009-2011 period was around 1% whereas for the 2013-2014 period daily volatility was merely 0.52%.

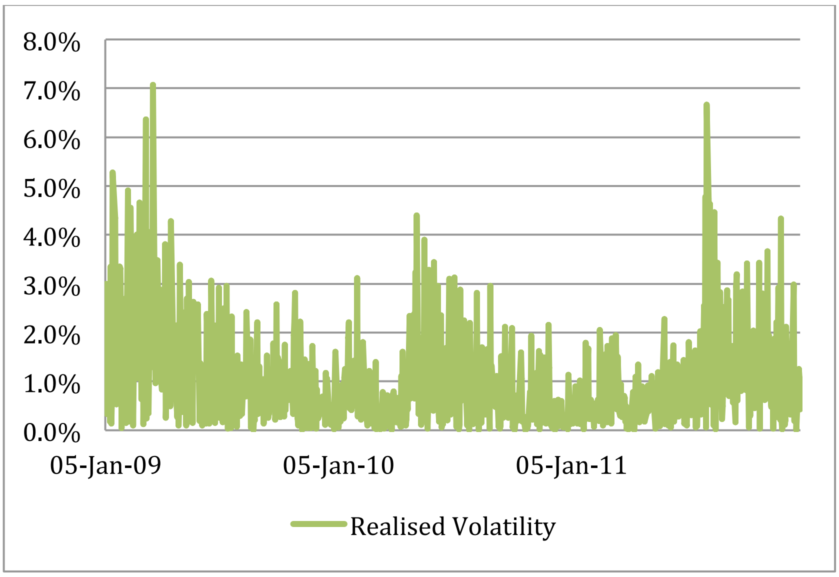

Figure 2. The Realised Volatility of the S&P500 January 2009 - December 2011

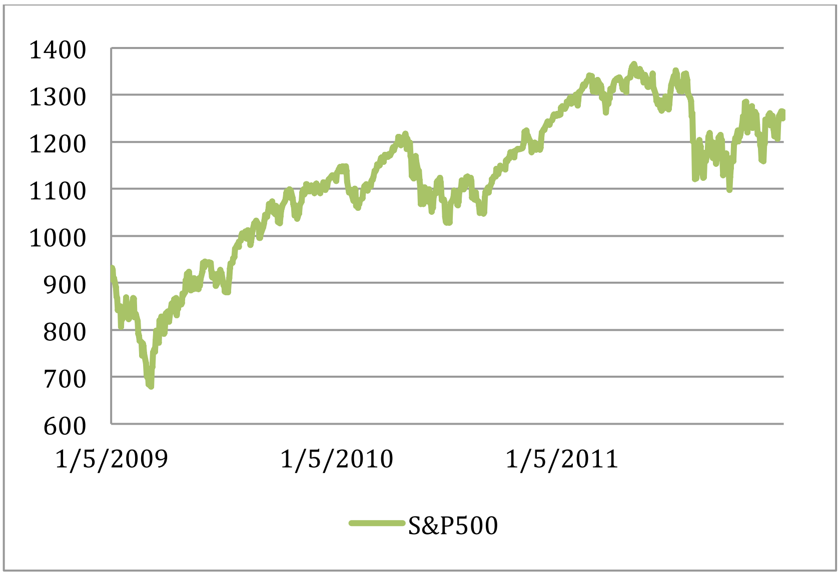

Figure 3. The S&P500 Index January 2009 - December 2011

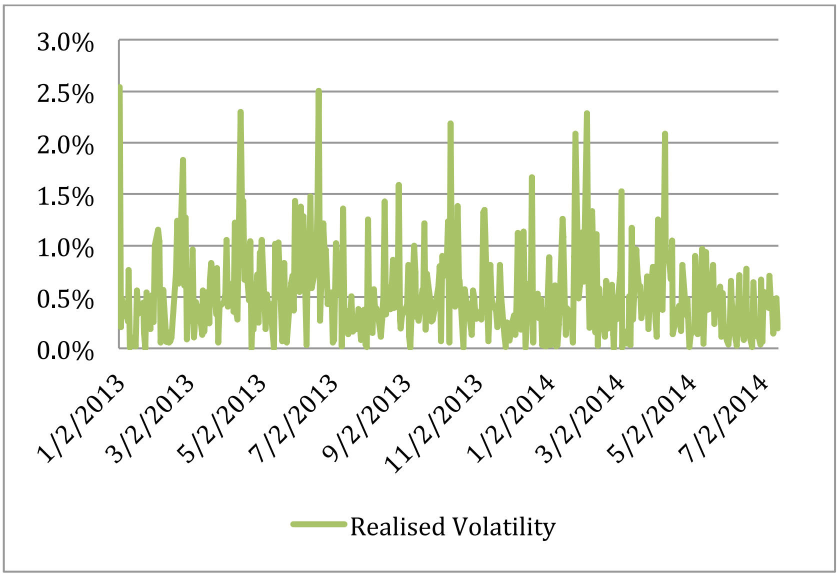

Figure 4. Realised Volatility of the S&P500 January 2013 - July 2014

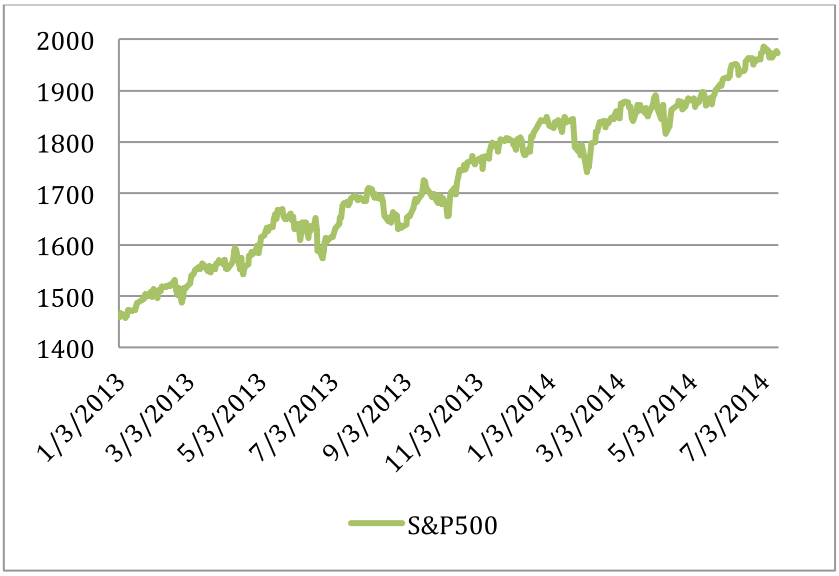

Figure 5. The S&P500 Index January 2013 - July 2014

The periods were not different only in terms of the market, but also in terms of financial news. The earlier period was characterised by more articles that exhibited extreme positivity or negativity. The mean negativity and positivity indicators for the 2013-2014 (table 5) period were 0.43 and 0.35 respectively, whereas for the 2009-2011 (table 6) period mean negativity and positivity were 0.52 and 0.48, respectively.

