QUCEH WORKING PAPER SERIES
http://www.quceh.org.uk/working-papers

NEWS MEDIA AND INVESTOR SENTIMENT OVER THE LONG RUN

Alan J. Hanna (Queen's University Belfast)
John D. Turner (Queen's University Belfast)
Clive B. Walker (Queen's University Belfast)

Working Paper 2017-06

QUEEN'S UNIVERSITY CENTRE FOR ECONOMIC HISTORY
Queen's University Belfast
185 Stranmillis Road
Belfast BT9 5EE
August 2017

News Media and Investor Sentiment over the Long Run*

Alan J. Hanna
Queen's University Belfast

John D. Turner
Queen's University Belfast

Clive B. Walker†
Queen's University Belfast

Abstract

This paper studies the effect of investor sentiment on the London stock market on a daily basis over the period 1899 to 2010. We use a broad mix of reporting from the *Financial Times* as our proxy for investor sentiment. The main contribution of this paper is threefold. First, newspaper commentary, which was sentiment-laden, but information-light, in the *Financial Times* affects returns. Second, we find evidence that sentiment plays a role in propagating price movements, particularly during bull markets. Third, we find little evidence that the effect of sentiment on the market differs in bear versus bull markets.

JEL Classification: G12, N23, N24

Keywords: News media, investor sentiment, stock market, bull, bear

* Thanks to Carsten Burhop, David Chambers, Narly Dwarksing, Richard S. Grossman, Les Hannah, Lyndon Moore, Mike Staunton, participants at the 2016 EurHistock Workshop, the Queen's Management School Finance Seminar, the 8th World Congress of Cliometrics and the Macrohistory Research Seminar at Bonn University for their comments. Thanks also to Gareth Campbell for help with the *Investor's Monthly Manual* data.

† Corresponding author. Address: Queen's Management School, Queen's University Belfast, 185 Stranmillis Road, Belfast BT9 5EE, Northern Ireland, UK. Email address: c.walker@qub.ac.uk.

1. Introduction

The important role of the press in shaping behavior in financial markets was highlighted by Bagehot (1873) in his famous *Lombard Street*. Economists, implicitly accepting this fact, have thus used the press to understand the news that moves financial markets (Niederhoffer, 1971; Cutler et al., 1989; and Elmendorf et al., 1996). More recently, however, economists have begun to examine how the press itself, rather than the news it reports, influences financial markets. Shiller (2000, 2017) argues that the press play an important role in propagating stock market booms and market sentiment through their narratives and hype. Tetlock (2007) formally tests this role of the press and finds that the negativity of reporting in a *Wall Street Journal* market column predicts daily stock returns. Similarly, García (2013) finds that the negativity and positivity of reporting in two financial columns in the *New York Times* over a century predicts daily stock returns, but that the effect is concentrated in recessions.

This paper contributes to this literature on the role of the press by studying a broad mix of financial and non-financial reporting in the *Financial Times* (FT) from 1899 to 2010. Because we examine a range of types of press reporting, including explicit opinion or commentary pieces, we are able to see whether particular types of reporting affect markets more and whether sentiment-laden but information-light reporting affects markets. We find that sentiment-laden and information-light commentary predicts stock returns and the number of trades on the market, which is strong evidence to suggest that pure newspaper hype is moving markets. In addition, because we study such a long period of time, we are able to see whether the effect of press reporting on returns is associated with the prevailing state of the stock market, i.e., whether it is in a bull or bear state. We find that the effect of negative reporting on returns does not differ across market states, but that the effect of positive reporting on returns is twice as large during bear markets. We

also find that, the effect of reporting on returns tends to be non-reversing in both bull and bear markets, which is evidence consistent with Shiller's press-as-propagators hypothesis.

In this paper, we use four principal sections from the FT—Editorials and Leaders, Lex, News in Brief, and Market news columns—and count the number of positive and negative words used in each section to generate a daily proxy of investor sentiment. The sections we include are broader than those used in previous studies in that they cover more than stock-market news. They also include explicit commentary and opinion pieces, which are almost perfect measures of market sentiment. In addition, the columns we include are unlikely to contain new information on individual companies; this is particularly the case for the FT's two commentary/opinion sections.

The FT has been the UK's premier and dominant financial newspaper for well over a century. During its early years, the *Financial News* was its main competitor, but this newspaper's circulation was only one quarter of the FT's by the 1910s and eventually, in October 1945, it merged with the FT. Consequently, given this dominance, the FT was the main newspaper read by investors and participants in financial markets and therefore if any newspaper was going to reflect investor sentiment, it would be the FT.

In order to test the effect of investor sentiment, as proxied by the FT's reporting, on stock returns, we create a daily blue-chip index for the UK market. From 1930 onwards, we use pre-existing blue-chip indices. Prior to 1930, we collect the relevant data to construct a daily index of the largest stocks in the UK in an attempt to imitate the post-1930 daily blue-chip index. Consequently, this paper presents the first daily stock-market index for the UK dating back to 1900. We also collect data on trading activity for our blue-chip index as far back as 1930. All our results are robust to the exclusion of the pre-1930 period.

Investor sentiment is highly correlated with the prevailing state of the stock market (Baker and Wurgler, 2007). Thus, the effect of investor sentiment on the market may vary with the market's prevailing state. If the stock market is in a bullish/bearish mood, positive/negative reporting by the press may exacerbate or dampen the overall mood of the market. While there is no formally agreed definition for 'bull market' and 'bear market' phases, these terms have become commonplace and are widely understood to mean periods of time when market prices are generally rising or falling respectively (Chauvet and Potter, 2000). Google's Ngram viewer reveals these specific terms were a well-established part of the financial lexicon by the end of the nineteenth century. These terms capture and express something akin to prevailing market sentiment. By contrast, business cycles capture aggregate activity across an entire economy. Moreover, the economic data used to define business cycles is released only periodically, lacks immediacy, is backwards looking, is subject to subsequent and frequent revision, and ultimately may not matter to investors—markets can sustain high returns during economic downturns and *vice versa*. We therefore test whether the overall state of the stock market affects investor sentiment as well as testing Shiller's (2000) press-as-propagators hypothesis and García's (2013) hypothesis, based on the personality and psychology literature, that newspaper sentiment has a greater effect during periods of hardship. To identify whether the market is in a bull or bear state, we develop algorithms based firstly on the performance of our blue-chip index and secondly on the FT's use of 'bullish' words.

Our findings suggest that there is some evidence to support Shiller's hypothesis. First, positive sentiment neither reverses in bull or bear states. Second, for the overall period, the initial effect of negative sentiment and pessimism (which is the negative minus the positive sentiment) is reversed over subsequent days. However, when we analyze the media effect by market state, we

find that negative sentiment and pessimism do not reverse in bear markets, suggesting that the media are propagating this market state. Third, we find that extreme positive sentiment during bull markets shifts trading activity, which is consistent with noise traders driving trade during speculative booms (Baker and Stein, 2004). We also discover some evidence to support García's hypothesis. Both positive sentiment and pessimism have a greater effect during bear markets than during bull markets. However, there is no evidence that negative sentiment has a greater effect during bear markets.

Our analysis of the different sections of the FT reveals several interesting findings. First, the more narrowly focused the section is on financial markets, the greater its influence on returns. Second, the words used to report on financial markets and the commentary therein are just as sentiment-laden as the pure commentary sections of the FT. Third, the sentiment of the pure commentary sections of the FT affect returns, suggesting that the sentiment in our media content measures are uncontaminated by information.

As well as using the various sections of the FT to ascertain whether our sentiment measures are contaminated by information, we test for the presence of information in our sentiment measures by exploring whether the effect of sentiment differs between consecutive and non-consecutive trading days. The premise of doing this indirect test is that information accumulates over non-consecutive trading days or, alternatively, information production is concentrated on consecutive trading days. Our results suggest that the effect of sentiment on returns between consecutive and non-consecutive trading days does not differ. However, the effect on non-consecutive days does not reverse, which may suggest that our sentiment measure on Mondays may be contaminated with some information.

We also examine whether or not the effect of our sentiment measures on the market reversed over subsequent days. If our sentiment measures contain information, then we would anticipate that the initial effect on returns would not be reversed because new information has been impounded into the market. However, if our measure is sentiment, then we would expect the initial impact upon returns to be reversed because after the sentiment shock, noise traders are prompted to sell or buy stocks, rational traders will subsequently drive prices back to their fundamental value unless there are limits to arbitrage which prevent them from doing this. On the whole, our evidence points in the direction of a reversal of negative sentiment shocks, which suggests that our negative sentiment measure is uncontaminated by information. However, positive sentiment shocks are not reversed, which could point to information in positive sentiment or limits to arbitrage with upward stock movements.

This paper is part of a growing literature which examines the role of the press in financial markets. This literature can be categorized into distinct areas, namely the role of the press as a financial watchdog (Dyck et al., 2008; Taylor, 2012); the role of the press as information providers to investors (Fang and Peress, 2009; Engelberg and Parsons, 2011; Griffin et al., 2011; Turner et al., 2017); the role of bias and incentives in the financial press and financial reporting (Dyck and Zingales, 2003; Bignon and Miscio, 2010); the role of the media in financial bubbles or speculation (Shiller, 2000; Huberman and Regev, 2001; Bhattacharya et al., 2009; Campbell et al., 2012); and the use of media content as a way of quantifying animal spirits or sentiment in the stock and housing markets (Tetlock, 2007; Dougal et al., 2012; García, 2013; Soo, 2013; Walker, 2014; Kräussl and Mirgorodskaya, 2017). The part of the literature which this paper is most closely related to is the sentiment work of Tetlock (2007) and García (2013). Similar to both of these papers, we find that newspaper sentiment has an economically and statistically significant

relationship with the returns on that day's stock market index. A one standard deviation rise in negative sentiment predicts a 4.3-basis point decrease in our blue-chip index and a 3.3-basis point increase in our blue-chip index for positive sentiment, which is of a similar order of magnitude to the effect found by both Tetlock (2007) and García (2013). However, unlike these two papers, we are able to use different types of reporting to produce better measures of sentiment, which are potentially less contaminated by information. We also produce stronger evidence in favor of Shiller's hypothesis that the press propagate upward swings in the market through their positive sentiment during bull markets. Finally, unlike previous studies, we test whether the prevailing state of the market affects the impact of sentiment on returns.

The paper also augments the investor sentiment literature, which suggests that some investors are not fully rational and that their demand for stocks is driven by sentiment which is not justified by fundamental news (De Long et al., 1990; Shleifer and Summers, 1990; Barberis et al., 1998; Brown and Cliff, 2005; Baker and Wurgler, 2006, 2007). For example, Kamstra et al. (2003) and Edmans et al. (2007) document how the weather and soccer results affect stock prices. We find that news media content, our proxy for investor sentiment, predicts stock returns as well as trading activity. Shleifer and Vishny (1997) argue that limits to arbitrage can prevent rational investors from driving prices back to fundamental levels. We find some evidence to suggest that stocks are not driven back to fundamentals, particularly during bull markets.

The rest of the paper is structured as follows. The next section describes our media and stock-market data. Section three examines whether FT sentiment affects the stock market. Section four asks whether there is information in our sentiment measure. Section five compares the effect of media sentiment across bull and bear markets. Section six compares the effect of the sentiment of the various sections of the FT newspaper upon stock returns. The final section concludes.

