Journal of Behavioral and Experimental Finance 39 (2023) 100825

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Finance

journal homepage: www.elsevier.com/locate/jbef

Full length article

Lost in translation. When sentiment metrics for one market are derived from two different languages

Robert B. Durand [a], Joyce Khuu [a], Lee A. Smales [b,*]

[a] Finance Discipline, Curtin University, Western Australia, Australia
[b] UWA Business School, University of Western Australia, Western Australia, Australia

ARTICLE INFO

Article history:
Received 16 February 2023
Received in revised form 2 May 2023
Accepted 12 June 2023
Available online 16 June 2023

Dataset link: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

JEL classification:
G12
G14
G41

Keywords:
Behavioral finance
Sentiment
News analytics
TOPIX

ABSTRACT

We compare two text-based proxies for the sentiment of investors in the Japanese market. Both proxies are constructed by Thomson Reuters using the same algorithm, and the only difference between them is that the first proxy is derived using only Japanese language items; the second is derived from English language items. The correlation between the proxies is low and this suggests that they measure different aggregate affective states. The English-language sentiment proxy is found to have a positive and statistically significant association with Japanese returns before a key date in the history of Abenomics relating to quantitative easing (April 2013). After this date, only the Japanese-language proxy has a positive and statistically significant association with the market. Studies of sentiment are predominantly based on proxies derived from English language sources; our results contribute to this nascent research field and suggest that not using both the English and local-language sources when constructing text-based sentiment indices may lead to misleading or mistaken inferences.

1. Introduction

The association between investor sentiment and market returns is now well accepted. How investor sentiment might be measured remains an open question. Hirshleifer and Shumway (2003) and Kamstra et al. (2003) propose using weather and, or, the hours of darkness as sentiment proxies. Sentiment is derived from observable market, or macroeconomic, phenomena or general surveys of consumer sentiment (Brown and Cliff, 2005; Baker and Wurgler, 2006). Sentiment metrics are also derived from text, either using a "bag of words" methodology (Tetlock, 2007) or machine learning. The standard interpretation of such sentiment metrics is that they capture the *Zeitgeist* (spirit) of the market. The key research question examined in this paper is whether using *only* English-language sources to derive sentiment is appropriate in non-English speaking countries. In particular, can using *only* English-language sources capture the *Zeitgeist* of a market where the local language is *not* English?

This paper studies the Japanese market from January 2003 until September 2020 and finds that analyses that do not use *both* Japanese and English-based sentiment indices will lose much in translation.[1] In doing so, we believe our major contribution to the literature is to provide guidance to researchers who wish to use text-based sentiment metrics for markets where English and local-language sources could be used. Not using both the English and local-language sources when constructing text-based sentiment indices may lead to misleading or mistaken inferences.

We find that each sentiment index has played a very different role in Japan's market. We run a structural break test and find a structural break on 5th April 2013. This statistical finding aligns with the Bank of Japan's announcement of a renewal of quantitative easing on 4th April 2013. This is an important date in the history of Abenomics (reforms introduced by Prime Minister Abe's government in 2012 aimed a revivifying Japan's moribund economy). Shinzo Abe began introducing these reforms after his election on the 16th of December 2012 witnessed the election of Shinzo Abe to be Prime Minister of Japan. Subsequently, the Bank of Japan announced a renewal of quantitative easing on 4th April 2013 (the date preceding the structural break our tests

* Corresponding author.
E-mail address: lee.smales@uwa.edu.au (L.A. Smales).

[1] Prior to this paper, only Khuu et al. (2016) have considered the effect of news sentiment on the Japanese market, and they derive sentiment *only* from the English-language version of TRNA.

https://doi.org/10.1016/j.jbef.2023.100825

identify), and further authority to enact economic policy was provided following the Diet elections of July 2013.[2] We explore the implications of this break in Section 3.3 and the most important finding reported in this section is that the sentiment index derived from English-language sources is associated with Japanese stock returns before Abenomics. After April 2013, the relationship between the English-language based index disappears. Instead, the Japanese language index is found to have a positive relationship with Japanese returns. The English language index and the Japanese index are measuring quite different affective states: the Japanese and English sentiment indices have low or no correlations. Chan et al. (2017) prove that proxies for the same underlying phenomenon should be correlated and, in their empirical demonstration of this finding, show that the well-known Thomson Reuters News Analytics (TRNA) sentiment index is uncorrelated with Baker and Wurgler's metric.[3] The two Japanese market indices do not appear to be reflecting a coherent view of feelings about the market; rather, they capture two very differing world views. Additionally, descriptive statistics reported in Section 3.3 (Table 7, Panel A) also report changes in the production and sentiment of the news. Daily publications in both languages increase in the post-Abenomics period. Japanese sentiment declines post-Abenomics while English language sentiment increases.

The literature has utilized a wide range of measures to proxy for investor sentiment. This includes surveys (Lee et al., 2002), consumer confidence (Lemmon and Portniaguina, 2006), trader positions (Wang, 2003), and other market-based indicators (Baker and Wurgler, 2006). Starting with Tetlock (2007), textual analysis has been utilized to derive sentiment measures from newspapers (Tetlock, 2007), Google (Da et al., 2015), and Twitter (Gu and Kurov, 2020).[4] Most recently, technological developments have enabled sentiment to be processed from photographs (Obaid and Pukthuanthong, 2022; Chiah et al., 2022). We have also witnessed the emergence of financial data providers, such as Thomson Reuters, Refinitiv, and Ravenpack, who utilize natural language processing algorithms to generate text-based sentiment score in real-time.

Our study utilizes text-based sentiment measures that are provided by Thomson Reuters News Analytics[5] (TRNA). TRNA uses neural network machine learning to classify the sentiment associated with news stories. By examining sentences, rather than individual words, the algorithm can provide a contextual word analysis. TRNA offers coverage of over 35,000 global equities and, in addition to academic research, is used by low latency traders via direct connections in Chicago, Frankfurt, London, New York, and Tokyo. Several recent studies have used this dataset.[6] However, the existing research has focused exclusively on English language news articles, while we are able to exploit a newly available Japanese language dataset. There is recent evidence that suggests it is difficult to precisely translate from one language into another, particularly for technical terms such as sentiment wordlists. Du et al. (2022) show that it is possible to create a sentiment dictionary directly from Chinese financial news articles that outperforms dictionaries based on English translations. The TRNA sentiment measures follow a similar process to Du et al. (2022) in producing a sentiment score directly from Japanese language articles.[7] We use *only* Japanese (local) language sources to create the Japanese (local) language sentiment index and only English articles to create the English language index. Given the growing number of high-frequency traders who trade based on signals from news analytics software (Von Beschwitz et al., 2020) the questions we pose are of interest to the wider investment community as well as academics.

Prior research has focused on the English language version of the TRNA sentiment proxy; this is one of the proxies that we apply to Japanese stock markets in this paper. Much of this prior analysis focuses on stock markets, with investors responding to new information from sentiment proxies in US (Dzielinski, 2011; Boudoukh et al., 2013; Uhl, 2014; Heston and Sinha, 2017; Allen et al., 2019), British (Groß-Klußmann and Hautsch, 2011), and Japanese (Khuu et al., 2016) markets. The magnitude of the response varies across industries and over time (Smales, 2015) and is asymmetric, whereby negative news induces a stronger market response (Riordan et al., 2013; Smales, 2014; Uhl, 2014). Leinweber and Sisk (2011) show that trading strategies can be constructed to produce alpha from news sentiment, while Hendershott et al. (2015) demonstrate that institutions may be informed in advance of news since their trading volume can predict news announcements and associated sentiment. Dang et al. (2015) use Ravenpack, another news analytics tool, to demonstrate that production of firm-specific information is dependent on a country's institutional environment.[8] Calomiris and Mamaysky (2019) use articles from the Thomson Reuters Machine Readable News archive (a more extensive source of information than that used by TRNA) to derive sentiment metrics for fifty-one markets and find that sentiment has predictive power for returns. Calomiris and Mamaysky's analysis is based on English word lists from English articles.

Japan provides an instructive market to analyze the relationship between a local-language based metric and an English-based metric due to the considerable proportion of foreign holdings of Japanese stocks. Foreign investors play a significant role in the Japanese market. Iwatsubo and Watkins (2021) note that the presence of foreign investors improves the informational efficiency of the market, and they appear to have an information advantage over domestic investors. In the decade to 2021, around 60% of trading value in the largest Japanese firms (TSE 1st section) was conducted by foreigners.[9] With 30% of trading value the TSE 2nd section, foreigners also play an increasingly important role in ownership of smaller Japanese firms. Additionally, Japan is a major world market; in 2020, the final year of our study, it was the third largest by market capitalization (after the United States and China).[10]

2 Further information may be found at: https://www.economist.com/finance-and-economics/2020/09/03/did-abenomics-work ; https://www.economist.com/leaders/2016/07/30/overhyped-underappreciated ; and https://www.economist.com/finance-and-economics/2022/07/14/the-legacy-of-abe-shinzo-will-shape-japans-economy-for-years.

3 Chan et al. argue that "…At least one, but perhaps all, of these are not valid proxies of sentiment" (p. 477).

4 Loughran and McDonald (2016, 2020) and Kearney and Liu (2014) provide surveys of the textual analysis literature in finance.

5 After the initiation of this project, Refinitiv has demerged from Thomson Reuters and the dataset is now called Refinitiv News Analytics.

6 See for instance: Dzielinski (2011), Groß-Klußmann and Hautsch (2011), Leinweber and Sisk (2011), Boudoukh et al. (2013) Riordan et al. (2013), Uhl (2014), Smales (2014, 2015), Khuu et al. (2016) Heston and Sinha (2017), and Allen et al. (2019).

7 Liu, Lee, Huang and Wu (2023) have recently used Chinese social media data to derive sentiment indices to analyze the relationship of Chinese returns and sentiment.

8 Sentiment signals generated from software analyzing newswire messages (i.e., Ravenpack, TRNA) are also examined in the context of other assets, including commodity and credit markets. News sentiment can explain movements in the benchmark US term structure (Gotthelf and Uhl, 2019), in CDS markets (Cathcart et al., 2020), and in bank credit risk (Smales, 2016). Commodity markets are also shown to respond to sentiment, with the response dependent on market conditions, such as credit availability, levels or reserves, and consumer confidence (Smales, 2014; Clements and Todorova, 2016; Gupta and Banerjee, 2019).

9 Source: https://www.jpx.co.jp/english/markets/statistics-equities/investor-type/00-02.html.

10 Source: https://www.world-exchanges.org/our-work/articles/2020-annual-statistics-guide

Table 1
Article filtering.