Therefore, the 2009-2011 period was a much more volatile and nervous period, which is captured by the sentiment analysis on the financial news of the time (table 5). The period of 2013-2014 represents a time of a calm bull market, also reflected in the results of the analysis. This result, in its own right, indicates that mass media is more extreme in its views in times of turbulence, perhaps further escalating bear markets, as subsequent analysis suggests.

Table 4. Descriptive Statistics of the Mass Media Sentiment Indicators 2013 - 2014

2013-2014	Positivity	Negativity	Uncertainty	Observations
10% percentile	0.48	0.58	0.52	0.69
20% percentile	0.42	0.52	0.46	0.62
5% percentile	0.53	0.63	0.58	0.77
Median	0.34	0.42	0.36	0.54
Mean	0.35	0.43	0.37	0.49
Standard Dev.	0.10	0.11	0.12	0.17

Table 5. Descriptive Statistics of the Mass Media Sentiment Indicators 2009 - 2011

2009-2011	Positivity	Negativity	Uncertainty	Observations
10% percentile	0.68	0.68	0.23	0.43
20% percentile	0.61	0.63	0.20	0.33
5% percentile	0.74	0.74	0.25	0.48
Median	0.46	0.52	0.16	0.29
Mean	0.48	0.52	0.17	0.22
Standard Dev.	0.15	0.12	0.05	0.14

Keeping the disparate nature of the two periods in mind, the results of the linear regressions are presented. The first group of regressions is:

$$SP500\ Returns_{(t+1)} = \alpha + \beta_1 NNI + \beta_2 NPI + \beta_3 NUI + \beta_4 ANA + u_t$$

$$SP500\ Implied\ Volatility_{(t+1)} = \alpha + \beta_1 NNI + \beta_2 NPI + \beta_3 NUI + \beta_4 ANA + u_t$$

$$SP500\ Realised\ Volatility_{(t+1)} = \alpha + \beta_1 NNI + \beta_2 NPI + \beta_3 NUI + \beta_4 ANA + u_t$$

If there is a significant linear relationship between the independent variable X and the dependent variable Y, the slope will not equal zero. The null and alternative hypotheses can be formulated as follows.

$$H0: \beta_n = 0$$

$$Ha: \beta_n \neq 0$$

1.1. Explaining Daily Returns Using Mass Media Sentiment Indicators

SP500 Returns 1st of January 2009 – 31 of December 2011				
Dependent Variable: Daily Returns SP500(1)				
Method: Least Squares				
Date: 08/11/14 Time: 11:59				
Sample (adjusted): 1 752				
Included observations: 752 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00054	0.00349	-0.154781	0.877
NNI	-0.002159	0.004599	-0.469464	0.6389
NPI	0.0000744	0.003711	0.020044	0.984
NUI	0.007808	0.011277	0.692352	0.4889
ANA	0.003839	0.003797	1.011081	0.3123
R-squared	0.002075	Mean dependent var		0.000515
Adjusted R-squared	-0.003269	S.D. dependent var		0.014573
S.E. of regression	0.014597	Akaike info criterion		-5.609371
Sum squared resid	0.159166	Schwarz criterion		-5.578635
Log likelihood	2114.123	Hannan-Quinn criter.		-5.597529
F-statistic	0.388257	Durbin-Watson stat		2.205498
Prob(F-statistic)	0.817126			

Table 6. Regression Results for the SP500 Returns 1st of January 2009 – 31st of December 2011

The regression in table 6 uses mass media sentiment indicators (t+0) to estimate returns of the S&P500 (t+1). The results of the FTSE100 regression were sufficiently similar and will not be discussed separately. The null hypothesis is accepted for daily returns for all independent variables based on the t-tests. No mass media sentiment indicator (MMSI) is significant at the 90% level. The result of the F-test, which estimates the significance of the

independent variables as a group, gives a similar result. Although a more sophisticated regression analysis could yield different outcomes, this result is consistent with the theory of Fama (1965) and Samuelson (1965). Stock returns appear to be a random walk. Adding information about mass media sentiments provides no insight into the direction of the market. This strongly contradicts the suggestion of Clark et al (2004) and the results of Tetlock (2007).

Based on this regression model it would seem that there is no statistically significant relationship between media sentiments and returns on the following day. However, if one analyses a sample of particularly positive or negative news days, a connection becomes apparent (see figure 6). This indicates that there indeed is a relationship, but the basic content analysis implemented in this study is not sufficient for unveiling that relationship in the entire sample. Analysing the 10% and 20% percentiles of the strongest positive and negative days from the 2009-2011 data, however, reveals the following relationship:

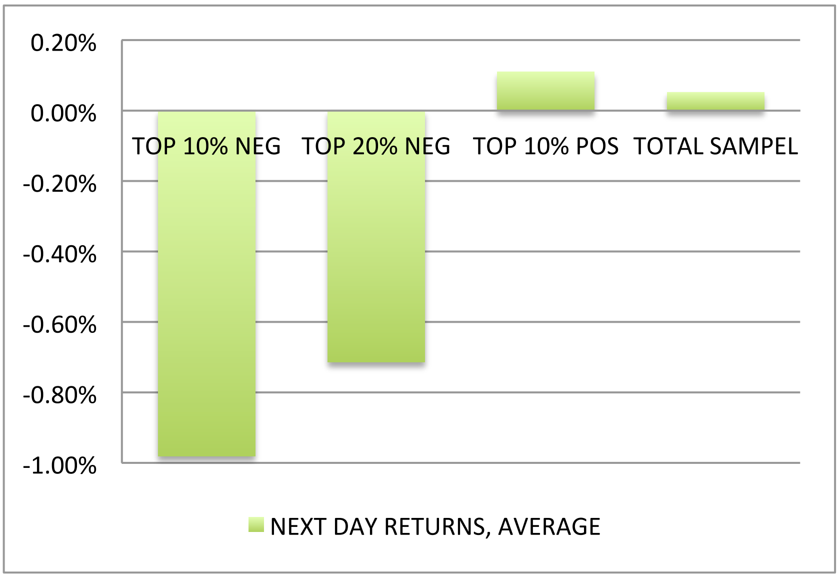

Figure 6. The Average Returns (t+1) of the S&P500 Grouped by Media Sentiment (t+0)