2. Data

2.1 News media

The *Financial Times* (FT) was first published in 1888. Facsimiles of the original print editions are available from the *Financial Times Historical Archive* (FTHA).[1] Our analysis commences in 1899 due to the low quality of the print in the FT prior to this period. The FTHA does not have articles after 2010, and though alternative sources are available after this date, they lack the section classification used in our analysis, and are thus not suitable. This 111-year sample provides an opportunity to consider the role of the media over a long period and through many different market cycles. Our sample period coincides with the era of 'New Financial Journalism', which recognized the responsibility of editors to inform the vastly increased number of investors (Porter, 1998).

Using the FT to quantify investor sentiment over the long run has several major advantages. Firstly, the UK is one of the world's largest and most established financial markets, with long standing freedoms for both markets and press alike. This reduces the chance of bias. Secondly, the FT is a daily financial newspaper, with specialist reporters and a reputation for high-quality journalism. As such, its content should more accurately reflect market activity and prevailing investor sentiment. Thirdly, the FT has a large circulation and a readership actively engaged in financial markets.[2] The first issue of the paper ran with the banner, "[This paper is] the friend of the honest financier, bona fide investor, respectable broker, genuine director and legitimate

[1] http://find.galegroup.com/ftha

[2] Kynaston (1988) reports average daily net sales in the early 1900s of about 12,000, rising to over 250,000 by 1986. The FT currently claim an 'average daily global audience' of 2.1 million across print and digital platforms. See http://aboutus.ft.com/corporate-information/history/

speculator". More recent demographic surveys confirm a professional and wealthy readership.[3] As a result, the potential for this news medium to influence market activity is high, making the FT a prime candidate for testing the effect of media coverage on investors and the stock market. Fourthly, while the format and content of the paper has evolved over time, the culture provided by long-standing features, journalists and editors provide considerable continuity and the make the FT a useful vehicle for a long-run analysis of media sentiment.

The final, and possibly most important, advantage of using the FT is that the FTHA classifies all articles. This allows us to identify articles which are focused primarily on the stock market as well as articles which provide wider macroeconomic and financial news. It also allows us to identify articles which are editorials and therefore provide explicit commentary which tends to speculate about the immediate past and the near future. For this study, we selected articles from the four areas of the FT which are most pertinent to the stock market and which are most likely to be read by traders. This approach means that we focus on those parts of the FT which are most likely to reflect investor sentiment in the London market. It also means that we ignore firm-specific reporting by the FT, which is much more likely to contain new information than the sections of the newspaper included in this study.

The four sections of the FT used in this study are as follows. Editorial articles were those classified within the FTHA as 'Editorials and Leaders'. Editorials are typically limited to two or three separate subjects per edition and tend to focus on key political and economic issues of the day. Lex articles were classified as 'Lex column' by the FTHA. The Lex column is described by the FT as setting 'the agenda on everything from company analysis and macroeconomics to

[3] http://fttoolkit.co.uk/d/audience/

financial markets to critical trends of the day'.[4] This feature first appeared in October 1945, following the FT's merger with its rival publication the *Financial News*. These opinion pieces in the FT's editorials and Lex column were written on a rotating basis by a small team under the direction of an editor, possibly providing some continuity to any measure of media sentiment.

'News in Brief' articles represent another FTHA classification and contained articles summarizing general macroeconomic and political news. Finally, stock-market articles were classified as 'Market News', with the additional criteria that articles contained the word 'market' in the title. These articles reported on what had been happening in the stock market during previous trading sessions.

In terms of any potential influence on the stock market, it is important to establish when the news in the FT was in the hands of investors and traders. In the case of the FT, the physical paper was published, printed and distributed prior to the opening of UK financial markets (Ferguson et al., 2015). The FT.com was launched in 1995, and some articles published in the print version of the newspaper were available online before market close of the preceding trading day. However, this is not an issue for this study because the four sections we examine were not available until well after the close of the markets.

Articles from the four categories were downloaded from the FTHA in PDF format and then converted into text files using ABBYY Optical Character Recognition (OCR) software. All subsequent processing and analysis of the text files was performed using bespoke Python programs developed by the authors. The main advantage of doing his compared with using an off-the-shelf textual analysis package is the ability to tailor the approach to address issues relating to the specific

[4] http://www.ft.com/lex (see 'About Lex')

data set. Indeed, Loughran and McDonald (2015) highlight various problems with using generic text analysis software.

Before sentiment analysis was conducted, a process of data cleansing was performed to address as many of the digitization issues and OCR errors as possible. Any errors remaining after the data cleansing process should be random and therefore have no effect on our results. Three major areas were corrected. Firstly, a list of substitutions was compiled to address common OCR errors and improve document quality. These were constructed based on a review of initial word counts to target high frequency issues, and then performed on a 'find and replace' basis. For example, substituting instances of the misspelt word 'diflicult' with the word 'difficult'. Such substitutions naturally arose from commonly occurring OCR errors such as confusion between the letters 'e'/'c'/'a', 'f'/'l'/'t', 'h'/'b', 'y'/'v', confusion between 'e' and the copyright symbol © and finally, confusion between the letter 'O' and the number '0'.

Secondly, we corrected hyphenated words used to accommodate newspaper justified formatting rules. The hyphen character was often omitted in the OCR translation or included along with spurious white space or random characters that did not appear in the original article. For example, the word 'demand' could easily be captured as 'de-mand', 'de mand' or even 'de-;mand'. Such errors lead to two separate words which may or may not be valid dictionary words. [5] A list of commonly split words was compiled. These were recombined when their two halves were

[5] Following Loughran and McDonald (2011), a base dictionary was compiled from the '2of12inf' word list available from http://wordlist.aspell.net. This was extended by including the '3of6all' word list to incorporate British-oriented dictionaries.

separated by selected characters associated with hyphenation errors. Hyphens were also removed when doing so created a valid dictionary word.

Thirdly, we filtered the content of articles. The OCR process produced text files separated into 'lines' that would naturally be delimited by carriage returns. Thus, a single line of text could represent an entire article paragraph. Filtering was used to exclude entire lines of text that did not directly relate to the article, or meaningfully contribute to its sentiment. Firstly, boiler-plate content that appears on a regular basis was removed. For example, the FT's 'without fear and without favor' motto. Secondly, text that was evidently not part of the article itself was removed. This issue relates to the image capture process used to digitize the FT. Thirdly, lines containing low quality or significant amounts of spurious text were removed.[6] This seeks to address issues arising from the digitization of non-text content such as tables, graphs and photographs embedded in the article.

We used a process of tokenization to decompose text into the single words and punctuation characters from which sentences are constructed. The Natural Language Toolkit Python package was used for this purpose.[7] While the comma and period character are typically used for punctuation, in financial publications they are frequently also used for abbreviation (e.g., U.S.A.), as a decimal point (e.g., 5.2%) and as a thousand separator (e.g., £10,000). To prevent excessive tokenization, allowances were made for such occurrences. A 'token' could therefore refer to any contiguous set of alphanumeric or symbolic characters that form a natural group, but is most often synonymous with a word.

---

[6] Criteria to exclude such lines were: 1. contain less than 5 words; 2. valid words make up less than 50% of the total word count; 3. valid words contribute less than 50% of the total character count of the line; 4. have more than 50% of letters in upper case; 5. contain more than one tab character.

[7] See http://www.nltk.org/

Tokens consisting of single punctuation characters, or those representing numeric quantities, clearly contain no sentiment and can therefore be ignored. Tokens matching entries in a standard dictionary were considered valid tokens. Remaining tokens consisted of valid words that did not appear in the standard dictionary, proper nouns, and tokens containing non-alphabetic characters. These remaining tokens were first sorted into descending order of frequency. Unclassified tokens occurring more than 30 times in any one of the four sections of the FT were manually inspected. Where appropriate, these tokens were added to an extended custom dictionary or, in the case of obvious OCR errors, appropriate substitutions were created. We reviewed the list to confirm their validity and spot checks were performed against the original source documents. Examples of words added to the custom dictionary include terms specific to finance, those with distinct British spellings, and some archaic terms. A total of 564 words were added to the custom dictionary through this process. Tokens appearing less frequently than the above threshold would be expected to appear on average at most every three years, thus reflecting Zipf's (1949) law.

The final step in the process was to count the number of tokens in each article expressing a particular sentiment. While earlier studies use the Harvard Psychosocial Dictionaries, Loughran and McDonald (2015) argue that use of non-specialist lists is inappropriate within a financial content. Consequently, sentiment categories are taken from Loughran and McDonald (2011). Of the 2,329 (354) words expressing negative (positive) sentiment in their dictionary, 38 (19) were modified to reflect the FT's use of British English (e.g., replacing 'penalize' with 'penalise').

The sentiment expressed by a word can be negated within the context of a larger phrase or sentence. To account for this, a list of negation terms was compiled. The six negation terms used by Loughran and McDonald (2011) were extended to include 'nor', 'cannot', 'lack of', 'loss of', 'end of' and contractions of words with the word 'not' (e.g., couldn't). The tallies of words

expressing a particular sentiment that were preceded by such negation terms were decremented rather than incremented.

Total and sentiment word counts were compiled at the article level and aggregated by publication date. Sentiment measures were then normalized by dividing the aggregated scores by the total sum of dictionary recognized words from the corresponding articles. Thus, sentiment *j* aggregated for dates *T* is:

$$s_j(T) = \frac{\sum_{t(i)\in T} as_{ij}}{\sum_{t(i)\in T} wc_i} \qquad (1)$$

where as_{ij} is the score for sentiment *j* in article *i* with total dictionary word count wc_i.

Table 1 contains the descriptive statistics for daily sentiment measures. The average and median positive sentiment is just under 1 per cent, whilst the average and median negative sentiment is 1.64 and 1.59 per cent respectively. The pessimism sentiment, which is simply the negative minus the positive sentiment, has a mean and median of 0.69 and 0.66 per cent respectively. These mean sentiment scores are slightly lower than those found by García (2013) for the *New York Times* financial column—1.2 for positive sentiment and 2.1 for negative sentiment. These differences could reflect cultural differences or the narrow financial media focus of García's study. Table 1 also shows that in bull (bear) markets positive sentiment is higher (lower) and negative sentiment lower (higher) than in bear (bull) markets.

<<<INSERT TABLE 1 HERE>>>

2.2 Stock-market data

We use a daily blue-chip index to assess the impact of media sentiment on the stock market. We use the FT30 index from 1 July 1935 to 2010; it is an unweighted price index of 30 leading

industrial and commercial shares and is the UK's oldest continuous daily index, dating to 1 July 1935. For the period January 1930 to June 1935, we use the forerunner to the FT30—the *Financial News* 30 or FN30, which was rebased to match the FT30. Replicating the FT30 before 1930 is not possible because the constituents of the index were chosen by journalists at the FT on an annual basis based on how widely held and popular the shares were. Instead, we create an equally-weighted daily index of the 30 largest companies listed in the *Investor's Monthly Manual* (IMM) and that had daily data in Global Financial Data (GFD). We identify companies with daily stock prices on GFD before 1930 and calculate log returns adjusting for weekends and non-trading days. These companies are then matched with those in the IMM. For each year from 1899, we use the December issue of the IMM to rank matched companies by market capitalization. We then form a portfolio of the 30 largest companies by market capitalization for every day over the next year.