	Japanese articles	English articles
All news	22,55,730	13,39,024
Novel	15,93,265	7,97,981
Relevant	12,00,406	8,38,532
Novel and relevant	8,09,549	5,08,662
Remove holidays/weekend	7,76,706	4,94,923
% articles during trading day	16.9%	22.2%
% articles outside trading day	83.1%	77.8%

Note: This table presents information on the filtering process for newswire articles published in Japanese and English languages. *Novel* shows the number of articles that remain in the sample after items with a novelty score greater than 0 are removed. *Relevant* shows the number of articles remaining after those with relevance scores of less than 0.7 are removed. *Novel and relevant* reports the number of articles that meet the criteria to be both novel and relevant. The proportion of articles published during, and outside, the Japanese trading day (9am–3pm) is also indicated.
Sample period: January 2003–September 2020.

2. Data and method

The TRNA dataset describes each item with more than ninety metadata fields, including identifiers, topic codes, and sector classification. For our purposes, the salient data comprises:

1. *Identifier* of the stock mentioned in the news article;
2. *Timestamp* indicating the arrival time of the news article;
3. *Sentiment class* which is a discrete variable indicating whether the news article was most likely to be positive (+1), neutral (0), or negative (−1);
4. *Sentiment probabilities* which are a set of variables signifying the probability of assigning the article portraying positive, neutral, or negative news;
5. *Relevance* which is a continuous variable on the [0,1] interval indicating how prominently the firm was mentioned in the news article. A relevance score of 1 generally indicates that the firm was mentioned directly in the headline; and
6. *Novelty* which shows whether the news item is linked to other similar articles or whether it is in a sequence of updated news items.

We collect news sentiment data for Japanese and English language articles related to Japanese stocks that are TOPIX constituents.[11] The articles related to firm-specific news and articles published in the two languages do not necessarily represent the same news event. Coinciding with TRNA data availability, our sample period runs from 1st January 2003 until 1st September 2020 (4315 trading days). Our initial sample contains 2,255,730 Japanese language articles and 1,339,024 English language articles. Since efficient markets suggest prices should only respond to new information, and the literature (e.g., Groß-Klußmann and Hautsch, 2011; Boudoukh et al., 2013) shows that relevant news has a significant impact on prices, we filter for novelty and relevance. Articles with a novelty score greater than 0[12] or a relevance score less than 0.7 are removed so that our final sample consists of 809,549 Japanese language articles and 508,662 English language articles. Table 1 shows the filtering process and indicates that most articles occur outside of the standard Japanese trading day.

We construct a daily sentiment score for the market by aggregating the sentiment of all articles, for all firms, arriving during that day. The 'day' is determined by the Tokyo time zone since we focus on news related to stocks traded there. Our principal measure is a probability weighted sentiment score, PSENT, that is constructed for each language k by multiplying the sentiment class attached to a news article published in that language by the assigned probability that it is correctly categorized.

$$PSENT_{k,t} = \frac{\sum(1)\cdot P\,(positive) + \sum(-1)\cdot P\,(negative)}{\eta_{positive} + \eta_{neutral} + \eta_{negative}} \in [-1, 1] \quad (1)$$

Although neutral news articles do not enter the sentiment calculation in the numerator (owing to their sentiment of 0) the number of articles appears in the denominator (i.e., $\eta_{positive} + \eta_{neutral} + \eta_{negative}$ = sum of all articles), placing downward bias on the measure. This ensures that a day with many neutral news articles will still have PSENT close to zero.[13] There are few news articles that arrive on a weekend or public holiday (32,843 Japanese and 13,739 English language articles in the whole sample period) and when this does happen the news sentiment is assigned to the next trading day. We also construct two alternative aggregate measures, TSENT and ASENT, which are used in robustness tests. TSENT is simply the cumulative total of sentiment class scores occurring during day t, while ASENT is an average sentiment score obtained by dividing TSENT by the number of articles published during day t.

Descriptive statistics for daily numbers of news articles and associated sentiment scores are shown in Table 2, Panel A.[14] Japanese language articles are more numerous and have more positive sentiment than English language articles. On average, there are 188 Japanese articles per day with PSENT of 0.057 compared to 118 English articles with PSENT of 0.020. Disaggregating by time of day (Panel B and Panel C) reveals that Japanese articles are more positive outside of the Japanese trading day (PSENT 0.055 against 0.007), while English language articles are more positive during the trading day (PSENT 0.014 against 0.002). Since the articles published in each language do not necessarily focus on the same event, the difference in sentiment is not necessarily due to a disparity in language, it could also be a result of Japanese and English writers focusing on different issues. For instance, English writers may focus on international aspects of a news event while Japanese writers frame events in terms of domestic concerns. We test for unit roots in our time series using augmented Dickey–Fuller with trend and intercept and confirm that all series are stationary.

11 TRNA provides sentiment scores for articles written in Japanese language and in English language. We do not attempt to compute our own scores by applying an algorithm to news articles.

12 A novelty score greater than 0 indicates that the news has previously been reported. Removing these news items from our sample fits with the notion of market efficiency such that asset prices should respond only to new information. An additional benefit of this filtering process is that it prevents the possibility of our sample becoming biased by articles that are updated many times. In Appendix A (column 1) we show that the results of our baseline regression are qualitatively similar if we do not exclude non-novel items.

13 Appendix A (column 2) shows that our results hold if we ignore articles deemed to have neutral sentiment.

14 Appendix B illustrates the times series of daily PSENT and TOPIX log returns.

Table 2
Descriptive statistics: Articles and sentiment in different languages.

Panel A: ALL	Japanese Language				English Language				diff in mean
Variable	Mean	Std Dev	Min.	Max.	Mean	Std Dev	Min.	Max.	(t-test)
No. articles	187.68	168.91	0.00	2298	117.86	121.21	0.00	1445	−22.06***
PSENT	0.057	0.10	−0.36	0.59	0.020	0.17	−0.66	0.66	−12.43***
TSENT	12.77	52.58	−982	1633	3.261	37.71	−314	421	−9.66***
ASENT	0.081	0.13	−0.50	0.75	0.040	0.22	−0.82	0.81	−10.33***
Panel B: During trading day (DAY)									diff in mean
Variable	Mean	Std Dev	Min.	Max.	Mean	Std Dev	Min.	Max.	(t-test)
No. articles	25.75	19.43	0.00	388	14.26	9.85	0.00	145	−34.65***
PSENT	0.002	0.03	−0.36	0.39	0.014	0.05	−0.31	0.43	13.14***
TSENT	11.39	41.02	−982	248	1.29	5.70	−50	130.00	−16.02***
ASENT	0.005	0.04	−0.51	0.49	0.021	0.08	−0.43	0.62	12.45***
Panel C: Outside trading day (OUT)									diff in mean
Variable	Mean	Std Dev	Min.	Max.	Mean	Std Dev	Min.	Max.	(t-test)
No. articles	161.92	157.79	0.00	2168	103.60	118.09	0.00	1411	−19.44***
PSENT	0.055	0.09	−0.36	0.59	0.007	0.15	−0.61	0.65	−17.60***
TSENT	12.11	51.21	−983	1580	1.97	36.14	−282	418	−10.63***
ASENT	0.076	0.12	−0.50	0.75	0.020	0.20	−0.76	0.79	−15.34***

Panel D: Average articles/ sentiment and TOPIX returns by day of week

	Japanese Language Articles			English Language Articles			TOPIX
	No. articles	ASENT	PSENT	No. articles	ASENT	PSENT	Returns
Monday	190.08	0.08	0.06	103.31	0.03	0.01	0.015%
Tuesday	169.33	0.09	0.06	110.36	0.07	0.04	0.005%
Wednesday	164.69	0.09	0.06	105.54	0.06	0.03	0.047%
Thursday	182.88	0.10	0.07	121.14	0.06	0.03	0.019%
Friday	231.57	0.05	0.03	147.33	**−0.02**	**−0.02**	−0.012%

Note: This table presents descriptive statistics for the number of articles and alternate sentiment measures used in this study. Articles appear in Japanese language and English language, and articles appearing during the weekend or public holiday are incorporated into the next trading day. *PSENT* is the probability weighted sentiment score, *TSENT* is cumulative total sentiment, and *ASENT* is average sentiment. Panel A reports statistics for the overall sample, Panel B reports statistics for articles occuring during the Japanese trading day (9am–3pm), Panel C reports statistics for articles occuring outside the Japanese trading day, and Panel D shows article statistics by day of the week together with TOPIX index returns.
Sample period: January 2003–September 2020.

Panel D of Table 2 shows the mean value of articles and sentiment for each day of the week, together with an indication of average TOPIX log returns for that day. Friday sees the peak in article numbers and the low in sentiment for both languages. Indeed, while PSENT is positive across all days for Japanese language articles, on Friday there is average negative sentiment for English language articles. The low sentiment shown for Friday coincides with the lowest average TOPIX return.

Table 3 reports correlation analysis for the key variables used in our study. Panel A shows that, for the whole sample, the sentiment measures are positively correlated with each other but in general are negatively correlated with the number articles published. Article numbers in each language are also positively correlated and to a much higher degree (0.93) than exists between sentiment measures.[15] More importantly, as we noted in the opening paragraph of this paper, measures of sentiment are commonly interpreted as representing the underlying *Zeitgeist* of the market. PSENT_JA and PSENT_EN have a small, statistically significant positive correlation (0.079), suggesting that they may be capturing the same underlying behavioral phenomenon (Chan et al., 2017) although this correlation would be consistent with very noisy perceptions of some underlying reality.

Fig. 1 provides further disaggregation of the data, illustrating the number of articles published and associated sentiment per hour. Panel A shows that article publications in both languages are concentrated in the morning, and peak in the 6am–8am interval which occurs before the Tokyo market opens. It is possible that this period is important owing to the propensity of Japanese firms make corporate announcements, including earnings releases, during this time. We note that 7am in Tokyo is 6pm in New York, creating the possibility that some of the articles relate to events that occurred during the prior day's New York session, and utilize this in one of our later tests. Panel B shows that Japanese language sentiment is highest before the market opens (and even tends to be negative just after it closes) when articles are more frequently published. In contrast, English language sentiment is highest after lunch in Tokyo, when the number of publications is lowest. This discrepancy is one reason why Japanese PSENT is higher on average than English PSENT.

3. Empirical analysis

Throughout our empirical analysis, we focus on examining the daily contemporaneous effect of sentiment on stock returns. To enable this, we construct a series of daily log returns, R_t = log(P_t / P_{t-1}), for TOPIX and individual firms using data obtained via DataStream.[16] Each of our regression models is then a variation of the following specification:

$$R_t = \beta_c + \beta_{JA} PSENT_{JA,t} + \beta_{EN} PSENT_{EN,t} + \sum_{k=1}^{2} \gamma_k log\left(articles_{k,t}\right) + \delta_j X_{j,t} + \varphi R_{t-1} + \varepsilon_t \tag{2}$$

Where R_t is the log return on day t, PSENT_JA and PSENT_EN are the sentiment measures for each language, articles $_k$ is the number of published articles in language k, X is a set of control

15 We discuss correlation pre- and post-Abenomics in Table 6 below.

16 Our analysis focuses on returns for the widely followed TOPIX price index. During our sample period, the correlation with the total return index is 0.9985. This suggests there would be a minimal impact on our reported results, and this is confirmed by the qualitatively similar baseline results shown in Appendix A (column 4).