Figure 6 demonstrates that the 10% percentile of most negative news days were followed by average returns of -1.0%, considerably lower than the total sample average 0.05%. Similar patterns were found when observing the top 20% percentile. The most positive news days were followed by average returns of 0.11%. Consistent with previous empirical studies, negative sentiments increase volatility, but also imply a downward pressure on prices. The time-series from 2013-2014 exhibits similar properties, although they are, in general, much weaker.

Overall, the evidence on the relationship between returns and mass media sentiment indicators is inconclusive. The null hypothesis for the regression has to be accepted, but this could be due to a misspecification in the model. Analysing particularly positive and negative days does suggest an intuitive relationship. However, because of the small sample size (n=43 for top 10% percentile) this could be due to chance alone. Furthermore, the regression for returns shows no autocorrelation. The Durbin-Watson[17] stat is sufficiently high. This implies that the F-test and t-tests are reliable, and further evidences an efficient market. All information is already absorbed in prices (t+0).

1.2. Explaining Volatility and Implied Volatility Using Mass Media Sentiment Indicators

SP500 Realised Volatility 1st of January 2009 – 31 of December 2011				
Dependent Variable: Realised Volatility SP500(1)				
Method: Least Squares				
Date: 08/11/14 Time: 11:59				
Sample (adjusted): 1 752				
Included observations: 752 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.006527	0.002437	2.678273	0.0076
NNI	0.005658	0.003212	1.761638	0.0485
NPI	-0.006089	0.002592	-2.349268	0.0191
NUI	0.007961	0.007876	1.010771	0.3125
ANA	0.010562	0.002652	3.982754	0.0001
R-squared	0.042298	Mean dependent var		0.010226
Adjusted R-squared	0.03717	S.D. dependent var		0.010389
S.E. of regression	0.010194	Akaike info criterion		-6.327319
Sum squared resid	0.077633	Schwarz criterion		-6.296582
Log likelihood	2384.072	Hannan-Quinn criter.		-6.315477
F-statistic	8.248063	Durbin-Watson stat		1.933245
Prob(F-statistic)	0.000002			

Table 7. Regression Results for the SP500 Realised Volatility 1st of January

[17] Values below 1 are generally considered as alarming. See Gujarati et al (2009) Basic Econometrics (5th ed.).

SP500 Implied Volatility 1st of January 2009 – 31 of December 2011				
Dependent Variable: Implied Volatility SP500(1)				
Method: Least Squares				
Date: 08/11/14 Time: 12:00				
Sample (adjusted): 1 752				
Included observations: 752 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.012813	0.001183	10.82826	0
NNI	0.006969	0.001559	4.469063	0
NPI	-0.004782	0.001258	-3.800194	0.0002
NUI	-0.001947	0.003824	-0.509063	0.6109
ANA	0.011561	0.001288	8.978761	0
R-squared	0.164011	Mean dependent var		0.016396
Adjusted R-squared	0.159534	S.D. dependent var		0.005399
S.E. of regression	0.00495	Akaike info criterion		-7.772394
Sum squared resid	0.0183	Schwarz criterion		-7.741657
Log likelihood	2927.42	Hannan-Quinn criter.		-7.760552
F-statistic	36.638	Durbin-Watson stat		0.282513
Prob(F-statistic)	0			

Table 8. Regression Results for the SP500 Implied Volatility 1st of January 2009 – 31st of December 2011

The regressions in table 7 and 8 use the mass media sentiment indicators (t+0) to estimate realised volatility (table 7) and implied volatility (table 9) of the S&P500 (t+1). The results from the FTSE100 are sufficiently similar so the different indices are not discussed separately. The null hypothesis is rejected for realised volatility and implied volatility of the S&P500 for negativity, positivity and aggregate number of articles (NNI, NPI, ANA)[18].

The NNI and NPI are significant for realised and implied volatility for both the S&P500 and FTSE100 at the 95% level (table 7 and 8), with the exception of NNI for FTSE100, which is significant at the 90% confidence level. The R-squared of SP500 realised volatility is 4.22% (table 7). Furthermore, according to the F-statistic both volatility models are significant. The F-test verifies that the slope coefficients are as a group statistically significantly different

[18] Reminder of the mass media sentiment indicators (MMSI).

NNI = Normalised negativity indicator.

NPI = Normalised positivity indicator

ANA = Aggregate number of articles

NUI = Normalised uncertainty indicator

from zero. The normalised uncertainty indicator (NUI) is not significant in any of the tests, most likely due to insufficient variance in the sample.

The estimated model for realised volatility gives the most robust results. The t-tests and p-values along with the F-test are within reasonable bounds. Furthermore, the Durbin-Watson statistic[19] does not exhibit significant autocorrelation within the model, so the t-tests are reliable (table 7). However, for implied volatility the Durbin-Watson statistic is alarmingly low (table 9). "Serial correlation in residuals of the model as indicated by the low Durbin-Watson statistics does not affect the unbiasedness or consistency of the estimated coefficients" (Asteriou & Hall, 2008:137). However, the estimates from the OSL model are inefficient. The estimated variances of the regression coefficients are inconsistent, and t-statistics indicate higher significance than would be accurate, possibly resulting to an incorrect rejection of the null hypothesis. However, based on the robust results from realised volatility, it is concluded that the probability statistics are likely to be inflated, but not to an unacceptable extent.

Table 9. Coefficients and Probability Statistics for the Mass Media Sentiment Indicators. Dependent Variable S&P500 Realised Volatility. Period 2009-2011

SP500 Realised Volatility				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
NNI	0.005658	0.003212	1.761638	0.0485
NPI	-0.006089	0.002592	-2.349269	0.0191
NUI	0.007961	0.007876	1.010771	0.3125
ANA	0.010562	0.002652	3.982754	0.0001
R-squared	0.042298			

[19] Values below 1 are generally considered as alarming. See Gujarati et al (2009) Basic Econometrics (5th ed.).