For the sake of robustness, we exclude the pre-1930 period from analysis. All of our results are robust to this exclusion and for the sake of brevity, they are not shown. The descriptive statistics for our blue-chip index are in Table 1. The mean daily return is 1.21-basis points and the standard deviation at 98.87-basis points is unsurprisingly high. During bull markets, mean daily returns increase to 9.26 basis points and during bear markets the mean daily return is -10.77-basis points.

Because we want to assess the effect of sentiment on abnormal trading activity, we obtained data from GFD which measures the number of trades that were executed on the constituent companies of the FN30 and FT30. The data begins on the 6 February 1930 and ends on the 30 June 2008 and for the period 7 November 1973 to 28 December 1979, the data is missing. Trading data is not available for the FT30, but the number of trades may better capture how traders react to extreme sentiment.

As can be seen from Table 1, because we have some missing data, there are only 20,615 days with trading data. On average, 40,757 trades were executed per day. The high standard deviation reflects the fact that the number of trades increased substantially from the 1970s onwards, with the maximum reaching 1,428,031 on 22 January 2008.

3. Does sentiment affect the market?

In this section, we test the effect of FT sentiment upon the returns on our blue-chip index. Our media measures are calculated as the number of positive or negative words as a fraction of total words written on that day. Pessimism is then calculated as the difference between the fractions of negative and positive words. To allow comparison with extant literature and to make the economic interpretation of coefficients simpler, we normalize our media measures to have a mean of zero and a unit standard deviation. Following García (2013), we date our media measures to the day written rather than the day published.

To analyze the effect of FT sentiment, we use the following model:

$$R_t = \beta \mathcal{L}_s(M_t) + \gamma \mathcal{L}_s(R_t) + \psi \mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t \qquad (2)$$

where the dependent variable R_t is daily returns on our blue-chip index. $\mathcal{L}_s$ is a lag operator set to five. The variable M_t is one of our media measures. As a set of exogenous variables X_t, we include a constant term, day-of-the-week dummies, month-of-the-year dummies, editor dummies, as well as a dummy for whether date t belongs to a bull or bear market.[8] We exclude public holidays, Sundays and Saturdays after 10 November 1929 and non-trading dates for calculating lags. The

[8] Editors were obtained from ft.com ("FT at 125: The world in focus", Lionel Barber, accessed 8 August 2017). We control for editors because they each may have had a different policy which affected choice of negative and positive words.

sample compromises 30,033 trading days for which we have return data for 27,708 days and volume data for 20,615.[9]

<<<INSERT TABLE 2 HERE>>>

Table 2 shows the relationship between sentiment for the FT and returns. The results in Panel A show that both the positive and negative sentiment of the FT published in the morning have an economically and statistically significant relationship with that day's returns. A one standard deviation rise in negative sentiment predicts a 4.3-basis point decrease in returns, a one standard deviation rise in positive sentiment predicts a 3.4-basis point increase in returns, and a one standard deviation rise in pessimism predicts a 5.9 decrease in returns. This is consistent with previous literature. Tetlock (2007), for example, found that a one standard deviation rise in negative sentiment of the "Abreast of the Market" column in the *Wall Street Journal* predicted a 4.4-basis point fall in the Dow Jones Industrial Average. Similarly, García (2013) found that a one standard deviation rise in the negative sentiment of the "Financial Markets" and "Topics in Wall Street" columns in the *New York Times* predicted a 4.3-basis point decrease in the Dow Jones Industrial Average, with a 3.9-basis point increase for positive sentiment and 5.5-basis point decrease for pessimism.

[9] We treat the following as non-trading dates: Sundays, Saturdays after 9 November 1929 (when no more trades are observed), 25 and 26 December (or the subsequent Monday/Tuesday if observed on Saturday Sunday), 1 January after 1920 (when no more trades are observed), Easter Monday and Good Friday, Whit Monday until 1964, the last Monday in May after 1964, the first Monday in May from 1978 onwards, the first Monday in August until 1965, thereafter the last Monday of the month. The emergency bank holiday following the Sterling crisis on 15 March 1968 is treated as a non-trading day.

Consistent with the extant literature, the effect of sentiment on returns appears to be temporary. Table 2 shows that only the first lag of our negative sentiment and the first two lags of pessimism measure are negative but the subsequent lags are positive. The first three lags of positive sentiment have positive coefficients. However, the fourth and fifth lags have negative coefficients with the former being the only lag other than the first to have a significant statistical relationship with returns. Collectively, with the exception of positive sentiment, the lags of sentiment greater than one basis point are statistically significant, with their coefficients in the table suggesting that they reverse the initial effect of media sentiment. This is consistent with the view that negative or pessimism shocks to sentiment create immediate downward pressure on prices because noise traders sell to arbitrageurs. However, after the negative/pessimism sentiment shock, rational traders drive the stock price back to its fundamental value. Hence the reversal of the initial media effect. In the case of positive sentiment shocks, there is no reversal, which either suggests that positive sentiment shocks contain information or that there are limits to arbitrage associated with positive sentiment shocks that prevent prices being driven back to fundamentals.

Table 2 also reports the relationships between the FT's sentiment and returns with corrections for time-varying volatility and outliers. In Panel B, the returns on our daily blue-chip index are discounted by conditional variance as calculated by a GARCH (1,1) model. Notably, the previously identified relationships between sentiment and returns are robust to this correction and the economic impact of sentiment appears to be greater, with a one standard deviation rise in negative (positive) sentiment predicting a 5.0-basis point decrease (4.0-basis point increase) in returns. Panel C corrects for the effect of outliers using Huber (1973) M-regressions. This robustness check reveals that there are still statistically significant relationships between both negative and positive sentiment and subsequent returns. However, the economic effect falls

slightly using this approach—a 3.6-basis point decrease in returns for a one standard deviation rise in negative sentiment and 2.8-basis point increase for a one standard deviation change in positive sentiment.

The results in Table 2 identify a relationship between sentiment and returns that is similar to that found in studies of the U.S. press (Tetlock, 2007; García, 2013). This is remarkable given that the US and UK markets experienced different performance and informational events and had different regulatory structures. Thus, the effect of media sentiment on markets exists across space and time. This suggests that a similar mechanism is determining the relationship between sentiment and returns.

Whereas the relationship between news media and returns can help us understand if the media influenced the prices at which traders bought or sold stock, the relationship of sentiment with trading activity can help us understand if the FT influenced traders to alter the amount of stocks that they bought or sold. In particular, one would expect that extreme positive or negative sentiment shocks would create large disagreement between rational and noise traders, moving the latter to buy or sell stocks. We follow Gallant, Rossi, and Tauchen (1992) in removing the time trend and known causes of variability in trading. Our results are robust to alternative approaches such as that of Campbell, Grossman, and Wang (1993).

As shown in Equation (3), we model log trading activity as a function of its previous five lags, $\mathcal{L}_s$, as well as a set of dummy variables, X_t, to control for day of the week and month of the year effects. The residuals from Equation (3), ε_t, are then standardized by their annual standard deviation, $\hat{\tau}_t = \frac{\varepsilon_t}{\bar{\sigma}_\varepsilon}$. This results in a detrended and white noise variable of abnormal trading, $\hat{\tau}_t$, which is standardized to have a mean of zero and a unit variance.

$$\tilde{\tau}_t = \beta \mathcal{L}_s(V_t) + \eta X_t + \varepsilon_t \qquad (3)$$

To assess the impact of media on abnormal trading, we model the relationship as per Equation (4). Where $\hat{\tau}_t$ is our measure of abnormal trading, R_t is our daily stock market index returns, R_t^2 is the squared return of our daily stock market index and X_t is a dummy for whether date t is a bull or bear market. $|M_t|$ is the absolute value of our standardized media measures, which is used to proxy extreme sentiment as per Tetlock (2007).

$$\hat{\tau}_t = \beta\mathcal{L}_s(|M_t|) + \gamma\mathcal{L}_s(\hat{\tau}_t) + \zeta\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t \qquad (4)$$

<<INSERT TABLE 3 HERE>>

Table 3 shows that extreme levels of sentiment are positively correlated with trading. A one standard deviation shock to pessimism (positive) sentiment in today's FT causes tomorrow's trading activity to increase by 3.6 (2.8) standard deviations and a one standard deviation shock to negative sentiment in today's FT causes today's trading activity to increase by 3.2 standard deviations. Interestingly, the effect of an extreme negative sentiment shock on trading activity is temporary. In other words, our results are consistent with the theory that extreme sentiment shocks result in a one-off increase in trading by noise traders.

4. Information in sentiment?

Could our measure of media sentiment actually contain information and is it this information which is moving markets? Although we have carefully selected the columns used in the analysis to avoid those which focus on individual company performance, it is still possible that our sentiment measures contain some information. Our results above have shown that FT sentiment has a short-term effect on returns. If this effect is due to a behavioral bias, then we would expect it to be similar on consecutive and non-consecutive (i.e., Mondays and days after public holidays) trading days. However, if there is informational content in sentiment to which traders are responding to, we would expect information to aggregate over non-consecutive trading days and

thus the effect to differ from when the market was open for consecutive days. Since the FT has never been published on a Sunday, Monday editions have two days of information to report. Alternatively, if information production is focused on week days, then we might expect the papers published on Tuesday through Friday to contain more information. Non-consecutive trading days are identified as trading dates having no return data on the previous day while consecutive trading days are those with return data on the previous day. Non-consecutive trading days are Mondays and days after market holiday closures, whereas consecutive trading days are Tuesdays to Fridays in the absence of any holidays. Saturday trading ceases in our sample on 9 November 1929. Prior to this, Saturday is treated as a trading date. After 1929, Saturdays are treated as non-trading dates. We do not incorporate the sentiment of Saturday editions into our measures of sentiment on Monday since the effect on traders may have already dissipated before trading resumes.

<<<INSERT TABLE 4 HERE>>>

Table 4 shows that the immediate effect of FT sentiment on returns does not depend on whether the market was closed on the preceding day. The findings reported in Table 4 are robust to the use of volatility-adjusted returns. From Table 4, we see that a one unit standard deviation rise in negative sentiment is associated with a 4.6-basis point decrease in returns on consecutive trading days and a 5.6-basis point decrease on non-consecutive trading days. A one unit standard deviation rise in positive sentiment is associated with a 3.3 and 2.7-basis point increase in returns on consecutive and non-consecutive trading days. While the immediate effect of FT sentiment on returns does not depend on whether the market was closed on the preceding day, the effect over subsequent days does. On consecutive trading days, the effect of lagged sentiment is the same as the results reported above; collectively they are statistically significant with the size and direction of coefficients suggesting that the immediate impact of FT sentiment is reversed. This is consistent

with FT sentiment creating a behavioral bias in traders. However, on non-consecutive trading days, lags 2 to 5 of each of our negative and pessimism sentiment measures is statistically insignificant, suggesting that the initial effect is not reversed. While collectively, lags 2 to 5 of our positive sentiment measure are statistically significant, all coefficients are positive, again suggesting that the initial effect is not reversed. According to the behavioral hypothesis of sentiment, one would expect reversals after sentiment shocks. Although only suggestive, it may be that there is more informational content in the sentiment of newspaper reporting than has been previously thought in that there is greater informational content following days when the market is closed.