Table 3
Correlation analysis.

	R_TOPIX	PSENT_JA	PSENT_EN	PSENT_NY	ARTICLES_JA	ARTICLES_EN	ARTICLES_NY
R_TOPIX		**0.113**	**0.082**	**0.122**	−0.018	−0.015	−0.025
PSENT_JA	**0.104**		**0.079**	**0.192**	**−0.140**	**−0.067**	**−0.207**
PSENT_EN	**0.079**	**0.102**		**0.153**	−0.026	**−0.132**	**0.134**
PSENT_NY	**0.137**	**0.209**	**0.169**		−0.038	−0.014	**−0.064**
ARTICLES_JA	−0.003	**−0.113**	0.008	−0.004		**0.733**	**0.285**
ARTICLES_EN	−0.011	**−0.075**	**−0.060**	0.008	**0.930**		**0.174**
ARTICLES_NY	−0.016	**−0.208**	**0.123**	**−0.106**	**0.240**	**0.179**	

Note: This table presents correlation analysis for the key variables used in this study. This includes returns on the TOPIX index (*R_TOPIX*), probability weighted measures of sentiment for articles published on TOPIX stocks in Japanese (*PSENT_JA*) and English (*PSENT_EN*) and US stocks in English (*PSENT_NY*), together with the number of articles for each language (*ARTICLES_JA, ARTICLES_EN, ARTICLES_JA*). Correlation coefficients for the whole sample period are shown, with the bottom left section obtained via Pearson correlation and the top right section via Spearman rank order correlation. **Bold** figures indicate statistical significance at the 1% level. Sample period: January 2003–September 2020.

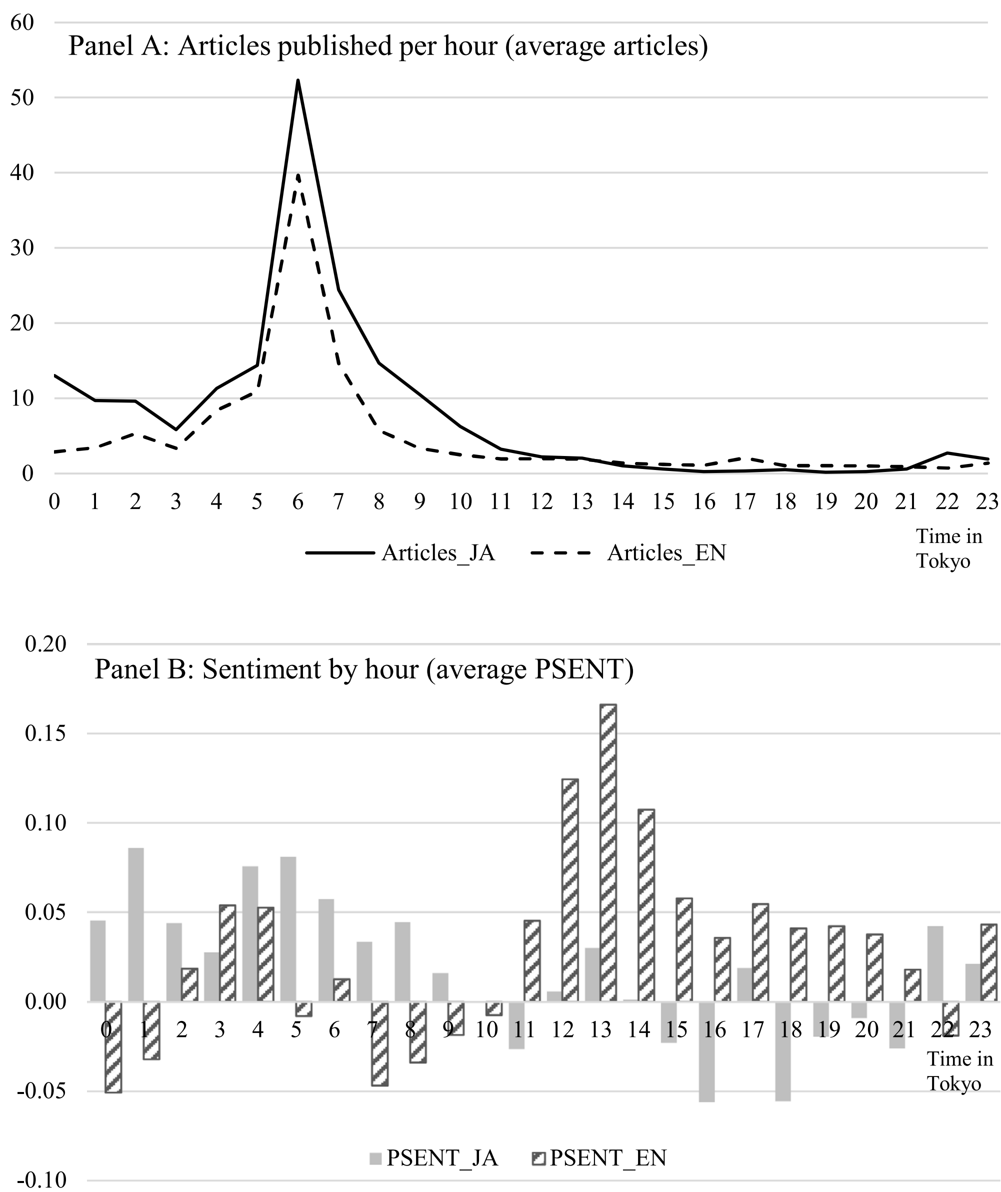

Fig. 1. Hourly sentiment and published articles. *Note:* The horizontal axis represents the time in Tokyo. 0 is midnight, 12 is noon, and 23 is 11pm.

Table 4
Effect of news sentiment and publications on TOPIX returns.

Panel A: PSENT	(1) R_t	(2) R_t	(3) R_t
Constant	−0.072	−0.048	−0.033
	(0.153)	(0.180)	(0.154)
PSENT_JA	1.343***	1.375***	1.375***
	(0.229)	(0.229)	(0.234)
PSENT_EN	0.602***	0.576***	0.560***
	(0.131)	(0.126)	(0.130)
log(articles_ja)	−0.048		
	(0.051)		
log(articles_en)	0.061		
	(0.055)		
log(articles_total)		0.001	−0.002
		(0.031)	(0.031)
R_{t-1}	0.008	0.008	0.008
	(0.023)	(0.023)	(0.023)
Day Control	YES	YES	YES
D-B Control	YES	YES	YES
No. Obs.	4313	4313	4313
Adj. R^2	0.014	0.014	0.014
F-stat	7.15	7.74	7.01
Panel B: ASENT	R_t	R_t	R_t
Constant	−0.113	−0.098	−0.075
	(0.155)	(0.182)	(0.156)
ASENT_JA	0.964***	0.985***	0.986***
	(0.168)	(0.168)	(0.172)
ASENT_EN	0.473***	0.453***	0.437***
	(0.097)	(0.093)	(0.096)
log(articles_ja)	−0.048		
	(0.051)		
log(articles_en)	0.069		
	(0.055)		
log(articles_total)		0.009	0.006
	(0.031)	(0.031)	
R_{t-1}	0.008	0.008	0.008
	(0.023)	(0.023)	(0.023)
Day Control	YES	YES	YES
D-B Control	YES	YES	YES
No. Obs.	4313	4313	4313
Adj. R^2	0.013	0.013	0.013
F-stat	6.72	7.20	6.48

Note: This table reports the estimated coefficients for a regression of the form specified in Eq. (2). The dependent variable is the daily return on the TOPIX index (R_t). In Panel A, the explanatory variables are the probability weighted measures of sentiment derived from Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles, and the natural log of the number of articles published in each language (*articles_ja*, *articles_en*) or across both languages (*articles_total*). Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. In the first two columns news occuring on the weekend or public holiday is incorporated into the next trading day, and in the third column this news is simply ignored. Panel B provides a robustness test by using a measure of average sentiment (*ASENT*). Newey–West standard errors are shown in parentheses.
*** Indicates statistical significance at the 1% level.
Sample period: January 2003–September 2020.

variables including dummy variables to account for day of the week and "Dekansho-bushi"[17] calendar effects, R_{t-1} is the lagged daily return to capture potential information transmission from the prior day, and ε is the error term. We use Newey and West (1987) standard errors to account for serial correlation and heteroskedasticity. We follow, and confirm the analysis presented in, Khuu et al. (2016) and use contemporaneous sentiment.

3.1. Baseline relationship between returns and sentiment

We begin by considering the effect of sentiment and articles published throughout the day. The estimated coefficients are shown in Table 4. We conduct additional diagnostic tests on our baseline model, shown in column 1, to guide subsequent analysis. We find no evidence of serial correlation (the Breusch–Godfrey serial correlation test statistic is 1.10) but heteroskedasticity does seem to be present (Breusch–Pagan–Godfrey test statistic of 12.62) so, to be cautious, we report Newey and West (1987) standard errors.[18] In addition, the high degree of correlation between the two article publication counts (0.93), which largely offset in column 1, suggests the presence of multicollinearity. In this model, the variance inflation factor is 3.29 for Japanese language articles and 3.18 for English articles. As a result, in columns 2 and 3, and subsequent analyses,[19] we focus on a single article count that represents the total number of articles published in both Japanese and English.

17 Sakakibara et al. (2013) report that the "sell in May effect" is not applicable to Japanese stocks. Instead, they find a the "Dekansho-bushi" effect whereby mean stock returns are significantly positive during the first half of the calendar year, and significantly negative during the second half.

18 The number of lags for each model is chosen to optimize AIC.

19 We check for multicollinearity in each of our models and find that none of the VIFs are above 1.98.

Table 5
Disaggregating news according to timing of release.

	(1) R_t	(2) R_t	(3) R_t	(4) $R(DAY)_t$	(5) $R(OUT)_t$
Constant	0.070	0.012	−0.035	0.051	−0.083
	(0.094)	(0.170)	(0.175)	(0.122)	(0.096)
PSENT_JA(DAY)	1.994**		1.262	0.589	0.693
	(0.863)		(0.856)	(0.663)	(0.421)
PSENT_EN(DAY)	0.806**		0.469	0.172	0.296
	(0.392)		(0.389)	(0.286)	(0.212)
log(articles_total(DAY))	−0.023		−0.004	0.011	−0.014
	(0.025)		(0.033)	(0.024)	(0.017)
PSENT_JA(OUT)		1.417***	1.373***	0.457***	0.928***
		(0.234)	(0.239)	0.172	(0.136)
PSENT_EN(OUT)		0.638***	0.598***	0.357***	0.248***
		(0.139)	(0.145)	(0.097)	(0.081)
log(articles_total(OUT))	−0.009	0.001	−0.026	0.023	
		(0.030)	(0.037)	(0.026)	(0.021)
R_{t-1}	0.012	0.008	0.008	−0.037	−0.001
	(0.023)	(0.023)	(0.023)	(0.024)	(0.016)
Day Control	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES
No. Obs.	4313	4313	4313	4313	4313
Adj. R^2	0.001	0.013	0.014	0.003	0.018
F-stat	1.52	7.46	5.92	2.36	7.76

Note: This table reports the estimated coefficients for a regression of the form specified in Eq. (2). The dependent variable is the daily return on the TOPIX index (R_t). The explanatory variables are the probability weighted measures of sentiment derived from Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles, and the natural log of the number of articles published across both languages (*articles_total*). Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. *DAY* indicates that the article was published, or the return occurred, during the Japanese trading day (9am–3pm). *OUT* indicates that the article was published, or the return occurred outside of the Japanese trading day. *R(DAY)* is calculated using the opening and closing prices on day *t*, and *R(OUT)* is calculated using the opening price on day *t* and the prior close. Newey–West standard errors are shown in parentheses.
*** and ** indicate statistical significance at the 1% and 5% level respectively.
Sample period: January 2003–September 2020.