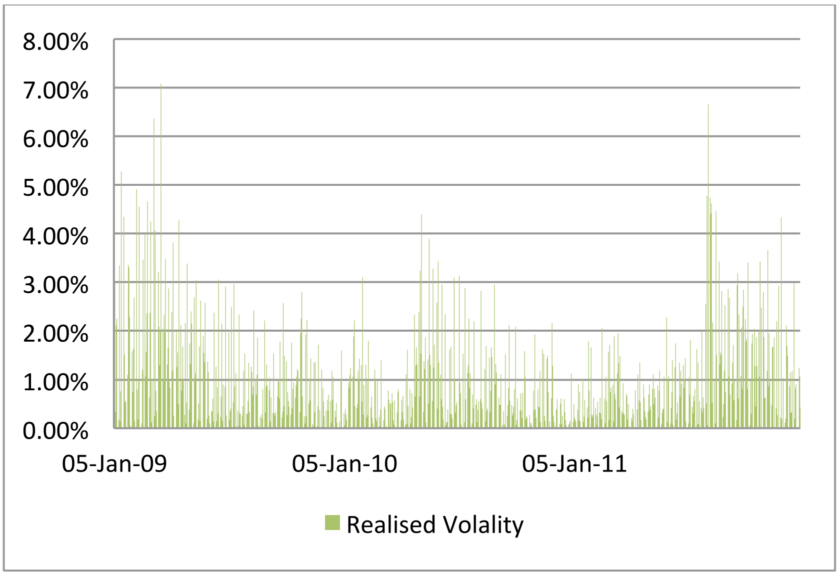

Figure 7. Realised Volatility of the S&P500 2009 – 2011

Table 9 reveals that the most negative news day (NNI of 1) would predict 0.56% more volatility for the following day, whereas the most positive news day (NPI of 1) would predict 0.60% less volatility for the following day.

Implied Volatility				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
NNI	0.006969	0.001559	4.469063	0.000
NPI	-0.004782	0.001258	-3.800194	0.0002
NUI	-0.001947	0.003824	-0.509063	0.6109
ANA	0.011561	0.001288	8.978761	0.000
R-squared	0.042298			

Table 10. Coefficients and Probability Statistics for the Mass Media Sentiment Indicators. Dependent Variable S&P500 Implied Volatility. Period 2009-2011

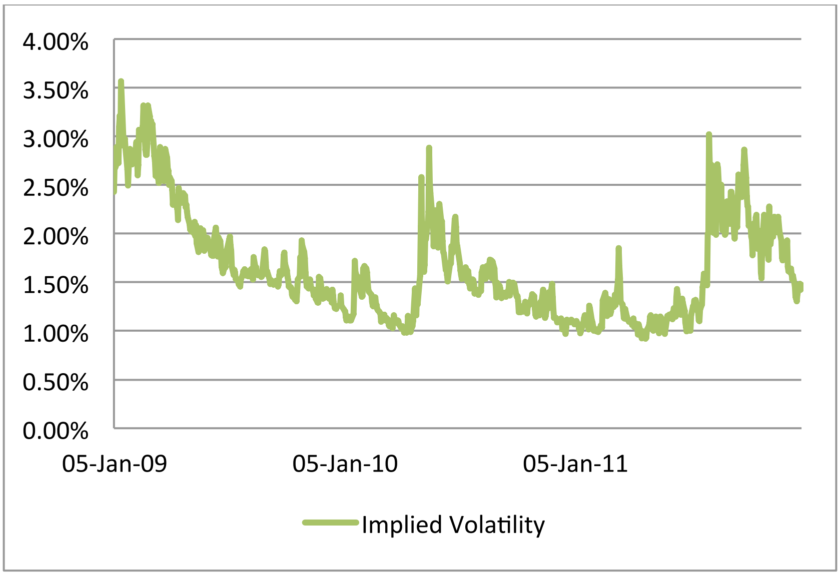

Figure 8. Implied Volatility of the S&P500 2009 - 2011

Table 10 reveals that the most negative news day (NNI of 1) predict 0.69% more implied volatility (t+1), whereas the most positive news day (NPI of 1) would predict 0.47% less implied volatility (t+1).

Descriptive Statistics			
Realised Volality		Implied Volatility	
Mean	1.02%	Mean	1.64%
Median	0.69%	Median	1.51%
Variance	0.01%	Variance	0.00%
Standard Deviaton	1.04%	Standard Deviaton	0.54%

Table 11. Descriptive Statistics for Realised Volatility and Implied Volatility. S&P500 Period 2009-2011

A look at table 11 demonstrates the implications of the results. The mean of realised volatility and implied volatility are 1.02% and 1.64%, respectively. An increase of 0.25 in media negativity as measured by the NNI indicator would predict realised volatility to increase by 0.14% and implied volatility to increase by 0.17%. The aggregate number of articles published on the day had the highest coefficient and an increase of 0.5 in the aggregate number of articles (ANA) indicator would predict 0.5% higher realised and implied volatility the following day. Given that noise-traders only make a fraction of the trades in any market, the mass media effect is surprisingly substantial. Analysing realised volatility on different types of news days reveals the following (see table 12):

Descriptive Statistic	NPI	NNI	NUI	NANA	SPRET
MEDIAN	0.462376238	0.516553413	0.161816401	0.19047619	0.001109
MEAN	0.477127623	0.520929826	0.165935623	0.221336875	0.00052428
MAX	0.926587302	0.923824244	0.416351469	1	0.070758
Standard Dev	0.149866706	0.125993761	0.048536668	0.144728772	0.01455629

Realised Volatility Regression Coefficients	
C (α)	0.006527
NNI (β_1)	0.005658
NPI (β_2)	-0.006089
NUI (β_3)	0.007961
ANA (β_4)	0.010562

$$Y(t) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + u_t$$

Negative day			
NNI X_1	NPI X_2	NUI X_3	ANA X_4
0.797377649	0.339188308	0.156426284	0.476190476
0.015248079	1.52%		
Y(t+1) or realised volatility (t+1) = 1.52%			

Let negativity go down two standard deviations and positivity up two standard deviations