5. Bull and bear markets

The reaction of traders and markets to FT sentiment may differ depending on the prevailing state of the stock market. On the one hand, if the prevailing state of the market is bearish, then investors may pay more attention to and be influenced by the news and its sentiment. As a result, sentiment will more easily move returns and trading activity during bear markets. For example, García (2013) finds that the effect of news sentiment on the stock market is much greater during recessions than during expansions. Notably, there is evidence from the experimental psychology literature that during times of anxiety that people are more open to advice (Gino et al., 2012). If traders are more prone to sentiment during times when the prevailing state of the stock market is negative, then we would expect to see sentiment have a greater effect during bear markets.

On the other hand, investors may be more susceptible to sentiment during times of excitement in the stock market. Indeed, the media may even create narratives to explain stock price movements, thus helping to shape and perpetuate prevailing market states (Shiller, 2017). If media sentiment perpetuates bull runs as per Shiller (2000, p.105), then we would expect to see the short-run effect of positive sentiment not being reversed. If we believe that media sentiment reinforces

the prevailing state of the market, then we would expect positive sentiment to have a greater effect than negative sentiment during bull markets and *vice versa* for bear markets. Alternatively, if the media propagate speculative price movements their effect may be greatest in bull markets (Shiller, 2000 p.105).

One would also anticipate that sentiment shocks might affect trading activity differently depending on the state of the market. Firstly, during a prolonged bull market, which may be marked by more noise traders entering the market, a positive sentiment shock may cause there to be substantially more disagreement between noise and rational traders and hence much more trading than there would be caused by positive sentiment shocks in bear markets. Secondly, if noise traders are more sensitive to negative news during bear markets, then this may result in greater disagreement between noise and rational traders and hence much more trading than there would be under a negative sentiment shock in a bull market.

Given that there is no commonly accepted definition of bull and bear markets, we identify bull and bear phases in the stock market using an algorithm similar to the approach of Pagan and Sossounov (2003), which in turn draws upon the methodology developed by Bry and Boschan (1971) to identify economic cycles for the National Bureau of Economic Research. Bull and bear market phases are naturally bounded by peaks and troughs achieved by the underlying time series of prices. This ex-post methodology first filters the time series to identify candidate turning points. Further filtering is applied to ensure that peak and trough turning points alternate, and satisfy minimum duration constraints including a minimum 16-month cycle length between successive peaks. Individual bull and bear phases are required to be at least 4 months in length with exceptions for shorter periods where the absolute price change exceeds 20 per cent. We have 30 bull and 30 bear phases during our period, with the average length of bull phases being 23.89 months and the

average length of bear phases being 16.21 months. In total, out of 30,033 trading days, 12,136 were during bear markets, and the remainder were in bull markets.

<<<INSERT TABLE 5 HERE>>>

Table 5 shows that FT sentiment has a significant impact on returns in both bull and bear markets. In terms of positive sentiment and pessimism, the effect during bear markets is much larger than that experienced during bull markets. This difference is statistically significant and is consistent with the findings of García (2013). However, in terms of negative sentiment, the impact is not statistically different between the two market states. Thus, there is some evidence to support the view that sentiment plays a greater role during bear markets, which is consistent with the findings of García (2013) for U.S. recessions.

The other notable finding from Table 5 is that the effect of positive sentiment is not reversed during bull markets. This could be because positive sentiment conveys information to the market. Alternatively, it is also consistent with a Shiller (2000) view of the world, which sees the media perpetuating upward movements in the stock market through a series of positive sentiment shocks, which are not reversed in the short run. The reason that rational investors do not correct the 'mistakes' of noise traders is due to limits of arbitrage (Shleifer and Vishny, 1997). Limits to arbitrage may be greater during bull markets, making it more difficult to correct mispricing arising from investor sentiment.

A final finding from Table 5 is that the effect of our sentiment measures does not reverse in bear markets. Again, this could be because sentiment conveys information to the market or that the media create narratives around market performance, thereby propagating broader market movements.

Table 6 shows how the relationship between extreme FT sentiment and abnormal trading activity varies with the prevailing state of the market. The first thing to note is that the effect of extreme sentiment on trading activity is statistically the same across bull and bear markets. This finding works against the view that investors are more sensitive to sentiment during bear markets. Although the differences are not statistically significant, the effect appears to be greater in bull markets. A one standard deviation increase in negative sentiment causes 0.033 (0.028) standard deviations increase in trading activity during bull (bear) markets. The combined effect of the first two lags of extreme positive sentiment is associated with an increase of 0.032 standard deviations during bear markets, whereas a one standard deviation increase in positive sentiment in today's and yesterday's newspaper during a bull market increases today's trading activity by 0.035 standard deviations. This finding is consistent with the view that noise traders drive trading activity during speculative booms and that the news media plays a role in propagating such booms. It is also consistent with Baker and Stein (2004) who suggest that irrational investors are more likely to trade when they are optimistic and speculate on stocks rising than the reverse.

<<<INSERT TABLE 6 HERE>>>

As an alternative way of classifying bull and bear markets, we use the occurrence of the words 'bull' and 'bullish' in the FT to define whether a particular month is in a bull or bear phase. We aggregated occurrences of these words to a monthly level and expressed them as a fraction of all words published in that month. Months are then classified as bull when the occurrence of bull words is greater than the five-year moving-average of bull words in the FT. This results in 567 months being classified as bull months with the remainder 775 as bear months. Table 7 shows that using this alternative classification produces results which are consistent with our previous findings in Table 5. While there is no statistically significant difference between bull and bear

months for the effect of sentiment on our blue-chip index, the effect appears to be greater during bull markets. Collectively, lags greater than one are insignificant, suggesting the initial effect of the media is not reversed and the media may help propagate the existing market state.

<<<INSERT TABLE 7 >>

As a robustness exercise, we use GDP data to classify expansions and contractions in the economy. We then use these classifications to see whether the effect of media sentiment differs across them. The Office for National Statistics has quarterly GDP data for the U.K. from 1955 onwards. We use the standard definition of a recession (i.e., two consecutive quarters of negative GDP growth) in our analysis. Because this analysis is limited to the post-1955 period, we also use annual GDP data back to 1899, which was obtained from the Bank of England. We classify each year as an expansion or contraction, depending on whether its GDP growth was positive or negative. We then compare the effect of FT sentiment across these various states of the economy—the results of this analysis are in Tables 8 and 9.

<<<INSERT TABLES 8 and 9 HERE>>

The results in Tables 8 and 9 broadly concur with those that we found for bull and bear markets. First, the state of the overall economy appears not to have a bearing on the effect of sentiment on the stock market. Second, as with bull markets, positive sentiment shocks do not reverse during periods of economic growth.

6. Commentary versus reporting

Our FT content analysis uses four principal sections from the newspaper. The four sections are the Markets section, the Lex column, News in Brief section, and the Editorials and Leaders section. The Markets section of the paper contains reporting and commentary upon what was happening in

the financial markets, whereas the Lex column provides pure commentary on macroeconomic and financial markets. The News in Brief section simply contains a summary of the major news items of the day, which includes, but is not limited to financial and economic news. Finally, the Editorials and Leaders section are opinion pieces commenting on a range of financial and non-financial issues.

In this section, we analyze the sentiment of each of the four sections and its effect on the market. We do so for two reasons. First, commentary, rather than reporting, is unlikely to contain new information on companies or the economy—they simply reflect the views of the FT's journalists and editors on the issues of the day. As a consequence, the sentiment of the two commentary sections of the newspaper provide a relatively uncontaminated measure to assess how the stock market responds to sentiment. Second, one might expect commentary sections of the FT to be more sentiment laden, containing the sort of speculative language that might reflect the excitement and dismay of investors.

Table 10 displays the key characteristics of the four sections. We use the *Harvard Psychosocial Dictionary* to quantify the use of economic and political words and Loughran McDonald (2011) word lists to measure words that express uncertainty in the four sections of the FT. Average word count and average uncertainty demonstrate that the Editorials and Leaders section and Lex columns are more commentary in nature; they are both longer than News in Brief and Market articles and express much more uncertainty, which is typical of commentary articles which are scanning the future or speculating about the present. Thus, it appears that the commentary articles are more subjective in their assessment of events than the articles from the Markets and News in Brief sections. The average political content of the four sections demonstrate that the scope of our non-financial columns (Editorials and Leaders and News in Brief) is broader

than the two narrowly-focused financial sections. Table 10 also shows the differences between the use of positive and negative words across sections. Editorials and Leaders contain the highest percentage of both negative and positive words, while News in Brief contains the lowest percentage of both measures. However, the differences are not large.

<<<INSERT TABLE 10 HERE>>

Table 11 shows the effect of the sentiment of each of the four sections on returns. The first thing of note is that the effect of the overall newspaper (i.e., the aggregation of the four sections) is greater than any of the individual sections. Thus, it is not just the narrow Markets section which is affecting investor sentiment. This finding extends the extant literature which has been focused on financial-market columns.

The next thing to note is that the effect of pessimism, which aggregates positive and negative content, is greatest for the Markets section. Interestingly, the results in Table 11 suggest that the initial returns are all reversed over following days. However, the economic effect of the Markets section is more than one half that of the FT.

The negative sentiment of the Lex column has the next greatest economic impact on returns. Similar to previous results, this effect appears to reverse over subsequent days. The positive sentiment of the Lex column does not affect the stock market. We put this down to the fact that the Lex column by its nature is critical in its stance and is not necessarily given to positive commentary. The fact that the Lex column affects returns suggests that our sentiment measure is not a proxy for new information being communicated to the market.

The sentiment of the FT's Editorials and Leaders does not have a meaningful economic effect on returns. The final thing to note is that the positive sentiment and pessimism of the News in Brief section affects returns. Notably, however, the initial returns over subsequent days are not

reversed. This implies that News in Brief is the section of the FT which is most likely to play a role in propagating upward movements of the stock market.

<<<INSERT TABLE 11 HERE>>>

Table 12 shows the effect of sentiment on abnormal trading volume. While there was limited evidence of Editorials and Leaders affecting market returns, increased positivity of that section is associated with increased abnormal trading. Similarly, increased negativity of Lex and Markets sections are associated with increased abnormal trading. Interestingly, the relationship between pessimistic sentiment and abnormal trading for News in Brief suggest that increased pessimism is associated with fewer trades occurring on the market, althought the strength of the statistical relationship is not strong.

<<<INSERT TABLE 12 HERE>>>

Overall, the findings from the sections suggest that the more narrowly focused the section is on financial markets, the more likely it is that sentiment has an effect on the stock market. In addition, the fact that the sentiment in commentary sections of the FT affects returns suggests that sentiment does not necessarily contain information. However, the sentiment-laden parts of the FT do not affect investors more than the other sections of the newspaper which focus on reporting news. This implies that the words used to report events and the commentary integrated into such reporting is just as sentiment laden as pure commentary sections of the FT.

7. Conclusions

This paper constructs a measure of investor sentiment using reporting from a broad range of articles in the FT over a 111-year period. We have three main findings. First, the positivity, negativity and pessimism of FT reporting predicts daily stock returns as well as stock trading. The weight of our evidence, as well as the use of sentiment-laden but information-light sections from

the FT in our analysis, suggests that our sentiment measure was uncontaminated by information and that is was sentiment and not information moving markets.

Second, the effect of negative sentiment on stock returns and stock trading does not differ substantially across bull and bear markets, suggesting that negative sentiment does not play a greater role during bear markets. However, we do find that the effect of our sentiment measures is not reversed in bear markets, which perhaps indicates that the news media propagate broader market movements, and that there are limits to arbitrage, particularly during bull markets, which prevent rational traders from driving stock prices back to fundamentals.