Table 4 shows that both Japanese language and English language sentiment are positively related to daily TOPIX returns at a high degree of significance. This is the case when we combine article counts (column 2) and ignore news that is published on weekends or public holidays (column 3) and the relationship also holds when ASENT is used in place of PSENT in Panel B.[20] The positive sentiment-return relationship is consistent with the prior literature (e.g., Smales, 2015; Heston and Sinha, 2017). The estimated coefficient for the number of articles published is positive but not significantly different from zero. Since article counts are sometimes attributed as a potential measure of investor attention (e.g., Dzielinski, 2011; Khuu et al., 2016; Heston and Sinha, 2017), the baseline results suggest that TOPIX returns are related to sentiment rather than attention.

In Appendix A we show that our baseline results are robust to a variety of different treatments. This includes structuring sentiment indexes to include non-novel news (column 1), excluding neutral items (column 2), weighting news items by relative market capitalization (column 3), incorporating dividends into returns (column 4) and winsorizing data to remove the largest returns (columns 5 and 6). In Appendix C we also show that our baseline results are robust to controlling for the COVID-pandemic in Japan, including the number of cases and number of COVID-related deaths.

3.2. News occurring outside of market hours

In the next stage of our analysis, we start by exploring an issue raised by our earlier observation that most news articles are published outside of market hours, primarily just before the market opens. We disaggregate our sample of news articles into articles occurring during the normal trading day (DAY) and outside market trading hours (OUT). Articles that are published between midnight and the market open (9am) are assigned to the current day *t*, while articles that are published after the market close (3pm) and midnight are assigned to the next day *t+1* since they cannot impact returns when the market is no longer trading.

The estimated coefficients are reported in Table 5. In the first three columns, returns are calculated as in Table 4, from close on day *t-1* to close on day *t*. There is a positive return-relationship for sentiment in both languages occurring during the trading day (column 1) and outside of the trading day (column 2).[21] We see the effect of most of the news occurring before the market opens, when the sentiment measures are combined into one specification (column 3) the news occurring outside of the trading day is statistically more important. We confirm that sentiment generated outside of market hours is most important even when returns are calculated for the day using the market open and close (column 4) or overnight using the prior close and market open (column 5).

3.3. Sensitivity of the results to a structural break

Having established a significant positive relationship between stock market returns and news sentiment, we investigate the stability of our baseline model (shown in Table 4). We conduct a Bai and Perron (1998, 2003) multiple breakpoint test and identify one structural break in our data that occurs on 5th April 2013

20 This is also true when using TSENT. The results are not reported here but are available upon request.

21 While we believe that we do not contradict Khuu et al.'s (2016) that sentiment has a contemporaneous effect on returns, this analysis adds subtlety to that interpretation. The significant relationships of overnight returns affecting returns that occur after the news is released in consistent with Granger causality (see Greene, 2012, p.358).

Table 6
Structural break tests.

Panel A: Bai-Perron Multiple Breakpoint Test			
Break Test	*F*-Statistic	Scaled *F*-Statistic	Critical Value
0 vs. 1	5.44	54.37	27.03
1 vs. 2	2.63	26.28	29.24
Break Date	**05-Apr-13**		
Panel B: Chow Breakpoint Test			
F-Statistic	4.34		
Probability	0.00		
Break Date	**05-Apr-13**		

Note: This table presents the results of the Bai-Perron multiple breakpoint test (Panel A) and the Chow breakpoint test (Panel B) for the baseline model shown in Table 4.
Sample period: January 2003–September 2020.

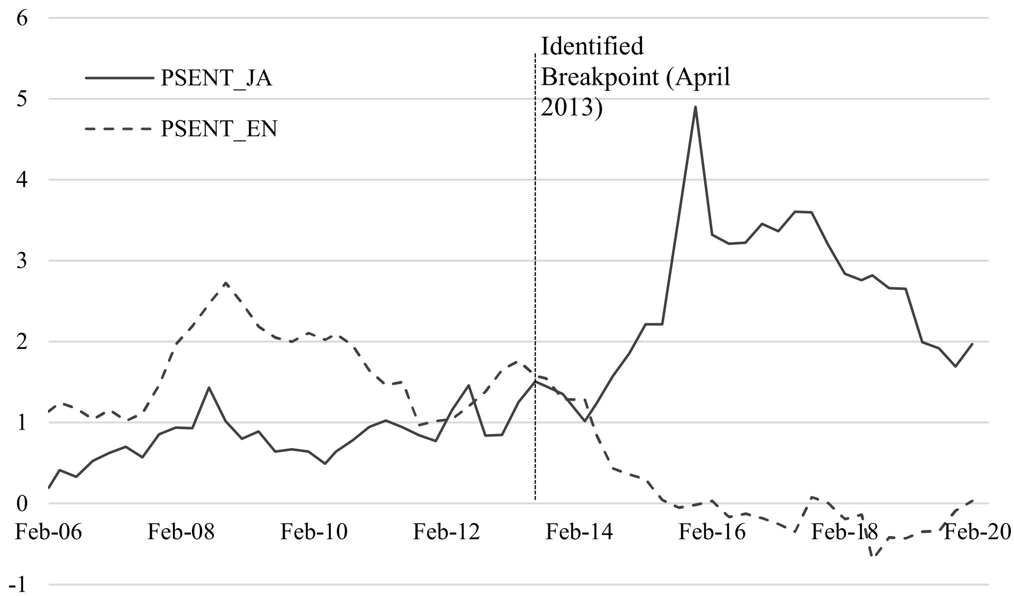

Fig. 2. Rolling estimated coefficients of baseline model.
Note: This figure depicts the estimated coefficients for Japanese language sentiment (*PSENT_JA*) and English language sentiment (*PSENT_EN*) for the baseline regression shown in Table 4. The coefficients are re-estimated for a rolling 3-year (750 trading day) window. The structural breakpoint in April 2013, identified by the Bai Perron test, is also shown.

(Table 6, Panel A). We confirm the presence of a break using a Chow (1960) breakpoint test (Table 6, Panel B) and seek additional clarification by performing a rolling correlation of the baseline model. Fig. 2 plots the estimated sentiment coefficients for fixed 3-year windows and demonstrates that the magnitude of the Japanese language coefficient increased substantially, while the English language coefficient declined towards zero, after April 2013.

The identified structural breakpoint aligns well with the 04 April 2013 announcement of expanded quantitative and qualitative monetary easing (QQE) that was introduced to help meet a new 2% inflation target. QQE involved significant expansion of the monetary base via purchase of Japanese Government Bonds (JGBs), exchange traded funds, and real estate investment trusts (Shirai, 2018). The market response to this surprise announcement was significant enough to elicit discussion at the April 2013 FOMC Meeting (FOMC, 2013).

As QQE is just one of many policies introduced during the Abenomics era, we use the more general term to refer to our breakpoint and divide our analyses using this break point to consider the effect pre- and post-Abenomics. Subsequently, we take the opportunity to revisit our initial analysis and assess whether, and how, the identified relationships for English and Japanese language sentiment change.

Panel A of Table 7 displays descriptive statistics for article publication and sentiment in both languages. The mean number of daily publications is significantly higher in the period post-April 2013 for both languages. Increasing from 155 per day to 233 per day in Japanese, and 111 to 127 per day in English. More significantly, the news sentiment portrayed in those articles differs considerably. While average Japanese language sentiment declines (from 0.069 to 0.039) there is a change of a much greater magnitude in English language sentiment as it increases from −0.056 to +0.126. This provides additional evidence that our results are likely to be driven by differences in sentiment. One possible explanation for this change is that domestic (international) reporters, who write articles in Japanese (English), were more pessimistic (optimistic) about the chance of policy success post-2013.

Fig. 3 shows that the increase in PSENT_EN is driven primarily by news arriving just prior to the opening of the Tokyo market. In addition, Panels B and C disaggregate the sample into periods pre- and post-Abenomics. The correlations are similar across the two panels. Two differences are worth pointing out. First, sentiment in the two languages (PSENT_EN and PSENT_JA) are highly correlated in the pre-Abenomics period (0.336) but not in the post-Abenomics period (0.053). Second, sentiment measures are negatively correlated with article numbers pre-Abenomics but tend to be positively correlated post-Abenomics.

In Table 8 we repeat our earlier test but divide the sample according to the pre-/ post-Abenomics structural break. Columns

Table 7
Descriptive statistics and correlation table: Pre- and Post-Abenomics.

Panel A: Descriptive statistics	Pre-Abe		Post-Abe		diff in mean
Variable	Mean	Std Dev	Mean	Std Dev	(t-test)
No. Articles_JA	155.15	131.79	232.86	201.35	15.31***
PSENT_JA	0.069	0.11	0.039	0.08	−10.17***
ASENT_JA	0.101	0.13	0.053	0.11	−12.33***
No. Articles_EN	111.05	103.82	127.32	141.37	4.36***
PSENT_EN	−0.056	0.13	0.126	0.15	42.25***
ASENT_EN	−0.059	0.18	0.179	0.20	40.62***
Panel B: Correlation Pre-Abenomics	ARTICLES_JA	PSENT_JA	ASENT_JA	ARTICLES_EN	PSENT_EN
PSENT_JA	**−0.162**				
ASENT_JA	**−0.195**	**0.989**			
ARTICLES_EN	**0.917**	**−0.095**	**−0.128**		
PSENT_EN	**−0.303**	**0.336**	**0.345**	**−0.276**	
ASENT_EN	**−0.237**	**0.309**	**0.316**	**−0.217**	**0.983**
Panel C: Correlation Post-Abenomics	ARTICLES_JA	PSENT_JA	ASENT_JA	ARTICLES_EN	PSENT_EN
PSENT_JA	0.003				
ASENT_JA	**−0.071**	**0.979**			
ARTICLES_EN	**0.960**	−0.033	**−0.114**		
PSENT_EN	−0.013	0.052	−0.014	0.029	
ASENT_EN	−0.048	**0.065**	0.013	−0.005	**0.982**

Note: Panel A presents descriptive statistics when the sample period is disaggregated according to the advent of Abenomics. A Bai-Perron multiple breakpoint test indicates a structural break as of 05 April 2013 and this is used to disaggregate our sample. The variables are the probability weighted measures of sentiment derived from Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles, average sentiment measures (*ASENT_JA, ASENT_EN*), and the number of articles published in each languages. Pearson correlation statistics are shown in Panel B and C, with Panel B representing the *pre*-Abenomics sample and Panel C representing the *post*-Abenomics sample. In Panel A statistical significance at the 1% level is indicated by ***. In Panels B and C significance at the 1% level is indicated by **bold** figures.
Sample period: January 2003–September 2020. Pre-Abenomics from January 2003–April 2013 and Post-Abenomics from April 2013–September 2020.