Neutral day			
NNI X_1	NPI X_2	NUI X_3	ANA X_4
0.545390126	0.63892172	0.156426284	0.476190476
0.011997256	1.20%		
Y(t+1) or realised volatility (t+1) = 1.20%			

Let negativity go further down two standard deviations and positivity further up two standard deviations

Positive day			
NNI X_1	NPI X_2	NUI X_3	ANA X_4
0.293402604	0.938655133	0.156426284	0.476190476
0.008746434	0.87%		
Y(t+1) or realised volatility (t+1) = 0.87%			

Table 12 Analysis of a Negative Day, Neutral Day, a Positive Day and Realised Volatility

It is apparent that a change of only one or two standard deviations in the mass media sentiment indicators can have a profound effect on both volatility (t+1) and the volatility expectations (t+1) in the model. Furthermore, the effect is consistent with the sentiment in the media. A negative news day predicts more volatility, characteristic for depressed prices. Positive news days predict less volatility, which in turn correlates with gains in prices. These

results alone cannot determine if markets already adjust for the mass media effect before the closing bell of (t+0), but they demonstrate that the sentiments in mass media (t+0) robustly predict market volatility in period (t+1).

Comparison of 2009-2011 Series and 2013-2014 Series

Comparing the more recent time-series, the calm bull market of 2013-2014 to the turbulent 2009-2011 period, demonstrates similar results although the observed mass media effect is much weaker. As discussed earlier, volatility and daily returns during this period were much lower than during the 2009-2011 period.

Table 13. The Results of the Regressions, Period 2013-2014

SP500 Returns 1st of January 2013 – 15th of July 2014				
Dependent Variable: SP500 Daily Returns				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002171	0.002166	1.002201	0.3169
NNI	-0.000458	0.003312	-0.138202	0.8902
NPI	-0.004498	0.003353	-1.341604	0.1805
NUI	-0.000939	0.003135	-0.299326	0.7649
ANA	0.001538	0.001963	0.78389	0.4336
R-squared	0.006434	Mean dependent var		0.000801

SP500 Realised Volatility 1st of January 2013 – 15th of July 2014				
Dependent Variable: SP500 Realised Volatility				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.0059	0.001382	4.269189	0
NNI	0.002453	0.002114	1.160583	0.2465
NPI	-0.005039	0.00214	-2.354888	0.019
NUI	0.000219	0.002001	0.109591	0.9128
ANA	-0.000129	0.001252	-0.103195	0.9179
R-squared	0.02059	Mean dependent var		0.005216

SP500 Implied Volatility (VIX) 1st of January 2013 – 15th of July 2014				
Dependent Variable: SP500 Implied Volatility				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12.76412	0.545088	23.41663	0
NNI	3.439283	0.833689	4.125378	0
NPI	-2.618586	0.843901	-3.102953	0.0021
NUI	1.424141	0.789125	1.804709	0.0719
ANA	-0.000129	0.001252	-0.103195	0.9179
R-squared	0.104719	Mean dependent var		14.07584

Financial news of the period 2013-2014 exhibited much weaker sentiments over all. The markets were much less volatile. Consequently, the results of the second period are more problematic. In the implied volatility (measured by the VIX index) regression, the MMSIs remain significant, but in terms of realised volatility only the normalised positivity indicator (NPI) remains significant at the 95% confidence level. Lower t- and R-squared values are to be expected, as not only is the sample smaller, but the time-series exhibits very different characteristics. T-values may indicate a different structure of the market, but also the simple fact that as the variance of the dependent variable decreases, it is more difficult to estimate the parameters reliably. The results from the latter time-series point to the same direction (realised volatility, implied volatility correlate with media sentiments), but these results should be met with caution due to a smaller sample size and less variance of all indicators in the time-series.

Introducing Lagged Values as an Independent Variable

If the mass media impact was significant, why would it not be reflected in the prices in (t+0)? The evidence suggests a mass media effect. However, if the closing price of (t+0) already reflects the mass media sentiments, and there is no observable effect in (t+1) that is not already reflected in the prices, one could argue that the mass media effect is insignificant or non-existent or markets are efficient. To study this – and whether prices adjust to reflect the possible media effect in (t+0) – another set of regressions was conducted. The tests were identical but this time each regression had an additional variable:. X_5 is the lagged value of Y_t (the dependent variable). The lagged value was the closing price of the previous day (t+0). In this analysis the adjusted R-squared was used instead of R-squared. The adjusted R-squared has been adjusted for the number of independent variables in the model. This is necessary, because adding irrelevant independent variables increases the R-squared unless it is adjusted for the increased degrees of freedom. The results of this comparison are as follows:

Table 14. Comparing the Adjusted R-Squared

Time-series	SP500 Returns	MMSI+X_5	X_5
2011	Adjusted R-squared	0.92%	0.93%
2014	Adjusted R-squared	-0.40%	0.20%

Time-series	SP500 Realised Volatility	MMSI+X_5	X_5
2011	Adjusted R-squared	3.91%	0.74%
2014	Adjusted R-squared	1.66%	1.09%

Time-series	SP500 Implied Volatility	MMSI+X_5	X_5
2011	Adjusted R-squared	93.49%	93.52%
2014	Adjusted R-squared	78.53%	77.95%

The R-squared of the regressions using both the MMSI and the X_5 lagged value are consistently higher than the ones using the lagged value alone. However, controlling for the increased amount of variables by using adjusted R-squared gives mixed results. For the S&P500 index (2009-2011) and realised volatility, the MMSI+X_5 is significantly better at estimating next day volatility than the lagged value (X_5) would be alone. Normalised negativity, normalised positivity and aggregate number of articles are significant at the 90% level, df. 747. For the 2013-2014 time-series the R-squared indeed increases, but only the NPI remains significant at the 90% level. This is probably due to an insufficient sample and insufficient variance in the sentiment indicators in the latter time-series.

Regarding implied volatility and returns, the MMSIs mostly lose their significance (p-values above the 90% threshold) and the adjusted R-squared levels imply that using MMSIs and lagged value does not improve the model compared to using lagged values alone. Attention should be paid to the extremely high R-squared of implied volatility. The numbers are indeed correct: the high R-squared is consistent with the high autocorrelation of implied volatility.