Our dataset opens up other research questions. While our paper reveals that the effect of sentiment does not always reverse, it would be insightful to understand how limits to arbitrage vary over the long run and across bull and bear markets and the extent to which limits to arbitrage explain the non-reversal. In addition, it would be helpful to understand how the narratives created and shaped by the *Financial Times* affected the stock market.

References

Bagehot, W. (1873) *Lombard Street: A Description of the Money Market*, Scribner, Armstrong & Company, New York.

Baker, M. and Stein, J. (2004) Market Liquidity as a Sentiment Indicator, *Journal of Financial Markets* **7**, 271–99.

Baker, M. and Wurgler, J. (2006) Investor Sentiment and the Cross Section of Stock Returns, *Journal of Finance* **61**, 1645–80.

Baker, M. and Wurgler, J. (2007) Investor Sentiment in the Stock Market, *Journal of Economic Perspectives* **21**, 129–151.

Barberis, N. Shleifer, A. and Vishny, R. W. (1998) A Model of Investor Sentiment, *Journal of Financial Economics* **49**, 307–43.

Bhattacharya, U. I., Galpin, N., Ray, R. and Yu, X. (2009) The Role of the Media in the Internet Ipo Bubble, *Journal of Financial and Quantitative Analysis* **44**, 657–682.

Bignon, V. and Miscio, A. (2010) Media Bias in Financial Newspapers: Evidence from Early Twentieth-century France, *European Review of Economic History* **14**, 383–432.

Brown, G. W. and Cliff, M. T. (2005) Investor Sentiment and Asset Valuation, *Journal of Business* **78**, 405–40.

Bry, G. and Boschan, C. (1971) Cyclical Analysis of Time Series: Selected Procedures and Computer Programs, *National Bureau of Economic Research*.

Campbell, G., Turner, J. D. and Walker, C. B. (2012) The Role of the Media in a Bubble: Evidence from the British Railway Mania, *Explorations in Economic History* **49**, 461–481.

Campbell, J. Y., Grossman, S. J. and Wang J. (1993) Trading Volume and Serial Correlation in Stock Returns, *Quarterly Journal of Economics* **108**, 905–939.

Chauvet M. and Potter S. (2000) Coincident and Leading Indicators of the Stock Market, *Journal of Empirical Finance* **7**, 87–111.

Cutler, D. M., Poterba, J. M. and Summers, L. H. (1989) What Moves Stock Prices?, *Journal of Portfolio Management* **15**, 4–12.

De Long, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J. (1990) Noise Trader Risk in Financial Markets, *Journal of Political Economy* **98**, 703–38.

Dougal, C., Engelberg, J., García, D. and Parsons, C.A. (2012) Journalists and the Stock Market, *Review of Financial Studies* ***25***, 639–679.

Dyck, A., Volchkova, N. and Zingales, L. (2008) The Corporate Governance Role of the Media: Evidence from Russia, *Journal of Finance* **63**, 1093–1135.

Dyck, A. and Zingales, L. (2003) The Media and Asset Prices, unpublished working paper, Harvard Business School, Cambridge, Massachusetts.

Edmans, A. García, D. and Oyvind, N. (2007) Sports Sentiment and Stock Returns, *Journal of Finance* **62**, 1967–98.

Elmendorf, D. W., Hirschfeld, M. L., and Weil, D. N. (1996) The Effect of News on Bond Prices: Evidence from the United Kingdom, 1900-1920, *Review of Economics and Statistics* **78**, 341–44.

Engelberg, J. E. and Parsons, C. A. (2011) The Causal Impact of Media in Financial Markets, Journal *of Finance* **66**, 67–97.

Fang, L. and Peress, J. (2009) Media Coverage and the Cross-Section of Stock Returns, *Journal of Finance* **64**, 2023–2052.

Ferguson, N. J., Philip, D., Lam, H. Y. T. and Guo, J. M. (2015) Media Content and Stock Returns: The Predictive Power of the Press, *Multinational Finance Journal* **19**, 1–31.

García, D. (2013) Sentiment during Recessions, *Journal of Finance* **68**, 1267–1300.

Gallant, A. R., Rossi, P. E. and Tauchen, G. (1992) Stock Prices and Volume, *Review of Financial Studies* ***5***, 199–242.

Gino, F., Brooks, A. W. and Schweitzer, M. E. (2012) Anxiety, Advice and the Ability to Discern: Feeling Anxious Motivates Individuals to Seek and Use Advice, *Journal of Personality and Social Psychology* **102**, 497–512.

Griffin, J. M., Hirschey, N. H. and Kelly, P. J. (2011) How Important Is the Financial Media in Global Markets?, *Review of Financial Studies* **24**, 3941–3992.

Huber, P. J. (1973) Robust Regression: Asymptotics, Conjectures and Monte Carlo, *The Annals of Statistics* **1**, 799–821.

Huberman, G. and Regev, T. (2001) Contagious Speculation and a Cure for Cancer: A Nonevent That Made Stock Prices Soar, *Journal of Finance* **56**, 387–396.

Kamstra, M. J., Kramer, L. A. and Levi, M. D. (2003) Winter Blues: A SAD Stock Market Cycle, *American Economic Review* **93**, 1257–63.

Kräussl, R. and Mirgorodskaya, E. (2017) Media, sentiment and market performance in the long run, *The European Journal of Finance* **23**, 1059**–**1082.

Kynaston, D., (1988). *The Financial Times: A Centenary History*, Viking Adult, London.

Loughran, T. and McDonald, B. (2011) When is a Liability Not a Liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* **66**, 35–65.

Loughran, T. and McDonald, B. (2015) The Use of Word Lists in Textual Analysis, *Journal of Behavioral Finance* ***16***, 1–11.

Niederhoffer, V. (1971) The Analysis of World Events and Stock Prices, *Journal of Business* **44**, 193–219.

Pagan, A. R. and Sossounov, K. A. (2003) A Simple Framework for Analysing Bull and Bear Markets, *Journal of Applied Econometrics* **18**, 23–46.

Porter, D. (1998) City Editors and the Modern Investing Public: Establishing the Integrity of the New Financial Journalism in Late Nineteenth-Century London, *Media History* **4**, 49–60.

Shleifer, A. and Summers, L. H. (1990) The Noise Trader Approach to Finance, *Journal of Economic Perspectives* **4**, 19–33.

Shleifer, A. and Vishny, R. W. (1997) The Limits to Arbitrage, *Journal of Finance* **52**, 35–55.

Shiller, R. J. (2000) *Irrational Exuberance*, Princeton University Press, Princeton, NJ.

Shiller, R. J. (2017) Narrative Economics, *Cowles Foundation Discussion Paper* No. 2069.

Soo, C. (2013) Quantifying Animal Spirits: News Media and Sentiment in the Housing Market. Unpublished working paper, Ross School of Business, Michigan.

Taylor, J. (2012) Watchdogs or Apologists? Financial Journalism and Company Fraud in Early Victorian Britain, *Historical Research* **85**, 632–52.

Tetlock, P. C. (2007) Giving Content to Investor Sentiment: The Role of Media in the Stock Market, *Journal of Finance* **62**, 1139–1168.

Turner, J. D., Ye, Q. and Walker, C. B. (2017) Media Coverage and Stock Returns on the London Stock Exchange, 1825–70, *Forthcoming Review of Finance*.

Walker, C. B. (2014) Housing Booms and Media Coverage, *Applied Economics* ***46***, 3954–3967.

Zipf, G.K. (1949) *Human Behaviour and the Principle of Least Effort*, Addison-Wesley Press, Cambridge, MA.

TABLE 1: Descriptive Statistics

Returns are log close to close from our blue-chip indices. From 1899 to 1930 we use a proprietary index generated from Global Financial Data price and Investor Monthly Manual market capitalization data to identify the 30 largest companies for which we have data. Trade data is taken from Global Financial Data and is available from February 1930 to June 2008, excluding November 6, 1973 to December 28, 1979. Word count, columns, negative, positive and pessimism are calculated from the Financial Times Historic Archive 1899 to 2010. It includes articles published in the News in Brief, Markets, Editorial and Leaders and Lex sections. Negative and Positive are based on 2,337 and 353 words from Loughran and McDonald (2011). Our pessimism measure is the daily difference between the percentage of negative and positive words.

		Obs.	Mean	Median	St. Dev.	25%	75%
				Panel A: All Periods			
Returns	Basis Points	27,708	1.21	0.10	98.87	-35.76	40.25
Trades	No. ('000s)	20,615	40.76	11.08	98.90	6.30	25.46
Word Count	No.	34,044	5,549	5,155	2,692	3,668	7,042
Columns	No.	34,044	9.31	8.00	7.58	5.00	10.00
Negative	%	34,044	1.64	1.59	0.51	1.27	1.98
Positive	%	34,044	0.95	0.93	0.27	0.77	1.11
Pessimism	%	34,044	0.69	0.66	0.61	0.25	1.11
				Panel B: Bull Periods			
Returns	Basis Points	16,531	9.26	7.38	83.07	-25.96	42.87
Trades	No. ('000s)	12,418	36.31	12.98	71.86	7.24	28.80
Word Count	No.	17,682	5,442	4,967	2,751	3,566	6,826
Columns	No.	17,682	8.36	7.00	5.50	5.00	10.00
Negative	%	17,682	1.60	1.55	0.50	1.23	1.94
Positive	%	17,682	0.97	0.95	0.27	0.79	1.13
Pessimism	%	17,682	0.62	0.59	0.61	0.19	1.06
				Panel C: Bear Periods			
Returns	Basis Points	11,163	-10.77	-5.65	117.50	-53.07	36.03
Trades	No. ('000s)	8,001	48.14	9.02	130.29	5.29	16.85
Word Count	No.	12,024	5,630	5,330	2,714	3,677	7,274
Columns	No.	12,024	10.64	8.00	9.93	5.00	11.00
Negative	%	12,024	1.66	1.60	0.51	1.30	1.98
Positive	%	12,024	0.91	0.89	0.26	0.74	1.06
Pessimism	%	12,024	0.74	0.70	0.61	0.33	1.14

TABLE 2: Feedback from FT to the blue-chip index, 1899–2010

This table reports β coefficients and t-stats from the model $R_t = \beta\mathcal{L}_s(M_t) + \gamma\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index log returns. M_t is one of our media measures. X_t is our exogenous variables including day-of-the-week dummies, month-of-the-year dummies, editor dummies and a dummy for whether date t is a bull or bear market. All results are based on 23,017 observations. Panel A and B report results with robust standard errors. Panel B uses returns on our daily stock index discounted by conditional variance as calculated by a GARCH (1,1) model. Panel C controls for outliers through Huber (1973) M-regression.