1 and 3 relate to articles published throughout the whole trading (24-hr) day. Columns 2 and 4 disaggregate articles according to whether they are published during the market opening hours (DAY) or outside of those hours (OUT). Considering the aggregated articles first, we again identify a positive relationship between returns and sentiment in both languages. However, whilst that relationship is significant for both languages pre-Abenomics, only the Japanese language sentiment is significant post-Abenomics. This pattern is repeated when disaggregating by time of article. Prior to the structural break, there is a significant relationship for Japanese language sentiment occurring during the market hours, and for English language sentiment occurring outside of market hours. After the structural break, only Japanese language sentiment occurring outside of market hours is important. Following the introduction of Abenomics, Japanese language news clearly takes on a more prominent role in determining stock market returns.

3.4. Asymmetry of negative and positive news effect

The literature suggests that stronger market responses are caused by negative news sentiment (Riordan et al., 2013; Smales, 2014; Uhl, 2014). We test whether this applies in our case by disaggregating the sentiment scores into negative and positive components. We do this by creating two variables for each language, one that takes a value equal to PSENT if PSENT is less than zero and zero otherwise (PSENT_NEG), and another that takes a value equal to PSENT is greater than zero and zero otherwise (PSENT_POS). The model output is shown in Table 9 and we find results that are consistent with the prior literature in non-Japanese markets. In both languages, and for both pre- and post-Abenomics periods, the coefficient for negative news has a larger magnitude than for positive news. Positive English language news is insignificant when considering the whole sample period or just the post-Abenomics period.

3.5. Importance of US news

We have highlighted the importance of news occurring before the Tokyo market opens, and the change that seems to have occurred in English language news during this period, so the determinants of that news sentiment warrant further investigation. Since the spike in news articles occurs after the New York market has just closed, it is possible that events in the US drive sentiment in media coverage of Japanese firms. If that is the case, then it is also possible that Japanese stock returns are related to US stock market sentiment.

We begin by creating an aggregate sentiment score for stocks traded on the NYSE (PSENT_NY) and counting the number of articles published each day for those stocks. We use this to investigate whether US returns and sentiment can explain TOPIX returns, and whether sentiment for Japanese stocks (PSENT_JA or PSENT_EN) is subsumed by this information. In addition to American sentiment, we also control for the returns of the S&P 500 in case this might influence returns. The American market is never open at the same time as Japan's and the subscript *t-1* reflects this. Thus, when considering the Tuesday return in Japan, we must look to Monday's news arrival and associated returns in the US.

We repeat our earlier analysis, incorporating NYSE-based measures into our regression model, and report results in Table 10. Panel A considers articles occurring throughout a 24-hour period and demonstrates that prior session NYSE sentiment and returns are positively related to TOPIX returns. We find further evidence of the importance of dividing the sample according to the (Abenomics-related) structural break point identified earlier. Prior to the structural break (column 1) NYSE sentiment (PSENT_NY) and English language TOPIX sentiment (PSENT_EN) are statistically significant, while the Japanese language TOPIX sentiment (PSENT_JA) is not significant. However, this relationship is reversed following the introduction of Abenomics, and only Japanese language sentiment is important. This again suggests that Japanese language articles better capture important

Table 8
Disaggregating the sample period to show the effect of Abenomics.

	Pre-Abe		Post-Abe	
	R_t	R_t	R_t	R_t
Constant	−0.171	−0.140	−0.144	−0.300
	(0.240)	(0.237)	(0.317)	(0.344)
PSENT_JA	0.568**		2.378***	
	(0.288)		(0.392)	
PSENT_EN	1.606***		0.055	
	(0.224)		(0.210)	
log(articles_total)	0.055		0.018	
	(0.043)		(0.056)	
PSENT_JA(DAY)		2.355**		−0.418
		(1.120)		(1.236)
PSENT_EN(DAY)		0.554		0.399
		(0.631)		(0.465)
log(articles_total(day))		−0.032		0.080
		(0.039)		(0.085)
PSENT_JA(OUT)		0.383		2.629***
		(0.300)		(0.427)
PSENT_EN(OUT)		1.801***		0.086
		(0.241)		(0.246)
log(articles_total(out))		0.076		−0.021
		(0.050)		(0.058)
R_{t-1}	0.002	−0.001	0.014	0.013
	(0.031)	(0.032)	(0.033)	(0.033)
Day Control	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES
No. Obs.	2507	2507	1806	1806
Adj. R^2	0.022	0.023	0.021	0.022
F-stat	7.17	5.85	5.40	4.42

Note: This table presents regression estimates when the sample period is disaggregated according to the advent of Abenomics. A Bai-Perron multiple breakpoint test, indicates a structural break as of the start of April 2013 and this is used to disaggregate our sample. For the regression model, the dependent variable is the daily return on the TOPIX index (R_t). The explanatory variables are the probability weighted measures of sentiment derived from Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles, and the natural log of the number of articles published across both languages (*articles_total*). Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. *DAY* indicates that the article was published, or the return occurred, during the Japanese trading day (9am–3pm). *OUT* indicates that the article was published, or the return occurred outside of the Japanese trading day. Newey–West standard errors are shown in parentheses.
*** and ** indicate statistical significance at the 1% and 5% level respectively.
Sample period: January 2003–September 2020. Pre-Abenomics from January 2003–April 2013 and Post-Abenomics from April 2013–September 2020.

information arising from Abenomics policies in this post-2013 period. The number of articles is not important regardless of whether we consider articles on NYSE stocks or TOPIX stocks.

In Panel B, we once again disaggregate the news for Japanese stocks into news arriving during and outside of the trading day. The results are like those reported in Table 8, whereby news occurring outside of the trading day is more important, with English language news statistically significant pre-Abenomics and Japanese language news significant post-Abenomics. One difference is that here, we report that Japanese language news occurring during the trading day is not statistically significant while in Table 8 we had significance at the 5% level. One explanation could be that the introduction of New York sentiment, which is most relevant at the start of the trading day, has subsumed the effect of news specific to Japanese stocks. This explanation is consistent with Hamao et al. (1990) who note 'evidence of price volatility spillovers from New York to Tokyo', and with Singh et al. (2010) finding that stock markets are mostly affected by markets that close just prior to opening.

3.6. Further tests on quintile portfolios

To this point we have concentrated on aggregate market sentiment. However, in doing this we fail to observe more nuanced return-sentiment relationships for stocks of different sizes. Understanding how the relationship varies by firm size may also provide insights into the mechanism by which sentiment in disparate languages affects stock returns. For instance, international investors constitute a larger proportion of trading in large firms, are more likely to buy stock in large firms and read English language news, so we might expect SENT_EN to be more relevant for big firms. Conversely, those foreign investors are less likely to buy small firms and so English language articles would be less relevant for smaller firms.

We construct five equally weighted quintile portfolios according to firm size, with data acquired from DataStream. The smallest 20% of firms by market capitalization form portfolio Q1, the next 20% of firms form portfolio Q2, etc. We rebalance the portfolio on a semi-annual basis, so the initial portfolios are formed as of 1st January 2003 and rebalanced on 1st July 2003, and so forth. If we are unable to ascertain a market value as of the rebalancing date, then the stock is removed for that period. On average there are 1792 TOPIX constituents each year, and so approximately 358 firms in each quintile portfolio.

Having constructed the quintile portfolios, we compute a time series of daily log returns, along with the number of articles and sentiment associated with firms in each quintile. Descriptive statistics are shown in Table 11. Panel A shows that smaller firms have higher returns than large firms, and this holds for periods

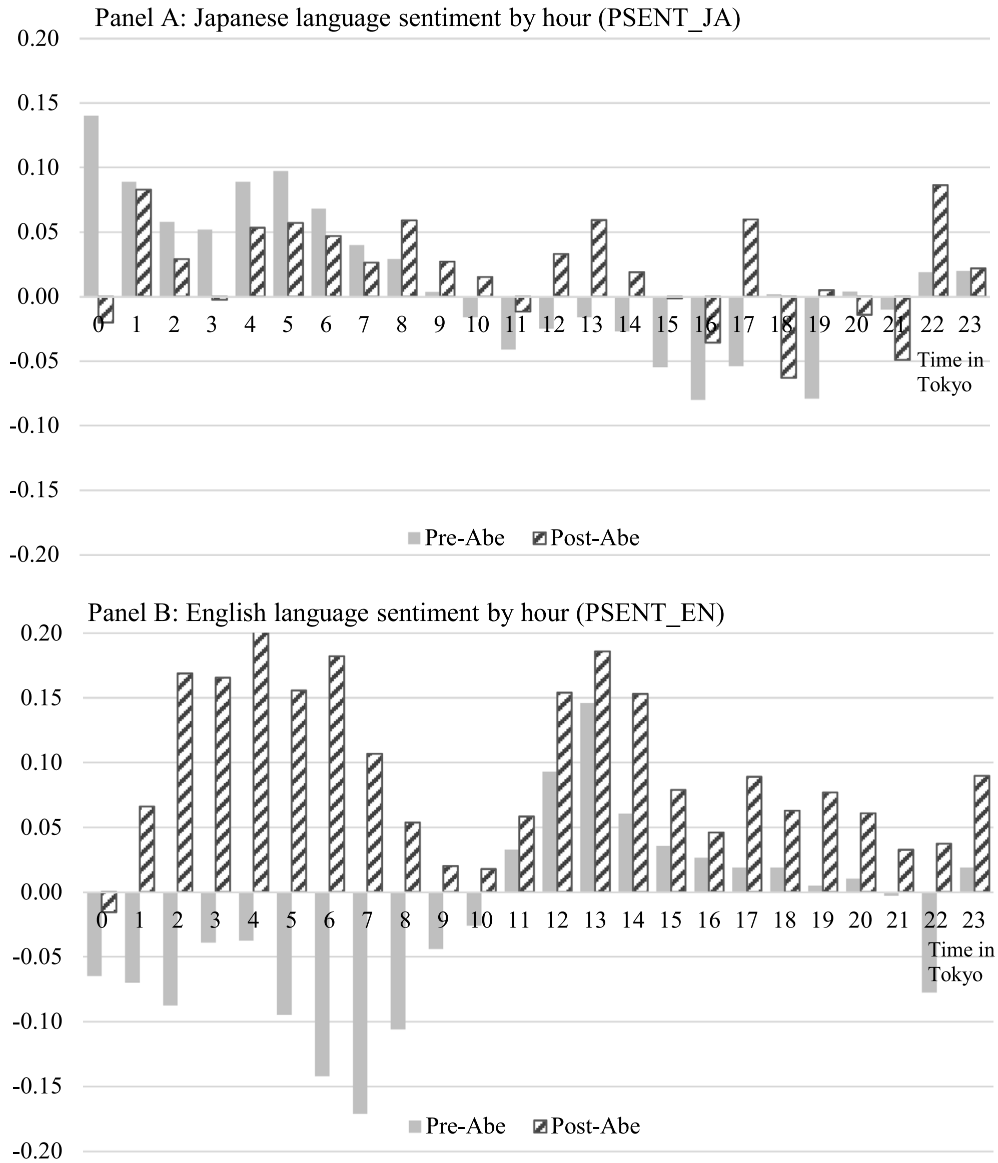

Fig. 3. Hourly sentiment pre- and post-abenomics. *Note:* The horizontal axis represents the time in Tokyo. 0 is midnight, 12 is noon, and 23 is 11pm.