The F-probability (table 15), indicating the likelihood that the null hypothesis should be accepted for the regressors as a group, gives concurring results. In table 15, values lower than 5% implies a significance level of 95%.

Table 15. The F-Statistics (probability) for the MMSI+X5 Regression and the X5 Regression.

Time Series	SP500 Returns	MMSI+X_5	X_5
2011	F-Stat	3.67%	0.47%
2014	F-Stat	62.98%	18.36%
Time-series	**SP500 Realised Volatility**	**MMSI+X_5**	**X_5**
2011	F-Stat	0.00%	1.04%
2014	F-Stat	4.46%	2.27%
Time-series	**SP500 Implied Volatility**	**MMSI+X_5**	**X_5**
2011	F-Stat	0.00%	0.00%
2014	F-Stat	0.00%	0.00%

EVALUATION OF HYPOTHESES AND DISCUSSION OF IMPLICATIONS

The main research question and the original hypotheses will be evaluated in light of the empirical data and the implications of the results will be discussed. The objective of research was to find evidence for a sentiment based media effect in period (t+1). To return to the hypotheses presented earlier, one can accept H1, H2, H3, H5. The rejected hypotheses are H4 and H6. The hypotheses will be evaluated in the same order as the empirical data was presented, beginning with returns and moving on to volatility and finally concluding with a discussion about the overall mass media effect.

H1: There is correlation between media pessimism in (t+0) and stock market volatility in (t+1).

H2: There is negative-correlation between media optimism in (t+0) and stock market volatility in (t+1).

H3: There is a correlation between news frequency in (t+0) and stock market volatility in (t+1).

H4: The effect is diminished if lagged values of the index analysed are introduced as independent variables but media sentiments improve the explanatory power of the model.

H5: The observed media effect is stronger during bear markets than during bull markets.

H6: Media sentiments predict returns.

Returns and H6

Based on the empirical evidence H6 has to be rejected. The model exhibited no linear correlation between mass media sentiment indicators and the returns of S&P500 or the FTSE100. This result agrees with the efficient market hypothesis and evidence from Fama

(1965) and Samuelson (1965) among others. It disagrees with Clark et al (2004) in their assessment of media's power to influence noise-traders. There are several possible explanations, which might explain why no correlation was found. The first group has to do with the conditions of felicity (Bourdieu, 1991) and networked power (Castells, 2009). First of all, it is possible that theorists like Clark et al (2004) have overstated the effect media has on individual investors. If investors do not regard the media as a trustworthy source, financial news will not have an effect on the investment decisions of noise-traders. This explanation does not imply that the market is efficient or that noise-traders do not trade on sentiments and irrational signals. However, it does imply that mass media do not have a significant role in transmitting noise: it does not possess the sufficient networked power to influence noise-traders; the conditions of felicity (Bourdieu, 1991) are not met.

The other group of explanations is linked to the efficient market hypothesis. Mass media might meet the conditions of felicity and impact noise-traders, but the market is efficient and arbitrageurs immediately exploit the irrational positions of noise-traders adjusting prices to reflect collective rational expectations. Alternatively, it might be that noise-traders make such a small fraction of trades on the market that their irrational trades do not impact the overall market enough for the impact to be observable in this analysis. In conclusion, this result implies that either mass media does not hold sufficient power to impact noise-traders or financial markets react much faster than this analysis could reveal. The latter would suggest that markets are indeed efficient.

However, as discussed earlier, these results must be met with caution. The model is likely to be misspecified in the sense that a non-stationary[20] dependent variable (returns) is explained with stationary independent variables. Non-stationary variables should not be explained with stationary independent variables (Asteriou & Hall, 2008) as the subsequent R-squared and probability statistics will not be reliable. This is not necessarily a problem, if the properties of the time-series so imply. However, as the nature of the time-series for returns and mass media sentiment indicators were not explored, the result should be re-evaluated in subsequent research. Specifically, properties of the error terms should be analysed to evaluate whether they are independent and evenly distributed. The non-stationary process should possibly be transformed into a stationary one by differencing.

Further motivation for the re-evaluation of the results can be found from the analysis of

---

[20] A stationary process is a stochastic process whose mean and variance do not change over time and do not follow any trends.

particularly positive and negative days. In the sample, particularly negative news days were followed by returns of -1.0% the following day, significantly lower than the sample average of 0.05%. Furthermore, positive news (t+0) days predicted better than average returns in (t+1). These results are not reliable because of the sample size of only n=43. However, the results encourage further analysis and it might indeed be that the mass media sentiment indicators predict returns in period (t+1). Nevertheless, based on the regressions it is concluded that this is not the case.

Volatility and H1 to H3

The strongest evidence for a sentiment-based mass media effect comes from the tests explaining volatility and implied volatility in (t+1) using mass media sentiment indicators (t+0). Hypotheses H1 to H3 are accepted. The null hypothesis is rejected based on the t-test for all variables except for the uncertainty indicator. Mass media positivity, negativity and the aggregate number of articles predict volatility consistent with the theory presented and previous empirical results (Tetlock, 2007; Campbell, 1992; De Long, 1990). The F-tests imply that not only are the independent variables significant individually, but also as a group. The probability statistics for implied volatility are likely to be inflated, but not to an unreasonable extent.

This evidence contradicts with the efficient market hypothesis, and agrees with the theoretical account of Clark et al (2004) along with the theoretical framework of this study. Mass media sentiments clearly impact noise-traders decisions and expectations. As financial news cannot be generally considered genuine new information, this effect has to be based on the networked power of mass media and the performativity of financial news. As information is disseminated from the right position in the communicative network of noise-traders, it has an impact on the markets beyond what the 'genuine news' would justify. Furthermore, it should be noted that this result implies that the 'media effect' is not only convenience at work (i.e. investors investing in assets that have strong brand recognition). Financial literature suggests that high volatility is usually paired with depressed returns. Consequently, the fact that negative news increase volatility suggests that noise-traders decisions do indeed follow media sentiments.

Nonetheless, these results should be re-evaluated in future research. The lower p-values and higher coefficients of the implied volatility model suggest that performativity might have a bigger impact on collective expectations than on collective actions. However, due to the autocorrelation in the model this result should be met with caution. Robust probability

statistics that consider the high autocorrelation in implied volatility should be calculate to adjust for the autocorrelated nature of implied volatility.