	(1)		(2)		(3)	
	Negative	*t-stat*	Positive	*t-stat*	Pessimism	*t-stat*
	Panel A: OLS Regression Coefficients					
$\mathcal{L}_1$	-0.043	-4.75	0.034	5.50	-0.059	-6.62
$\mathcal{L}_2$	0.002	0.17	0.009	1.59	-0.005	-0.54
$\mathcal{L}_3$	0.010	1.14	0.002	0.31	0.007	0.82
$\mathcal{L}_4$	0.012	1.33	-0.014	-2.35	0.020	2.25
$\mathcal{L}_5$	0.029	3.22	-0.002	-0.35	0.026	2.87
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	22.51	0.000	30.25	0.000	43.79	0.000
$\mathcal{L}_{2-5} = 0$	4.20	0.002	1.94	0.101	4.17	0.002
	Panel B: Volatility-Adjusted Returns Coefficients					
$\mathcal{L}_1$	-0.050	-5.47	0.040	5.36	-0.069	-7.30
$\mathcal{L}_2$	0.002	0.23	0.005	0.67	-0.001	-0.10
$\mathcal{L}_3$	0.018	1.98	-0.004	-0.60	0.018	1.96
$\mathcal{L}_4$	0.007	0.73	-0.008	-1.10	0.012	1.25
$\mathcal{L}_5$	0.015	1.62	0.003	0.38	0.011	1.16
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	29.90	0.000	28.74	0.000	53.28	0.000
$\mathcal{L}_{2-5} = 0$	2.31	0.055	0.54	0.708	2.24	0.062
	Panel C: Huber Regression Coefficients					
$\mathcal{L}_1$	-0.036	-4.85	0.028	4.62	-0.048	-6.48
$\mathcal{L}_2$	-0.008	-1.06	0.007	1.11	-0.010	-1.36
$\mathcal{L}_3$	0.015	1.96	-0.002	-0.40	0.014	1.89
$\mathcal{L}_4$	0.005	0.70	-0.005	-0.81	0.008	1.09
$\mathcal{L}_5$	0.020	2.65	-0.003	-0.42	0.018	2.50
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	23.52	0.000	21.31	0.000	42.00	0.000
$\mathcal{L}_{2-5} = 0$	3.64	0.006	0.53	0.713	3.74	0.005

TABLE 3: Feedback from Extreme News Content to Abnormal Trading

This table reports β coefficients and t-stats from the model $\hat{\tau}_t = \beta\mathcal{L}_s(|M_t|) + \gamma\mathcal{L}_s(\hat{\tau}_t) + \zeta\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + +\eta X_t + \epsilon_t$. R_t is our daily stock market index returns. $|M_t|$ is one the absolute value of our standardized media measures. X_t is our exogenous variables, including editor dummies and a dummy for whether date t is a bull or bear market. $\hat{\tau}_t$ is abnormal trading which is calculated from the equation $\tilde{\tau}_t = \beta\mathcal{L}_s(\tau_t) + \eta X_t + \varepsilon_t$. Where τ is log number of trades (000's). The residuals are standardized by the annual standard deviation in the residuals. Thus $\hat{\tau}_t = \frac{\varepsilon_t}{\bar{\sigma}_\varepsilon}$ which results in a detrended and white noise variable of abnormal trading. Data begins on February 6, 1930 and ends June 30, 2008. Trading data is not available from November 6, 1973 to December 28, 1979. Results based on 16,973 obs.

	(1) Negative	(2) *t-stat*	(3) Positive	(4) *t-stat*	(5) Pessimism	(6) *t-stat*
$\mathscr{L}Abs_1$	0.032	3.76	0.007	0.93	0.006	0.77
$\mathscr{L}Abs_2$	0.006	0.75	0.028	3.62	0.036	4.41
$\mathscr{L}Abs_3$	-0.016	-1.92	0.009	1.13	-0.009	-1.11
$\mathscr{L}Abs_4$	-0.009	-1.06	-0.016	-2.01	-0.010	-1.30
$\mathscr{L}Abs_5$	0.015	1.89	-0.005	-0.59	0.005	0.63
	t-stat.	*p-val.*	*t-stat.*	*p-val.*	*t-stat.*	*p-val.*
$\mathscr{L}_1 = 0$	14.16	0.000	0.86	0.354	0.60	0.440
$\mathscr{L}_{2-5} = 0$	2.15	0.072	4.50	0.001	5.46	0.000

TABLE 4: Feedback from FT to the blue-chip index on consecutive and non-consecutive trading days

This table reports β coefficients and t-stats from the model $R_t = \beta\mathcal{L}_s(M_t) + \gamma\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index log returns. M_t is one of our media measures. X_t is our exogenous variables including day-of-the-week dummies, month-of-the-year dummies, editor dummies and a dummy for whether date t is a bull or bear market. All results are reported with robust standard errors. Non-consecutive trading days are identified as having no return data on the previous day. Consecutive trading days are those with return data on the previous day.

	(1)		(2)		(3)	
	Negative	*t-stat*	Positive	*t-stat*	Pessimism	*t-stat*
	Panel A: Consecutive Trading Days (18,461 obs.)					
$\mathcal{L}_1$	-0.046	-4.50	0.033	4.83	-0.061	-6.01
$\mathcal{L}_2$	-0.003	-0.32	0.005	0.75	-0.006	-0.65
$\mathcal{L}_3$	0.021	2.07	0.000	0.06	0.016	1.66
$\mathcal{L}_4$	0.011	1.03	-0.020	-2.89	0.022	2.21
$\mathcal{L}_5$	0.030	2.94	-0.006	-0.91	0.028	2.91
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	20.28	0.000	23.32	0.000	36.09	0.000
$\mathcal{L}_{2-5} = 0$	4.41	0.002	2.52	0.039	4.87	0.001
	Panel B: Non-Consecutive Trading Days (4,556 obs.)					
$\mathcal{L}_1$	-0.056	-2.78	0.027	1.87	-0.063	-3.17
$\mathcal{L}_2$	0.024	1.07	0.028	2.01	0.001	0.07
$\mathcal{L}_3$	-0.019	-0.90	0.017	1.18	-0.026	-1.29
$\mathcal{L}_4$	0.013	0.67	0.005	0.34	0.008	0.44
$\mathcal{L}_5$	0.027	1.29	0.013	0.96	0.015	0.69
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	7.74	0.005	3.51	0.061	10.02	0.002
$\mathcal{L}_{2-5} = 0$	1.31	0.264	2.50	0.041	0.57	0.688
	Panel C: Differences Between Consecutive and Non-Consecutive Trading Days					
	χ^2	*p*-val	χ^2	*p*-val	χ^2	*p*-val
$\mathcal{L}_{1Con} = \mathcal{L}_{1Non-Con}$	0.37	0.540	0.00	0.973	0.26	0.608
$\mathcal{L}_{2-5Con} = \mathcal{L}_{2-5Non-Con}$	0.02	0.894	11.59	0.001	2.49	0.115

TABLE 5: Feedback from FT to the blue-chip index by bull/bear market

This table reports β coefficients and t-stats from the model $R_t = \beta\mathcal{L}_s(M_t) + \gamma\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index returns. M_t is one of our media measures. X_t is our exogenous variables including day-of-the-week dummies, month-of-the-year dummies, editor dummies and a dummy for when the market state switches between a bull and bear state. Differences are tested by running the above model.

	(1)		(2)		(3)	
	Negative	*t-stat*	Positive	*t-stat*	Pessimism	*t-stat*
	Panel A: Bull Market (13,695 obs.)					
$\mathcal{L}_1$	-0.045	-3.73	0.022	2.74	-0.052	-4.49
$\mathcal{L}_2$	0.008	0.68	0.010	1.31	-0.000	-0.03
$\mathcal{L}_3$	0.005	0.44	0.009	1.15	-0.002	-0.19
$\mathcal{L}_4$	0.022	1.89	-0.013	-1.70	0.028	2.45
$\mathcal{L}_5$	0.031	2.62	-0.002	-0.20	0.026	2.30
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	13.93	0.000	7.51	0.006	20.18	0.000
$\mathcal{L}_{2-5} = 0$	3.56	0.007	1.40	0.232	3.23	0.012
	Panel B: Bear Market (9,322 obs.)					
$\mathcal{L}_1$	-0.049	-3.34	0.057	5.82	-0.079	-5.60
$\mathcal{L}_2$	-0.012	-0.81	0.014	1.51	-0.019	-1.33
$\mathcal{L}_3$	0.012	0.84	-0.002	-0.16	0.012	0.83
$\mathcal{L}_4$	-0.005	-0.36	-0.009	-0.90	0.002	0.12
$\mathcal{L}_5$	0.023	1.57	0.001	0.07	0.019	1.32
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	11.18	0.001	33.82	0.000	31.39	0.000
$\mathcal{L}_{2-5} = 0$	0.92	0.449	0.70	0.591	0.97	0.425
	Panel C: Testing Effect Differences					
	χ^2	*p*-val	χ^2	*p*-val	χ^2	*p*-val
$\mathcal{L}_{1Bull} = \mathcal{L}_{1Bear}$	0.67	0.413	13.85	0.000	5.86	0.016

TABLE 6: Effect of FT abs sentiment on abnormal trading by bull/bear

This table reports β coefficients and t-stats from the model $\hat{\tau}_t = \beta\mathcal{L}_s(|M_t|) + \gamma\mathcal{L}_s(\hat{\tau}_t) + \zeta\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index returns. $|M_t|$ is one the absolute value of our standardized media measures. X_t is our exogenous variables including editor dummies and a dummy for when the market state switches between a bull and bear state. $\hat{\tau}_t$ is abnormal trading which is calculated from the equation $\tilde{\tau}_t = \beta\mathcal{L}_s(\tau_t) + \eta X_t + \varepsilon_t$. Where τ is log number of trades (000's). The residuals are standardized by the annual standard deviation in the residuals. Thus $\hat{\tau}_t = \frac{\varepsilon_t}{\bar{\sigma}_\varepsilon}$ which results in a detrended and white noise variable of abnormal trading. $\hat{\tau}_t$ is standardized to have a mean of zero and unit variance. Data begins on February 6, 1930 and ends June 30, 2008. Trading data is not available from November 6, 1973 to December 28, 1979. Results based on 16,725 obs.

	(1)		(2)		(3)	
	Negative	*t-stat*	Positive	*t-stat*	Pessimism	*t-stat*
			Panel A: Bull (10,750 obs.)			
$\mathcal{L}Abs_1$	0.033	3.20	0.003	0.27	0.007	0.75
$\mathcal{L}Abs_2$	0.008	0.80	0.032	3.31	0.038	3.77
$\mathcal{L}Abs_3$	-0.015	-1.49	0.008	0.76	-0.009	-0.86
$\mathcal{L}Abs_4$	-0.006	-0.61	-0.016	-1.55	-0.013	-1.31
$\mathcal{L}Abs_5$	0.023	2.45	-0.006	-0.64	0.012	1.25
			Tests			
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}Abs_1 = 0$	10.24	0.001	0.07	0.791	0.57	0.451
$\mathcal{L}Abs_{2-5} = 0$	2.21	0.065	3.44	0.008	4.47	0.001
			Panel B: Bear (6,223 obs.)			
$\mathcal{L}Abs_1$	0.028	1.80	0.014	1.04	0.001	0.05
$\mathcal{L}Abs_2$	0.002	0.12	0.018	1.42	0.029	2.04
$\mathcal{L}Abs_3$	-0.021	-1.33	0.008	0.67	-0.012	-0.83
$\mathcal{L}Abs_4$	-0.018	-1.29	-0.017	-1.33	-0.009	-0.63
$\mathcal{L}Abs_5$	-0.006	-0.41	-0.003	-0.26	-0.012	-0.83
			Tests			
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}Abs_1 = 0$	3.23	0.072	1.08	0.299	0.00	0.957
$\mathcal{L}Abs_{2-5} = 0$	1.01	0.402	1.02	0.395	1.46	0.211
			Panel C: Testing Effect Differences			
	χ^2	*p*-val	χ^2	*p*-val	χ^2	*p*-val
$\mathcal{L}_{1Bull} = \mathcal{L}_{1Bear}$	0.12	0.725	0.85	0.357	0.05	0.829
$\mathcal{L}_{2-5Bull} = \mathcal{L}_{2-5Bear}$	2.52	0.113	0.21	0.649	1.16	0.282

TABLE 7: Feedback from FT to the blue-chip index by alternative bull/bear definitions, 1899–2010

This table reports β coefficients and t-stats from the model $R_t = \beta\mathcal{L}_s(M_t) + \gamma\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index returns. M_t is one of our media measures. X_t is our exogenous variables including day-of-the-week dummies, month-of-the-year dummies and editor dummies. Differences are tested by running the above model. We identify bull states as any month for which the FT used more bullish than bearish and vice versa for bear states.