before and after the introduction of Abenomics. Panel B shows that there are more articles, and fewer days without articles, related to large firms. Approximately 55% of all articles concern the largest firms (Q5). Indeed, while there are no days without articles for the largest firms, there are 462 days (10.7% of sample) in which no English language articles are published concerning the smallest firms. The smallest firms (Q1) have the lowest sentiment in both languages and is negative for the English language articles. While the largest firms (Q5) have higher sentiment, the highest sentiment is in the above-average size firms (Q4). Panel C reveals that sentiment has increased even though the number of articles has declined for the smallest firms. The negative English language sentiment for smaller firms is concentrated in the pre-Abenomics period. For all other quintiles, the number of articles has increased in both languages while sentiment has declined in the Japanese articles and increased in the English articles.

Our testing then consists of two components. First, we test whether news sentiment related to stocks in specific quintiles helps to explain returns for the TOPIX index. Since the TOPIX is a value-weighted index we would expect news for the largest firms to provide the greatest explanatory power. Second, we test whether news sentiment related to each quintile portfolio explains returns for that portfolio. In both cases, we also conduct analysis that incorporates the identified "Abenomics" structural break.

Table 12 shows how news from each quintile portfolio contributes to TOPIX returns. TOPIX returns are clearly related to news concerning the largest firms (quintile 5) and are not associated with news for small firms in either language. Separating the sample according to the structural break shows that, consistent with our earlier results, only English language news for the largest firms is important pre-Abenomics, and only Japanese news for the largest firms is significant post-Abenomics.

The first three columns of Table 12 use sentiment computed from articles occurring throughout a whole day. The latter three columns disaggregate sentiment according to whether the associated articles were published during the Tokyo trading day, or outside of the trading day. This allows us to determine whether the estimates for quintile portfolios follow a similar pattern to that established earlier in our analysis (Tables 5 and 8). None of

Table 9
Testing for asymmetry in negative and positive news.

	ALL R_t	Pre-Abe R_t	Post-Abe R_t
Constant	0.025	−0.135	−0.186
	(0.187)	(0.263)	(0.309)
PSENT_JA_NEG	1.948***	0.741	3.913***
	(0.585)	(0.751)	(0.978)
PSENT_JA_POS	1.060***	0.506	1.677***
	(0.302)	(0.381)	(0.589)
PSENT_EN_NEG	1.357***	1.726***	0.446
	(0.285)	(0.362)	(0.683)
PSENT_EN_POS	0.023	1.346**	0.030
	(0.213)	(0.542)	(0.254)
log(articles_total)	0.011	0.054	0.040
	(0.032)	(0.046)	(0.053)
R_{t-1}	0.013	0.008	0.018
	(0.015)	(0.020)	(0.053)
Day Control	YES	YES	YES
D-B Control	YES	YES	YES
No. Obs.	4313	2507	1806
Adj. R^2	0.019	0.021	0.022
F-stat	7.43	5.86	4.75

Note: This table presents regression estimates when news sentiment is disaggregated into positive and negative components. The dependent variable is the daily return on the TOPIX index (R_t). The explanatory variables are the probability weighted measures of sentiment derived from negative and positive Japanese (*PSENT_JA_NEG, PSENT_JA_POS*) and English (*PSENT_EN_NEG, PSENT_EN_POS*) language articles, and the natural log of the number of articles published across both languages (*articles_total*). Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. Newey–West standard errors are shown in parentheses. *** and ** indicate statistical significance at the 1% and 5% level respectively.
Sample period: January 2003–September 2020. Pre-Abenomics from January 2003–April 2013 and Post-Abenomics from April 2013–September 2020.

Table 10
The effect of New York sentiment and returns on Japanese market.

Panel A: Whole Day PSENT	Pre-Abe R_t	Post-Abe R_t	*Panel B: Disaggregated PSENT*	Pre-Abe R_t	Post-Abe R_t
Constant	0.018	−0.393	Constant	0.041	−0.451
	(0.313)	(0.426)		(0.318)	(0.436)
rSP500 $_{t-1}$	0.503***	0.425***	rSP500 $_{t-1}$	0.503***	0.424***
	(0.029)	(0.100)		(0.029)	(0.094)
PSENT_NY$_{t-1}$	0.666***	0.270	PSENT_NY$_{t-1}$	0.672***	0.282
	(0.166)	(0.306)		(0.170)	(0.306)
log(articles_NY)$_{t-1}$	0.020	0.061	log(articles_NY)$_{t-1}$	0.021	0.064
	(0.056)	(0.090)		(0.064)	(0.089)
PSENT_JA	0.075	1.822***	PSENT_JA(DAY)	1.572	−0.586
	(0.248)	(0.369)		(1.068)	(1.171)
PSENT_EN	1.147***	0.009	PSENT_EN(DAY)	0.680	0.562
	(0.189)	(0.192)		(0.508)	(0.442)
log(articles_total)	0.002	0.011	log(articles_total(day))	0.001	0.039
	(0.037)	(0.055)		(0.037)	(0.076)
R_{t-1}	−0.091**	−0.070	*PSENT_JA(OUT)	0.015	2.007***
	(0.036)	(0.036)		(0.261)	(0.404)
			PSENT_EN(OUT)	1.199***	0.036
				(0.218)	(0.228)
			log(articles_total(out))	−0.003	−0.008
				0.043	(0.052)
			R_{t-1}	−0.092**	−0.071
				(0.036)	(0.035)

(*continued on next page*)

the quintile portfolios have news sentiment during the trading day that is significantly related to TOPIX returns and so we do not report those estimated coefficients in Table 12. Instead, we confirm our results are driven by news for the largest firms (quintile 5) occurring outside of the trading day. This news predominantly occurs in the hours prior to the start of trading.

Table 10 (*continued*).

Panel A: Whole Day PSENT	Pre-Abe R_t	Post-Abe R_t	*Panel B: Disaggregated PSENT*	Pre-Abe R_t	Post-Abe R_t
Day Control	YES	YES	Day Control	YES	YES
D-B Control	YES	YES	D-B Control	YES	YES
No. Obs.	2507	1806	No. Obs.	2507	1806
Adj. R^2	0.242	0.156	Adj. R^2	0.242	0.156
F-stat	67.92	28.78	F-stat	54.48	23.27

Note: This table reports the estimated coefficients for a regression of the form specified in Eq. (2) The dependent variable is the daily return on the TOPIX index (R_t). The explanatory variables are the probability weighted measures of sentiment for Japanese firms derived from Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles, the natural log of the number of articles on Japanese firms published across both languages (*articles_total*), similar measures for US stocks traded in New York (*PSENT_NY*, *articles_NY*), and log returns of the S&P500 index (*rSP500*). Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. DAY indicates that the article was published, or the return occurred, during the Japanese trading day (9am–3pm). OUT indicates that the article was published, or the return occurred outside of the Japanese trading day. A Bai-Perron multiple breakpoint test, indicates a structural break as of the start of April 2013 and this is used to disaggregate our sample. Newey–West standard errors are shown in parentheses.
*** indicates statistical significance at the 1% level.
Sample period: January 2003–September 2020.

Table 11
Descriptive statistics for quintile portfolios.

Panel A: Daily quintile returns

Quintile	ALL	Pre-Abe	Post-Abe
1	0.037	0.043	0.028
2	0.026	0.026	0.027
3	0.021	0.022	0.020
4	0.018	0.021	0.014
5	0.014	0.009	0.021

Panel B: News for quintiles in each language

Quintile	Japanese Language			English Language		
	No. Articles	PSENT	No News Days	No. Articles	PSENT	No News Days
1	19.51	0.025	44	13.23	−0.254	462
2	13.11	0.057	71	8.97	−0.126	517
3	15.47	0.055	42	9.72	0.025	414
4	21.51	0.067	9	12.17	0.144	233
5	83.11	0.045	0	56.40	0.094	0

Panel C: News disaggregated by pre- and post- Abenomics

Quintile	Japanese Language				English Language			
	Pre-Abe		Post-Abe		Pre-Abe		Post-Abe	
	No. Articles	PSENT	No. Articles	PSENT	No. Articles	PSENT	No. Articles	PSENT
1	20.69	0.038	17.78	0.073	14.61	−0.281	11.24	0.010
2	12.63	0.083	13.79	0.076	8.97	−0.204	8.97	0.081
3	13.03	0.094	18.99	0.053	8.19	−0.144	11.94	0.211
4	16.68	0.102	28.46	0.051	9.09	−0.082	16.59	0.278
5	68.97	0.077	103.45	0.030	55.86	0.046	57.20	0.162

Note: This table provides summary statistics for equally-weighted quintile portfolios constructed by semi-annual sorting on size. The smallest firms are in quintile 1 and the largest firms are in quintile 5. Panel A shows daily log returns for equally-weighted portfolios. Panel B shows the daily number of articles, the average sentiment, and the number of days with no news for each quintile and in each language. Panel C disaggregates the sample into periods occuring pre-Abenomics introduction (January 2003–April 2013) and post-Abenomics introduction (April 2013–September 2020).
Sample period: January 2003–September 2020.

In Table 13 we focus on quintile portfolio, rather than TOPIX returns. Since we are no longer focusing on the market index, we can also consider the Fama–French factors[22] (FFF) in our regression specification. Panel A displays results for the whole sample, with the FFF incorporated in the latter five columns. Initially, we observe that returns for the smallest firms (Q1 and Q2) are positively related to Japanese language sentiment, perhaps because domestic investors focus on this news and are more likely to invest in smaller firms. Consistent with larger firms having more international investors, their returns are positively related to both Japanese and English language sentiment. However, once we control for the FFF, the statistical significance of the identified relationships largely disappears, at least when considering the whole sample period. However, when we consider the effect of the Fama–French factors in the two sub-periods (Panel B) we find that PSENT_EN is significant in the pre-Abenomics period for the largest stocks in the Japanese market and PSENT_JA is significant for the largest stocks in post-Abenomics period. Therefore, our inferences regarding the two sentiment indices in the sub-periods, are robust to the inclusion of the Fama–French factors.