H4: Evidence from Using MMSI and Lagged Values

Based on this analysis, H4[21] has to be rejected, except for realised volatility. H4 arose from a lack of studies using lagged values (for example, Tetlock 2007) and therefore, a lack of empirics regarding to what extent and how quickly the mass media effect is absorbed in prices in (t+0). The results are inconclusive.

H4 was tested by a comparison of adjusted R-squared of the regressions using mass media sentiment indicators and lagged values of the index to regressions that only used the lagged values. Should the efficient market hypothesis hold true and stock markets truly follow a random walk, the best estimate of S&P500 volatility in period (t+1) would be the volatility of S&P500 in period (t+0). However, as adding the mass media sentiment indicators into the regression improves the capability of the model to predict realised volatility, it is evident that not only do the mass media have an impact, but this impact is not fully absorbed into prices in (t+1).

The lag is considerably long, as the stock market is open for 6.5 hours every day. The delay from the news in period (t+0) to the closing bell of (t+1) is from 12 to 24 hours. Financial news continue to have an effect on volatility for hours after they have been published by mass media. This result is consistent with the notion of mass media nodes transmitting a message in the communicative network of noise-traders. The message is relatively slowly disseminated, evaluated and reflected in the collective decisions of noise-traders. This result is consistent not only with the theoretical account of Clark et al (2004) and the framework of this study, but also with the empirical results from Tetlock (2007), Fang and Peres (2009) and Antweiler and Frank (2004).

When analysing realised volatility, the 'media effect' i.e. the mass media sentiment indicators remain relevant. When paired up with the lagged values, they allow a model to explain a greater portion of realised volatility (t+1) than using the lagged values alone. It is therefore the conclusion of this analysis that mass media sentiments have an effect on realised

[21] H4 : The mass media effect is diminished if lagged values of the index analysed are introduced as independent variables but media sentiments improve the explanatory power of the model

volatility. Prices do not immediately adjust to reflect this effect, conversely to what the efficient market hypothesis would suggest. As for returns and implied volatility, a more sophisticated model both in terms of regressions and in terms of media sentiment analysis should be constructed for any conclusive evidence to be found.

H5: Comparing Bulls and Bears

Comparing the two time-series allows one to accept H6[22]. It is obvious that the mass media effect is stronger during bear markets. This result is similar to previous empirical results (e.g. Hardouvelis & Theodossiou, 2002). This can be observed on a daily basis, as the negative impact of pessimistic news on returns is much stronger than the positive impact of optimistic news. Furthermore, this can be observed on a longer time scale as well, since the more recent time-series exhibited a much lower media effect in general, across all tests. However, this could to an extent be due to measuring errors arising from the nature of the second time-series (low variance in dependent and independent variables). The sample size might have been insufficient and the content analysis incapable of revealing the nuances of the calmer period 2013-2014.

The stronger impact of negative news is evidence for the performativity of financial reporting. As the notion of performativity suggests, negative news create more days of depressed returns and high volatility in the markets. The effect is much weaker during periods of optimism. A growing body of research suggests that negative news has a disproportionate impact on individuals' attitudes (Soroka, 2006). This suggests that performativity is stronger when the forecast is negative. It is an empirical fact that bear markets tend to be more intense and shorter in duration than bull markets. An economic explanation for the asymmetric and stronger response to negative news is risk-aversion of noise-traders and arguably it is likely that both of these phemomena are at work.

CONCLUSIONS

This study evaluated the influence of mass media sentiments on noise-traders in the broader theoretical context of communicative networks influencing financial markets and decision-making. The efficient market hypothesis (Samuelson, 1956; Fama, 1965) was juxtaposed with the notion of performativity arising from sociological literature (Bourdieu, 1991; MacKenzie,

[22] H5: The observed media effect is stronger during bear markets than during bull markets

2003, 2006). The aspiration of the study was to evaluate whether financial news influences the market and suggest a mechanism of action. The sentiment analysis on 7000 financial news articles and the subsequent regression analysis demonstrate the importance of studying mass media and the communicative networks of noise-traders. The most important insight this study has provided is the connection between realised volatility and media sentiments. Even when adjusting for prices in (t-0), mass media sentiments robustly predict market volatility in (t+1). The correlation intuitively followed the sentiments (negative days predicting depressed markets). Consequently, it was argued that the mass media effect is not based on convenience alone. Within the communicative network of noise-traders, mass media affects noise-traders collective expectations and actions, subsequently having an impact on financial markets that goes beyond disseminating new information.

The evidence regarding the effects of mass media sentiments on market returns and implied volatility was less conclusive. In terms of returns, the F-test and lack of autocorrelation indicate that in terms of mass media sentiments and returns, markets are indeed efficient. The strong autocorrelation in the model for implied volatility renders the results inoperable. Consequently, further analysis is needed. Despite being inconclusive, regressions on returns and implied volatility exhibit characteristics that suggest a mass media effect consistent with the theory presented as well as prior research.

This seems to indicate that the efficient market hypothesis should be re-evaluated and evidences the performativity of financial news. Consequently, the communicative network of noise-traders emerges as a topic for future research. However, as the methodological problems of this study and previous research demonstrate, future research should apply rigorous econometric analysis controlling for heteroscedasticity and carefully analyse the statistical properties of the time-series. The content analysis could be improved by emphasising different articles more than others, adding data from different media, using a more sophisticated algorithm to compute sentiments and adjusting for the change in sentiment rather than the absolute value. Moreover, it would be ideal to conduct the entire analysis on an hourly basis. However, hourly analysis is challenging, as neither financial data nor sentiment data is readily available. Finally, the intercoder reliability tests conducted should be improved to encompass a degree of agreement instead of using binaries.

Despite the need for numerous improvements, this study has demonstrated the importance of studying the communicative networks of noise-traders. The empirical evidence evaluated suggests that mass media have an effect on the market. The mass media effect is not absorbed into prices as fast as the efficient market hypothesis would suggest. Furthermore, it

trails media sentiments in an intuitive fashion demonstrating the performativity of financial news.

REFERENCES

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors.

Andreassen, P., & Kraus, S. (1987). Intuitive prediction by extrapolation: The stock market, exponential growth, divorce, and the salience of change. *Mimeographed. Cambridge, Mass.: Harvard University*.

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

Asteriou, D., & Hall, S. G. (2007). Applied Econometrics: a modern approach using eviews and microfit. New York: Palgrave Macmillan

Austin, J. L. (1975). How to do things with words (Vol. 1955). Oxford University Press.