	(1)		(2)		(3)	
	Negative	*t-stat*	Positive	*t-stat*	Pessimism	*t-stat*
	Panel A: FT Bull Months (9,414 obs.)					
$\mathcal{L}_1$	-0.055	-3.90	0.045	4.62	-0.075	-5.45
$\mathcal{L}_2$	-0.002	-0.18	0.021	2.12	-0.015	-1.11
$\mathcal{L}_3$	0.008	0.57	0.009	0.99	0.002	0.14
$\mathcal{L}_4$	0.011	0.81	-0.009	-0.92	0.018	1.33
$\mathcal{L}_5$	0.022	1.65	0.006	0.61	0.017	1.27
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	15.24	0.000	21.39	0.000	29.65	0.000
$\mathcal{L}_{2-5} = 0$	1.11	0.349	1.69	0.148	1.21	0.303
	Panel B: FT Bear Months (13,603 obs.)					
$\mathcal{L}_1$	-0.046	-3.77	0.035	4.32	-0.061	-5.21
$\mathcal{L}_2$	-0.005	-0.38	0.010	1.33	-0.010	-0.84
$\mathcal{L}_3$	0.001	0.09	0.005	0.66	-0.002	-0.21
$\mathcal{L}_4$	0.004	0.30	-0.010	-1.29	0.010	0.84
$\mathcal{L}_5$	0.023	1.86	0.002	0.24	0.018	1.46
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	14.20	0.000	18.62	0.000	27.17	0.000
$\mathcal{L}_{2-5} = 0$	0.95	0.435	0.91	0.460	0.85	0.493
	Panel C: Testing Effect Differences					
	χ^2	*p*-val	χ^2	*p*-val	χ^2	*p*-val
$\mathcal{L}_{1Exp} = \mathcal{L}_{1Con}$	0.36	0.547	0.86	0.354	0.93	0.336

TABLE 8: Feedback from FT to the blue-chip index over the business cycle, 1955–2010

This table reports β coefficients and t-stats from the model $R_t = \beta\mathcal{L}_s(M_t) + \gamma\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index returns. M_t is one of our media measures. X_t is our exogenous variables including day-of-the-week dummies, month-of-the-year dummies and editor dummies. Differences are tested by running the above model.

	(1)		(2)		(3)	
	Negative	*t-stat*	Positive	*t-stat*	Pessimism	*t-stat*
	Panel A: Expansion (11,675 obs.)					
$\mathcal{L}_1$	-0.024	-2.02	0.019	2.10	-0.032	-2.73
$\mathcal{L}_2$	-0.004	-0.34	0.013	1.44	-0.011	-0.96
$\mathcal{L}_3$	0.001	0.11	0.007	0.79	-0.003	-0.29
$\mathcal{L}_4$	0.018	1.54	-0.006	-0.67	0.019	1.62
$\mathcal{L}_5$	0.031	2.65	-0.002	-0.22	0.027	2.38
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	4.09	0.043	4.39	0.036	7.44	0.006
$\mathcal{L}_{2-5} = 0$	2.58	0.035	0.78	0.538	2.35	0.052
	Panel B: Recession (1,732 obs.)					
$\mathcal{L}_1$	-0.058	-1.76	0.036	1.57	-0.070	-2.24
$\mathcal{L}_2$	0.047	1.59	-0.000	-0.01	0.042	1.46
$\mathcal{L}_3$	0.013	0.45	0.036	1.63	-0.009	-0.29
$\mathcal{L}_4$	-0.021	-0.65	-0.009	-0.44	-0.009	-0.29
$\mathcal{L}_5$	0.030	0.98	0.025	1.11	0.013	0.42
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	3.10	0.078	2.46	0.117	5.03	0.025
$\mathcal{L}_{2-5} = 0$	1.08	0.363	1.08	0.365	0.62	0.648
	Panel C: Testing Effect Differences					
	χ^2	*p*-val	χ^2	*p*-val	χ^2	*p*-val
$\mathcal{L}_{1Exp} = \mathcal{L}_{1Rec}$	1.32	0.251	0.98	0.322	1.96	0.162

TABLE 9: Feedback from FT to the blue-chip index by GDP expansion/contraction market, 1899–2010

This table reports β coefficients and t-stats from the model $R_t = \beta\mathcal{L}_s(M_t) + \gamma\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index returns. M_t is one of our media measures. X_t is our exogenous variables including day-of-the-week dummies, month-of-the-year dummies and editor dummies. Differences are tested by running the above model.

	(1) Negative	*t-stat*	(2) Positive	*t-stat*	(3) Pessimism	*t-stat*
	Panel A: GDP Expansion (19,333 obs.)					
$\mathcal{L}_1$	-0.053	-5.32	0.040	5.89	-0.070	-7.40
$\mathcal{L}_2$	-0.004	-0.38	0.015	2.30	-0.013	-1.29
$\mathcal{L}_3$	0.008	0.75	0.003	0.48	0.004	0.46
$\mathcal{L}_4$	0.016	1.58	-0.009	-1.30	0.019	2.05
$\mathcal{L}_5$	0.022	2.17	0.000	0.01	0.019	1.90
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	28.29	0.000	34.66	0.000	54.80	0.000
$\mathcal{L}_{2-5} = 0$	2.34	0.053	1.75	0.136	2.54	0.038
	Panel B: GDP Contraction (3,694 obs.)					
$\mathcal{L}_1$	-0.034	-1.40	0.036	2.42	-0.054	-2.22
$\mathcal{L}_2$	0.005	0.24	0.008	0.54	-0.001	-0.05
$\mathcal{L}_3$	0.003	0.13	0.019	1.35	-0.011	-0.50
$\mathcal{L}_4$	-0.022	-0.91	-0.011	-0.72	-0.010	-0.40
$\mathcal{L}_5$	0.031	1.38	0.013	0.87	0.016	0.72
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	1.95	0.163	5.84	0.016	4.92	0.027
$\mathcal{L}_{2-5} = 0$	0.64	0.638	0.94	0.437	0.21	0.932
	Panel C: Testing Effect Differences					
	χ^2	*p*-val	χ^2	*p*-val	χ^2	*p*-val
$\mathcal{L}_{1Exp} = \mathcal{L}_{1Con}$	0.04	0.844	0.24	0.624	0.00	0.990

TABLE 10: FT articles by section, 1899–2010

Uncertainty is based on 297 words from the Loughran and McDonald (2011) word lists. Negative and Positive are based on 2,337 and 353 words from Loughran and McDonald (2011). Our pessimism measure is the daily difference between the percentage of negative and positive words. Economic and political words are taken from the Harvard Psychosocial Dictionary ECON and POLIT categories. "ECON@ comprises 510 words of an economic, commercial, industrial, or business orientation, including roles, collectivities, acts, abstract ideas, and symbols, including references to money. Includes names of common commodities in business." "POLIT@ 263 words having a clear political character, including political roles, collectivities, acts, ideas, ideologies, and symbols." http://www.wjh.harvard.edu/~inquirer/homecat.htm

	All Sections	News in Brief	Markets	Editorials	Lex
Total Columns	316,787	115,537	118,014	58,304	24,932
Total Words	188,900,000	44,870,000	71,430,000	48,960,000	22,770,000
Words per Column	596	388	605	840	913
Uncertainty, %	0.76	0.55	0.56	1.01	1.2
Economic, %	4.30	4.53	4.37	4.01	5.25
Political, %	1.36	1.79	1.01	1.88	0.95
Negative, %	1.64	1.59	1.46	2.06	1.69
Positive, %	0.95	0.68	0.99	1.05	0.97
Pessimism, %	0.68	0.91	0.48	1.02	0.73

TABLE 11: Feedback from FT sections to the blue-chip index

This table reports β coefficients and t-stats from the model $R_t = \beta\mathcal{L}_s(M_t) + \gamma\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index returns. M_t is one of our media measures. X_t is our exogenous variables including day-of-the-week dummies, month-of-the-year dummies, editor dummies and a dummy for whether date t is a bull or bear market.

	(1)	(2)	(3)	(4)	(5)	(6)
	Negative	*t-stat*	Positive	*t-stat*	Pessimism	*t-stat*
	PANEL A: News in Brief (13,626 Obs.)					
$\mathcal{L}_1$	-0.007	-0.72	0.021	2.73	-0.017	-1.71
$\mathcal{L}_2$	-0.014	-1.43	0.002	0.28	-0.012	-1.36
$\mathcal{L}_3$	-0.001	-0.08	-0.002	-0.20	0.001	0.05
$\mathcal{L}_4$	0.002	0.22	-0.001	-0.09	0.002	0.25
$\mathcal{L}_5$	0.015	1.53	0.011	1.47	0.008	0.78
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	0.52	0.470	7.45	0.006	2.94	0.087
$\mathcal{L}_{2-5} = 0$	0.99	0.413	0.60	0.666	0.56	0.690
	PANEL B: Markets (16,557 Obs.)					
$\mathcal{L}_1$	-0.015	-1.60	0.017	2.056	-0.023	-2.44
$\mathcal{L}_2$	0.017	1.80	0.017	1.98	0.004	0.40
$\mathcal{L}_3$	0.006	0.64	-0.002	-0.27	0.006	0.68
$\mathcal{L}_4$	0.024	2.63	-0.003	-0.40	0.022	2.39
$\mathcal{L}_5$	0.007	0.71	-0.018	-2.16	0.016	1.77
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	2.56	0.110	4.23	0.040	5.94	0.015
$\mathcal{L}_{2-5} = 0$	3.06	0.016	2.03	0.087	2.56	0.037
	PANEL C: Editorials and Leaders (17,312 Obs.)					
$\mathcal{L}_1$	-0.009	-0.93	0.009	1.32	-0.012	-1.35
$\mathcal{L}_2$	0.001	0.16	-0.009	-1.22	0.006	0.70
$\mathcal{L}_3$	0.008	0.83	0.006	0.77	0.003	0.40
$\mathcal{L}_4$	0.006	0.62	-0.006	-0.78	0.008	0.86
$\mathcal{L}_5$	0.015	1.65	0.014	1.91	0.005	0.55
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	0.86	0.354	1.74	0.187	1.82	0.177
$\mathcal{L}_{2-5} = 0$	1.06	0.373	1.41	0.228	0.45	0.772
	PANEL D: Lex (11,686 Obs.)					
$\mathcal{L}_1$	-0.023	-2.11	0.002	0.20	-0.018	-1.84
$\mathcal{L}_2$	0.024	2.23	-0.006	-0.62	0.020	2.12
$\mathcal{L}_3$	0.006	0.51	0.020	2.14	-0.007	-0.75
$\mathcal{L}_4$	-0.004	-0.39	-0.018	-1.94	0.007	0.71
$\mathcal{L}_5$	0.015	1.41	-0.025	-2.73	0.025	2.58
	Tests					
	t-stat	*p*-val	*t-stat*	*p*-val	*t-stat*	*p*-val
$\mathcal{L}_1 = 0$	4.45	0.035	0.04	0.841	3.38	0.066
$\mathcal{L}_{2-5} = 0$	2.25	0.061	3.89	0.004	3.48	0.008