When we split the sample according to the structural break (Panel B), we note that the smallest quintile returns are only related to Japanese language news in the post-Abenomics period, and even this disappears when we control for the FFF. In contrast, the largest quintile returns relate to news in English pre-Abenomics and to Japanese language news post-Abenomics. Even

[22] Obtained from the website of Prof. Ken French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Table 12
Quintile portfolio news and TOPIX returns.

	All R_t	Pre-Abe R_t	Post-Abe R_t	All R_t	Pre-Abe R_t	Post-Abe R_t
Constant	0.162	0.035	−0.088	0.135	0.143	−0.319
	(0.224)	(0.285)	(0.449)	(0.238)	(0.310)	(0.460)
rSP500 $_{t-1}$	0.466***	0.484***	0.411***	0.463***	0.480***	0.408***
	(0.042)	(0.027)	(0.094)	(0.042)	(0.028)	(0.094)
PSENT_NY$_{t-1}$	0.392***	0.553***	0.197	0.390***	0.555***	0.237
	(0.146)	(0.159)	(0.306)	(0.148)	(0.159)	(0.315)
log(articles_NY)$_{t-1}$	0.003	0.037	0.027	0.004	0.029	0.074
	(0.049)	(0.061)	(0.089)	(0.050)	(0.062)	(0.091)
PSENT_JA_Q1	0.053	0.045	0.093	0.038	0.021	0.080
	(0.058)	(0.081)	(0.083)	(0.056)	(0.081)	(0.079)
PSENT_EN_Q1	−0.095	−0.035	−0.073	−0.079	0.031	−0.111
	(0.054)	(0.077)	(0.086)	(0.057)	(0.079)	(0.086)
PSENT_JA_Q2	0.046	−0.025	0.152	0.035	−0.031	0.134
	(0.065)	(0.085)	(0.102)	(0.062)	(0.084)	(0.094)
PSENT_EN_Q2	0.004	0.001	0.063	−0.049	−0.094	0.058
	(0.053)	(0.077)	(0.079)	(0.057)	(0.080)	(0.081)
PSENT_JA_Q3	−0.090	−0.129	−0.048	−0.104	−0.130	−0.061
	(0.071)	(0.082)	(0.127)	(0.070)	(0.079)	(0.124)
PSENT_EN_Q3	0.134	0.155	0.120	0.084	0.108	0.143
	(0.077)	(0.081)	(0.078)	(0.054)	(0.079)	(0.077)
PSENT_JA_Q4	0.112	0.007	0.235	0.073	−0.022	0.205
	(0.089)	(0.112)	(0.148)	(0.084)	(0.106)	(0.142)
PSENT_EN_Q4	0.016	0.163**	−0.099	−0.009	0.099	−0.084
	(0.058)	(0.081)	(0.082)	(0.056)	(0.079)	(0.081)
PSENT_JA_Q5	0.604***	0.188	1.157***	0.696***	0.268	1.194***
	(0.160)	(0.178)	(0.355)	(0.153)	(0.176)	(0.305)
PSENT_EN_Q5	0.394***	0.890***	−0.086	0.379***	0.852***	−0.130
	(0.122)	(0.176)	(0.181)	(0.119)	(0.178)	(0.164)
log(articles_all)	−0.047	−0.036	−0.018	−0.066	−0.079**	−0.035
	(0.027)	(0.031)	(0.057)	(0.032)	(0.039)	(0.052)
Day Control	YES	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES	YES
Time of Day Sentiment Disaggregation	NO	NO	NO	YES	YES	YES
No. Obs.	4313	2507	1806	4313	2507	1806
Adj. R^2	0.200	0.237	0.149	0.201	0.238	0.150
F-stat	78.18	56.68	23.63	37.09	27.03	11.63

Note: This table reports the estimated coefficients for a regression of the form specified in Eq. (2) The dependent variable is the daily return on the TOPIX index (R_t). The explanatory variables are the probability weighted measures of sentiment for firms in size quintile k derived from Japanese (*PSENT_JA_k*) and English (*PSENT_EN_k*) language articles, and the natural log of the number of articles published across both languages (*articles_total*). Similar measures for US stocks traded in New York (*PSENT_NY*, *articles_NY*), and log returns of the S&P500 index (*rSP500*) are also included. Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. The sample is disaggregated into periods occuring pre-Abenomics introduction and post-Abenomics introduction. In the first three columns the sentiment measures are derived with articles published throughout a whole day. In the last three columns the sentiment measures are disaggregated according to whether the associated articles were published during the Tokyo trading day (*DAY*) or outside of the trading period (*OUT*). Since none of the *DAY* coefficients are statistically significant we only report the *OUT* coefficients. Newey–West standard errors are shown in parentheses.
*** and ** indicates statistical significance at the 1% and 5% level.
Sample period: January 2003–September 2020.

after controlling for FFF, we see that returns for large firms are related to English language news pre-Abenomics and only Japanese news post-Abenomics. This echoes our earlier results and confirms the importance of Japanese language news (and lower importance of English language news) following the introduction of Abenomics.

3.7. News disagreement

Our analysis has shown that over the period 2003–2020 there has been a significant difference in average news sentiment relating to Japanese and English language articles. Further, we have shown that the relationship between returns and news published in the disparate languages differs over time, with English taking precedence in early years, and Japanese post-Abenomics. One question we have not fully addressed is whether these observations matter? That is, is there any market impact when there is a difference in sentiment captured from each language. One interpretation of the difference in sentiment is that it represents a difference in opinion among journalists writing in each language; a view consistent with our proposition that articles in Japanese are presenting a very different worldview to those written in English (and *vice versa*). If that is the case, then the models of Harris and Raviv (1993) and Kandel and Pearson (1995) would suggest that there should be a positive relationship between the magnitude of the opinion difference and both trading volume and volatility. We use the absolute value of daily returns to proxy for volatility and compute a measure of abnormal trading volume as the difference between JPY trading volume (i.e., turnover) on day t and average trading volume over the preceding 60-days. In Table 14 we find evidence in support of this model, at least for the largest stocks and the aggregate TOPIX index. This implies that the difference in news sentiment *does* matter.

Table 13
Relationship of quintile portfolio news with quintile portfolio returns.

Panel A: Whole Sample	Small				Large	Small				Large
	1	2	3	4	5	1	2	3	4	
Constant	0.573**	0.341**	0.332	0.235	0.209	0.339**	0.156	0.107	0.112	0.128
	(0.253)	(0.230)	(0.226)	(0.223)	(0.225)	(0.144)	(0.123)	(0.121)	(0.122)	(0.123)
rSP500 $_{t-1}$	0.350***	0.359***	0.372***	0.393***	0.427***	0.026	0.005	−0.005	−0.007	−0.003
	(0.031)	(0.036)	(0.039)	(0.041)	(0.041)	(0.020)	(0.022)	(0.023)	(0.024)	(0.022)
PSENT_NY$_{t-1}$	0.313**	0.353***	0.365***	0.378***	0.304**	0.063	0.049	0.049	0.030	0.007
	(0.134)	(0.136)	(0.139)	(0.142)	(0.145)	(0.073)	(0.073)	(0.079)	(0.079)	(0.074)
log(articles_NY)$_{t-1}$	−0.123**	−0.069	−0.051	−0.036	−0.002	−0.084***	−0.045	−0.025	−0.031	−0.030
	(0.053)	(0.050)	(0.051)	(0.050)	(0.050)	(0.031)	(0.027)	(0.027)	(0.028)	(0.027)
PSENT_JA	0.151***	0.126**	−0.059	0.221**	0.522***	0.043	0.088***	−0.057	0.032	0.083
	(0.057)	(0.061)	(0.073)	(0.091)	(0.152)	(0.034)	(0.033)	(0.039)	(0.048)	(0.090)
PSENT_EN	−0.045	0.002	0.042	0.085	0.440***	0.004	0.047	0.059	0.037	0.109
	(0.052)	(0.046)	(0.050)	(0.056)	(0.117)	(0.033)	(0.027)	(0.030)	(0.027)	(0.066)
log(articles_all)	−0.041**	−0.035	−0.045**	−0.041	−0.060	−0.014	−0.001	−0.007	−0.001	−0.00
	(0.018)	(0.019)	(0.019)	(0.022)	(0.034)	(0.011)	(0.011)	(0.012)	(0.013)	(0.020)
Day Control	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
FFF Control	NO	NO	NO	NO	NO	YES	YES	YES	YES	YES
No. Obs.	4313	4313	4313	4313	4313	4313	4313	4313	4313	4313
Adj. R^2	0.126	0.128	0.130	0.138	0.175	0.711	0.738	0.729	0.730	0.760
F-stat	57.67	58.59	59.48	63.58	83.91	758.65	867.78	831.29	834.18	978.6

Panel B: Disaggregated Sample	Pre-Abe		Post-Abe		Pre-Abe		Post-Abe	
	Smallest (Low 20%)	Largest (Top 20%)	Smallest (Low 20%)	Largest (Top 20%)	Smallest (Low 20%)	Largest (Top 20%)	Smallest (Low 20%)	Largest (Top 20%)
Constant	0.616	0.055	0.252	0.259	0.429**	0.230	0.030	−0.044
	(0.333)	(0.293)	(0.390)	(0.415)	(0.200)	(0.158)	(0.189)	(0.246)
rSP500 $_{t-1}$	0.339***	0.433***	0.367***	0.404***	0.040	0.011	0.002	−0.034
	(0.026)	(0.028)	(0.077)	(0.097)	(0.025)	(0.022)	(0.030)	(0.042)
PSENT_NY$_{t-1}$	0.359**	0.416***	0.316	0.212	0.024	−0.008	0.128	0.092
	(0.148)	(0.152)	(0.270)	(0.315)	(0.084)	(0.085)	(0.129)	(0.135)
log(articles_NY)$_{t-1}$	−0.104	0.018	−0.084	0.031	−0.104**	−0.036	−0.022	−0.018
	(0.071)	(0.061)	(0.083)	(0.092)	(0.045)	(0.036)	(0.038)	(0.041)
PSENT_JA	0.114	0.103	0.229***	1.298***	0.067	−0.022	0.040	0.272**
	(0.078)	(0.178)	(0.088)	(0.325)	(0.047)	(0.106)	(0.047)	(0.138)
PSENT_EN	−0.023	0.903***	−0.023	−0.063	0.007	0.219**	−0.002	−0.093
	(0.082)	(0.175)	(0.075)	(0.186)	(0.058)	(0.093)	(0.042)	(0.100)
log(articles_all)	−0.082***	−0.044	0.015	−0.084	−0.021	−0.023	0.002	0.028
	(0.026)	(0.038)	(0.024)	(0.068)	(0.016)	(0.022)	(0.014)	(0.045)
Day Control	YES	YES	YES	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES	YES	YES	YES
FFF Control	NO	NO	NO	NO	YES	YES	YES	YES
No. Obs.	2507	2507	1806	1806	2507	2507	1806	1806
Adj. R^2	0.128	0.202	0.135	0.142	0.691	0.755	0.757	0.779
F-stat	34.35	58.54	26.52	28.27	400.36	553.86	402.02	455.75

Note: This table reports the estimated coefficients for a regression of the form specified in Eq. (2) Equally-weighted quintile portfolios are constructed by semi-annual sorting on size, with the smallest 20% of firms in quintile 1 and the largest 20% of firms in quintile 5. The dependent variable is the daily return on quintile portfolio *k*. The explanatory variables are the probability weighted measures of sentiment for firms in the same quintile portfolio *k* derived from Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles, and the natural log of the number of articles published across both languages (*articles_total*). Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. The rightmost columns also control for the three Fama–French factors. Panel A includes the whole sample period and Panel B disaggregates the sample into periods occuring pre-Abenomics introduction (pre-April 2013) and post-Abenomics introduction (post-April 2013). Newey–West standard errors are shown in parentheses. *** and ** indicate statistical significance at the 1% and 5% level respectively.
Sample period: January 2003–September 2020.