Barnes, B. (1983). Social life as bootstrapped induction. *Sociology*, 17(4), 524-545.

Bourdieu, P. (1991). Language and symbolic power. Harvard University Press.

Callon, M. (1998). Introduction: the embeddedness of economic markets in economics. The Sociological Review, 46(S1), 1-57.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1992). Trading volume and serial correlation in stock returns (No. w4193). National Bureau of Economic Research.

Carletta, J. (1996). Assessing agreement on classification tasks: the kappa statistic. *Computational linguistics*, *22*(2), 249-254.

Castells, M. (2007). Communication, power and counter-power in the network society. *International journal of communication*, *1*(1), 29.

Castells, M. (2009). Communication Power. Oxford, New York: Oxford University Press.

Chandler, D. (1994). The transmission model of communication. University of Wales, Aberystwyth. Online: http://www. aber. ac. uk/media/Documents/short/trans. html.

Clark, G., Thrift, N., & Tickell, A. (2004). Performing finance: the industry, the media and its image. *Review of International Political Economy*, 11(2), 289-310.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, 703-738.

Dimitrova, D. V., & Strömbäck, J. (2005). Mission accomplished? Framing of the Iraq War in the elite newspapers in Sweden and the United States. *Gazette*, *67*(5), 399-417.

Elliott, G., & Timmermann, A. (Eds.). (2013). *Handbook of economic forecasting* (Vol. 2). Newnes.

Fama, E. F. (1965). The behavior of stock-market prices. *Journal of business*, 34-105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work*. *The journal of Finance*, *25*(2), 383-417.

Fang, L., & Peress, J. (2009). Media Coverage and the Cross-section of Stock Returns. The Journal of Finance, 64(5), 2023-2052.

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical methods for rates and proportions*. John Wiley & Sons.

Frieder, L., & Subrahmanyam, A. (2005). Brand perceptions and the market for common stock. *Journal of financial and Quantitative Analysis*, 40(01), 57-85.

Granovetter, M. (1985). Economic action and social structure: the problem of embeddedness. *American Journal of Sociology*, 481-510.

Gujarati, D. N. & Porter D. C. (2009). Basic Econometrics. Fifth Edition, McGraw-Hill.

Haller, B., & Ralph, S. (2001). Not worth keeping alive? News framing of physician-assisted suicide in the United States and Great Britain. *Journalism Studies*, *2*(3), 407-421.

Hansen, G. S., & Hill, C. W. (1991). Are institutional investors myopic? A time-series study of four technology-driven industries. *Strategic Management Journal*, 12(1), 1-16.

Hanson, E. C. (1995). Framing the world news: The Times of India in changing times. *Political Communication*, *12*(4), 371-393.

Hardouvelis, G. A., & Theodossiou, P. (2002). The asymmetric relation between initial margin requirements and stock market volatility across bull and bear markets. *Review of Financial Studies*, 15(5), 1525-1559.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, Addison-Wesley.

Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. *Journal of financial economics*, *6*(2), 95-101.

Kim, Y. H., & Meschke, F. (2011). CEO interviews on CNBC. In *Fifth Singapore International Conference on Finance*.

Kirstensen, N., Orsten, M. (2007). Danish media at war. *Journalism*. 8(3). 323-343.

Kleidon, A. W. (1986). Anomalies in Financial Economics: Blueprint for Change?. *Journal of Business*, S469-S499.

Klibanoff, P., Lamont, O., & Wizman, T. A. (1998). Investor Reaction to Salient News in Closed-End Country Funds. *The Journal of Finance*, 53(2), 673-699. Chicago

Cetina, K., & Preda, A. (2004). *The sociology of financial markets*. Oxford University Press.

Krippendorff, K. (1980). *Content analysis*. Beverly Hills, California, Sage Publications. *7*, l-84.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

LeRoy, S. F., & Porter, R. D. (1981). The present-value relation: Tests based on implied variance bounds. *Econometrica: Journal of the Econometric Society*, 555-574.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35-65.

MacKenzie, D. (2003). Long-Term Capital Management and the sociology of arbitrage. *Economy and Society*, *32*(3), 349-380.

MacKenzie, D. (2006). Is economics performative? Option theory and the construction of derivatives markets. *Journal of the History of Economic Thought*, *28*(1), 29-55.

MacKenzie, D., & Millo, Y. (2003). Constructing a market, performing theory: The historical sociology of a financial derivatives exchange1. *American journal of sociology*, *109*(1), 107-145.

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, 11(1), 193-197.

Matthes, J., & Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. *Journal of Communication*, *58*(2), 258-279.

Marsh, T. A., & Merton, R. C. (1986). Dividend variability and variance bounds tests for the rationality of stock market prices. *The American Economic Review*, 483-498.

Marshall, B. R., & Cahan, R. H. (2005). Is technical analysis profitable on a stock market which has characteristics that suggest it may be inefficient?.*Research in International Business and Finance*, 19(3), 384-398.

Marshall, B. R., Qian, S., & Young, M. (2009). Is technical analysis profitable on US stocks with certain size, liquidity or industry characteristics?. *Applied Financial Economics*, 19(15), 1213-1221.

Neuendorf, K. A. (2002). *The content analysis guidebook*. London, Sage.

Parsons, T., & Smelser, N. J. (2003). *Economy and society: A study in the integration of economic and social theory* (Vol. 4). Psychology Press.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2), 41-49.

Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *The Journal of Economic Perspectives*, 19-33.

Shiller, R. J. (1980). Do stock prices move too much to be justified by subsequent changes in dividends?. NBER Working Paper No. 456 Issued in February 1980.

Soroka, S. N. (2006). Good news and bad news: Asymmetric responses to economic information. *Journal of Politics*, *68*(2), 372-385.

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.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, *76*(2), 105.

Wanta, W., Golan, G., & Lee, C. (2004). Agenda setting and international news: Media influence on public perceptions of foreign nations. *Journalism & Mass Communication Quarterly*, *81*(2), 364-377.

Whaley, R. E. (2000). The investor fear gauge. *The Journal of Portfolio Management*, 26(3), 12-17.

White, H. C. (1981). Where do markets come from?. *American Journal of Sociology*, 517-547.

Electronic MSc Dissertation Series