TABLE 12: Feedback from Extreme News Content to Abnormal Trading

This table reports β coefficients and t-stats from the model $\hat{\tau}_t = \beta\mathcal{L}_s(|M_t|) + \varphi\mathcal{L}_s(\hat{\tau}_t) + \zeta\mathcal{L}_s(R_t) + \psi\mathcal{L}_s(R_t^2) + \eta X_t + \epsilon_t$. R_t is our daily stock market index returns. $|M_t|$is the absolute value of one of our standardized media measures. X_t is our exogenous variables including editor dummies and a dummy for whether date t is a bull or bear market. $\hat{\tau}_t$ is abnormal trading which is calculated from the equation $\tilde{\tau}_t = \beta\mathcal{L}_s(\tau_t) + \eta X_t + \varepsilon_t$. Where τ is log number of trades (000's). The residuals are standardized by the annual standard deviation in the residuals. Thus $\hat{\tau}_t = \frac{\varepsilon_t}{\bar{\sigma}_\varepsilon}$ which results in a detrended and white noise variable of abnormal trading. $\hat{\tau}_t$ is standardized to have a mean of zero and unit variance. Data begins on February 6, 1930 and ends June 30, 2008. Trading data is not available from November 6, 1973 to December 28, 1979.

VARIABLES	(1) Negative	(2) *t-stat*	(3) Positive	(4) *t-stat*	(5) Pessimism	(6) *t-stat*
	Panel A: News in Brief (11,465 obs.)					
$\mathcal{L}Abs_1$	-0.002	-0.23	-0.014	-1.43	-0.017	-1.78
$\mathcal{L}Abs_2$	-0.008	-0.87	0.016	1.67	0.009	0.94
$\mathcal{L}Abs_3$	-0.002	-0.26	0.012	1.05	0.005	0.55
$\mathcal{L}Abs_4$	-0.007	-0.77	0.015	1.56	0.004	0.43
$\mathcal{L}Abs_5$	0.006	0.68	-0.008	-0.91	-0.002	-0.19
	t-stat.	*p-val.*	*t-stat.*	*p-val.*	*t-stat.*	*p-val.*
$\mathcal{L}_1 = 0$	0.05	0.817	2.03	0.154	3.16	0.076
$\mathcal{L}_{2-5} = 0$	0.49	0.744	1.81	0.124	0.41	0.804
	Panel B: Markets (12,095 obs.)					
$\mathcal{L}Abs_1$	0.027	2.94	-0.003	-0.36	0.016	1.82
$\mathcal{L}Abs_2$	-0.013	-1.36	0.012	1.32	0.018	1.85
$\mathcal{L}Abs_3$	0.004	0.48	0.013	1.41	0.013	1.54
$\mathcal{L}Abs_4$	-0.008	-0.95	-0.000	-0.05	0.004	0.49
$\mathcal{L}Abs_5$	-0.006	-0.68	0.009	1.08	-0.009	-1.05
	t-stat.	*p-val.*	*t-stat.*	*p-val.*	*t-stat.*	*p-val.*
$\mathcal{L}_1 = 0$	8.63	0.003	0.13	0.717	3.30	0.069
$\mathcal{L}_{2-5} = 0$	0.88	0.475	1.27	0.280	1.76	0.134
	Panel C: Editorials and Leaders (12,828 obs.)					
$\mathcal{L}Abs_1$	0.011	1.23	0.000	0.01	0.008	0.90
$\mathcal{L}Abs_2$	0.009	1.07	0.005	0.55	0.010	1.12
$\mathcal{L}Abs_3$	-0.022	-2.31	0.014	1.52	-0.013	-1.41
$\mathcal{L}Abs_4$	-0.007	-0.77	-0.014	-1.57	-0.010	-1.09
$\mathcal{L}Abs_5$	0.001	0.11	0.016	1.86	0.003	0.30
	t-stat.	*p-val.*	*t-stat.*	*p-val.*	*t-stat.*	*p-val.*
$\mathcal{L}_1 = 0$	1.52	0.218	0.00	0.993	0.80	0.370
$\mathcal{L}_{2-5} = 0$	1.77	0.131	2.16	0.071	1.12	0.345
	Panel D: Lex (9,496 obs.)					
$\mathcal{L}Abs_1$	0.009	0.79	0.012	1.10	0.000	0.02
$\mathcal{L}Abs_2$	0.014	1.22	-0.003	-0.32	0.012	1.04
$\mathcal{L}Abs_3$	0.008	0.73	0.012	1.14	0.014	1.29
$\mathcal{L}Abs_4$	0.030	2.99	-0.001	-0.12	0.015	1.54
$\mathcal{L}Abs_5$	0.006	0.64	0.010	0.92	0.008	0.84
	t-stat.	*p-val.*	*t-stat.*	*p-val.*	*t-stat.*	*p-val.*
$\mathcal{L}_1 = 0$	0.62	0.432	1.20	0.273	0.00	0.985
$\mathcal{L}_{2-5} = 0$	3.43	0.008	0.56	0.691	1.54	0.188

Appendix 1—Sample Articles & Textual Output

BOER LEADERS AT PRETORIA.

RUMOURED RECEIPT OF IMPORTANT NEWS IN LONDON.

Pretoria, Sat., 12th April.

Mr. Schalk Burger, General Louis Botha, Commandant Lucas Meyer, General Delarey, Mr. Steyn and General De Wet arrived here this morning by special train from Klerksdorp.—[Reuter.]

Utrecht, Sat., 12th April.

Dr. Leyds had a conference yesterday evening with the members of the Boer deputation here. The Boer delegates had another conference with Dr. Leyds this afternoon. At the close of the conference Dr. Leyds, on being asked whether the meeting which he came to attend had been summoned with reference to the peace negotiations, said that the conference had been merely an ordinary one such as he was in the habit of holding with the members of the Boer deputation.—[Reuter.]

PEACE NEGOTIATIONS.
BOER LEADERS AT PRETORIA
RUMOCRED RECEIPT OP IMPORTANT XE WS IN LONDON.
Pretoria, Sat., 12th AprIL Mr. Schalk Burger, General Ixmis Botha, Commandant Locaa Meyer, General Delarey, Mr. Steyn and General De Wet arrired her« this morning by special train from Klerka-dorp.—[Reuter.]
Utrecht, Rat., 12th April.
Dr. Leyds had a Conference vesterday erening with the membera of the Bo er deputation here. The Boer delegates had , another oonfcrcnce with Dr. Leyds this I afternoon. At the closc of the oonferenc« 1 Dr. Leyda, on being asked whether the mect-ing which he came to attend had been sum-moned with re fe re nee to the peace negotia-tiona, said tliat the oonference had been merely an ordinary one such as he was in the habit of holding with th« membere of the Boer deputation.—

Example of 'News in Brief' article published on 15th June 1903.

NEW YORK BANK STATEMENT.

The return of the New York Associated Banks gives further evidence of a welcome improvement in the financial conditions in Wall Street, and, though its effect had been so discounted that the market weakened after the publication of the statement, this was due merely to profit-taking sales and did not reflect any disappointment with the position disclosed. There is a further sharp reduction in the loan account, which is matched by an almost identical decrease in net deposits, while the cash reserves are actually increased by $1,760,000. The gain in surplus reserves amounts to $4,700,000, and the balance above the legal minimum is now back again practically at the figure at which it stood a fortnight ago.

OCR output after cleansing processes applied:

Tho return of the New York Associated Banks givcs further evidonoe of a welcomo improvement in the financial condition» in Wall Street, and, thougli it» eflect had boen &•> discountcd that the market weakcned aftor the publication of the Statement, this was due mcrely to profit-taking »nies and did not reflect any disappointment with the position disclose<1. There is a further sharp reduc-tion in the loan account, which is matched by an almost identical decrease in not deposits, while the cash reserves an; actually increosed by .$1,760,000. The gain in surpIus reserves amount» to $4,700.000. and the halanco aboro the legal minimum is now back again practica 1 ly nt the tigure at which it stood a fortnight ago.

Example ‘Markets’ article published 22nd December 1992

Iranian call for Opec output cut gives oil market brief fillip

By Deborah Hargreaves

OIL PRICES edged upwards yesterday in light trading in the run-up to Christmas, only to slip back again later in the day to close unchanged on the day. North Sea Brent crude for February delivery was at $18.55 a barrel.

Prices got a brief fillip early in the day from reports in the Middle East Economic Survey, which quoted Mr Gholamreza Aqazadeh, Iran's oil minister, calling for a 2 per cent cut in output by the Organisation of Petroleum Exporting Countries.

A 2 per cent cut in production would take Opec's ceiling to 24.09m barrels a day from its present level of 24.58m b/d. But many market traders are sceptical that members have even implemented the slight cutbacks required under their latest agreement.

Traders at one international oil company said there was little physical evidence of Opec cutbacks.

"Iran is talking up the oil price," said Mr Fareed Mohammedi at the Petroleum Finance Company in Washington, "I can't see them getting anyone to agree to a meeting before February".

Oil prices have found some support in the past week after being extremely weak since October, but traders believe that, if Opec members fail to cut their production in the New Year, prices will continue to slide.

● Russia and Kazakhstan are looking at plans to join together energy producing states from the former Soviet Union in a "mini Opec", the Reuter news agency reported yesterday. The agreement was reached at the weekend in talks between Mr Viktor Chernomydrin, Russian prime minister and Mr Nursultan Nazarbayev, Kazakhstan's president.

Russia and Kazakhstan were reported last month to be interested in joining Opec.

OCR output after cleansing processes applied:

OIL PRICES edged upwards yesterday in light trading in the run-up to Christmas, only to slip back again later in the day to close unchanged on the day. North Sea Brent crude for February delivery was at $18.55 a barrel.
Prices got a brief fillip early in the day from reports in the Middle East Economic Survey, which quoted Mr Gholamreza Aqazadeh, Iran's oil minister, calling for a 2 per cent cut in output by the Organisation of Petroleum Exporting Countries.
A 2 per cent cut in production would take Opec's ceiling to 24.09m barrels a day from its present level of 24.58m bid. But many market traders are sceptical that members have even implemented the slight cutbacks required under their latest agreement.
Traders at one international oil company said there was little physical evidence of Opec cutbacks.
"Iran is talking up the oil price." said Mr Fareed Moham-medi at the Petroleum Finance Company in Washington, “I can’t see them getting anyone to agree to a meeting before February".
Oil prices have found some support in the past week after being extremely weak since October, but traders believe that, if Opec members fail to cut their production in the New Year, prices will continue to slide.
• Russia and Kazakhstan are looking at plans to join together energy producing states from the former Soviet Union in a “mini Opec", the Reuter news agency reported yesterday. The agreement was reached at the weekend in talks between Mr Viktor Chemomydrin, Russian prime minister and Mr Nursultan Nazarbayev. Kazakhstan’s president.
Russia and Kazakhstan were reported last month to be interested in joining Opec.