4. Concluding remarks

The important role of sentiment in markets has been established in a large, and still growing, body of academic literature. Sentiment is of interest not just to academics, but to the wider investment community. For example, sentiment is particularly important for the growing number of high-frequency traders who rely on the real-time generation of trading signals from news analytics software to trade on company-specific news (Von Beschwitz et al., 2020).

We take advantage of recently available sentiment data from Thomson Reuters News Analytics (TRNA) which uses only Japanese language data and create a sentiment index for the Japanese market. We then use the same method to create a sentiment index using only TRNA data from English language sources. Many studies have, to date, only had access to similar English language sources to create sentiment indices for market located

Table 14
Disagreement between news in disparate languages.

	|rTOPIX|	|rQ1|	|rQ2|	|rQ3|	|rQ4|	|rQ5|	Abn. Volume
Constant	0.245	0.736***	0.575***	0.420**	0.390	0.272	0.011
	(0.226)	(0.248)	(0.214)	(0.208)	(0.203)	(0.218)	(0.628)
$\|rSP500_{t-1}\|$	0.369***	0.289***	0.302***	0.305***	0.320***	0.344***	0.278***
	(0.032)	(0.026)	(0.028)	(0.028)	(0.030)	(0.031)	(0.062)
$\|PSENT\_NY_{t-1}\|$	0.690***	0.236	0.318	0.401**	0.510***	0.627***	0.087
	(0.180)	(0.166)	(0.173)	(0.177)	(0.192)	(0.187)	(0.577)
$log(articles\_NY)_{t-1}$	−0.052	−0.070	−0.022	0.003	0.004	−0.043	−0.295**
	(0.040)	0.049	(0.046)	(0.045)	(0.045)	(0.046)	(0.141)
|PSENT_JA - PSENT_EN|	0.310**	−0.009	−0.038	0.055	0.014	0.316**	0.227
	(0.133)	(0.052)	(0.052)	(0.046)	(0.052)	(0.140)	(0.432)
log(articles_all)	0.102***	0.014	0.009	0.024	0.041	0.096**	0.422***
	(0.032)	(0.017)	(0.017)	(0.020)	(0.022)	(0.041)	(0.088)
Day Control	YES	YES	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES	YES	YES
No. Obs.	4314	4314	4314	4314	4314	4314	4254
Adj. R^2	0.152	0.101	0.110	0.109	0.115	0.136	0.208
F-stat	78.47	49.83	54.06	53.57	57.00	68.95	102.46

Note: This table reports the estimated coefficients for a regression where the dependent variable is the absolute value of the daily return on the TOPIX index (*rTOPIX*) or a size-sorted portfolio (*rQk*), or abnormal TOPIX trading volume (*Abn. Volume*). Abnormal trading volume is measured as the difference between JPY turnover on day t and average turnover during the preceding 60-days, scaled by 10^8. The explanatory variables are the absolute value of the *difference* in the probability weighted measures of sentiment for firms in size quintile k derived from Japanese and English language articles (*|PSENT_JA-PSENT_EN|*), and the natural log of the number of articles published across both languages (*articles_total*). Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. Newey–West standard errors are shown in parentheses.
*** and ** indicate statistical significance at the 1% and 5% level respectively.
Sample period: January 2003–September 2020.

in non-English speaking markets (see, for example, Calomiris and Mamaysky, 2019).

While the positive sentiment-return relationship we find is consistent with the existing literature, we also discover that we do not fully understand the relationship of Japanese returns and sentiment since much is lost in translation if we only use an English language source to derive sentiment. In Japan, we see that it is the English sentiment index which drives returns *before* a key date in Abenomics (namely, the Bank of Japan's announcement of expanded quantitative and qualitative monetary easing). After this, we find that it is only the Japanese sentiment index that is associated with Japanese returns; the English language sentiment index's role disappears. Furthermore, we find that this sentiment index is associated with the returns of Japan's largest firms. The timing of the articles we use to derive both the English and Japanese sentiment indices is also important: the sentiment of news arriving before the market open seems more influential than that derived using news appearing during the trading day.

While we only study Japan, we believe that our analysis provides a cautionary lesson to researchers and investors using only English language sources to model market sentiment. Researchers and traders need to ensure that their algorithms consider information presented in all relevant languages.

CRediT authorship contribution statement

Robert B. Durand: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Joyce Khuu:** Conceptualization, Data curation. **Lee A. Smales:** Conceptualization, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing.

Data statement

The data that supports the findings of this study has been obtained from three different sources: (a) sentiment data from Japanese language and English language news articles is obtained from Thomson Reuters News Analytics (TRNA) now renamed Refinitiv News Analytics (RNA); (b) data for TOPIX returns together with firm-level returns and characteristics is obtained from DataStream; and (c) Fama–French factors are obtained from the website of Prof. Ken French. The first two sets of data are available only on license from Refinitiv, the third may be accessed at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Acknowledgments

We are grateful for comments from seminar participants at the 2022 West Coast Finance Colloquium, the 2022 New Zealand Finance Meeting, and the 2023 AFAANZ Conference. We also thank Refinitiv for making the data used in this paper available to us.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Appendix A

See Table A.1.

Appendix B

See Fig. B.1.

Appendix C

See Table C.1.

Table A.1
Robustness of baseline model.

Panel A: Whole Sample	(1)	(2)	(3)	(4)	(5)	(6)
	R_t	R_t	R_t	TR_t	R_t	R_t
Constant	0.027	0.094	−0.007	−0.091	0.019	0.031
	(0.177)	(0.175)	(0.183)	(0.181)	(0.140)	(0.145)
PSENT_JA	0.751***	0.157***	0.131***	0.973***	1.147***	1.243***
	(0.198)	(0.061)	(0.048)	(0.167)	(0.173)	(0.192)
PSENT_EN	0.759***	0.149***	0.095***	0.447***	0.476***	0.497***
	(0.128)	(0.055)	(0.029)	(0.093)	(0.102)	(0.109)
log(articles_total)	−0.003	−0.014	−0.004	0.009	−0.006	−0.011
	(0.031)	(0.031)	(0.032)	(0.031)	0.024	(0.025)
R_{t-1}	0.011	0.011	0.012	0.011	0.021	0.012
	(0.023)	(0.023)	(0.023)	(0.023)	(0.015)	(0.015)
Day Control	YES	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES	YES
No. Obs.	4313	4313	4313	4313	4313	4313
Adj. R^2	0.013	0.002	0.004	0.013	0.017	0.016
F-stat	7.53	2.94	3.11	7.10	9.14	8.60
Panel B: Pre-Abe	(1)	(2)	(3)	(4)	(5)	(6)
	R_t	R_t	R_t	TR_t	R_t	R_t
Constant	−0.132	0.198	0.016	−0.165	−0.043	−0.020
	(0.228)	(0.228)	(0.248)	(0.239)	(0.183)	(0.195)
PSENT_JA	0.384	0.012	0.061	0.555*	0.440**	0.591**
	(0.253)	(0.073)	(0.067)	(0.287)	(0.212)	(0.245)
PSENT_EN	2.331***	0.346***	0.089***	1.580***	1.348***	1.376***
	(0.209)	(0.079)	(0.024)	(0.224)	(0.189)	(0.202)
log(articles_total)	0.076	−0.023	−0.017	0.054	0.032	0.027
	(0.041)	(0.040)	(0.042)	(0.043)	(0.034)	(0.037)
R_{t-1}	0.001	0.004	0.045	0.005	0.018	0.009
	(0.031)	(0.032)	(0.032)	(0.032)	(0.020)	(0.020)
Day Control	YES	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES	YES
No. Obs.	2507	2507	2507	2507	2507	2507
Adj. R^2	0.044	0.006	0.009	0.021	0.028	0.027
F-stat	13.84	2.76	3.44	6.96	8.90	8.51
Panel C: Post-Abe	(1)	(2)	(3)	(4)	(5)	(6)
R_t	R_t	R_t	TR_t	R_t	R_t	
Constant	0.006	−0.448	−0.205	−0.121	−0.137	−0.132
	(0.315)	(0.313)	(0.329)	(0.317)	(0.233)	(0.239)
PSENT_JA	2.005***	0.537***	0.216***	2.373***	2.039***	2.042***
	(0.365)	(0.107)	(0.080)	(0.389)	(0.299)	(0.322)
PSENT_EN	0.169	0.062	0.045	0.054	0.016	0.045
	(0.188)	(0.078)	(0.095)	(0.211)	(0.162)	(0.165)
log(articles_total)	0.004	0.077	0.040	0.054	0.021	0.018
	(0.055)	(0.055)	(0.058)	(0.211)	(0.040)	(0.041)
R_{t-1}	0.017	0.014	0.021	0.017	0.019	0.011
	(0.033)	(0.034)	(0.034)	0.033	(0.025)	(0.026)
Day Control	YES	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES	YES
No. Obs.	1806	1806	1806	1806	1806	1806
Adj. R^2	0.017	0.011	0.001	0.022	0.027	0.023
F-stat	4.57	3.23	2.80	5.41	6.48	5.62

Note: This table reports the estimated coefficients for a regression of the form specified in Eq. (2) In columns 1-3 the dependent variable is the daily total return on the TOPIX price index (R_t). In column 4, the dependent variable is the daily *total return* (including dividends) for the TOPIX index (TR_t). In columns 5 and 6 the dependent variable is the daily total return on the TOPIX *price* index (R_t) winsorized at the 5%/95% level and the 1%/99% levels respectively. The explanatory variables are the probability weighted measures of sentiment derived from Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles, and the natural log of the number of articles published in both languages (*articles_ja + articles_en*). In column 1 the sentiment indexes include non-novel news, in column 2 neutral news items are excluded, and in column 3 the sentiment indexes are constructed by weighting news items by relative market capitalization. Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. News occuring on the weekend or public holiday is incorporated into the next trading day. Panel A reports results for the whole sample period. Panel B reports results for the period prior to Abenomics, and Panel C reports results for the period after the start of Abenomics (April 2013). Newey–West standard errors are shown in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.
Sample period: January 2003–September 2020.

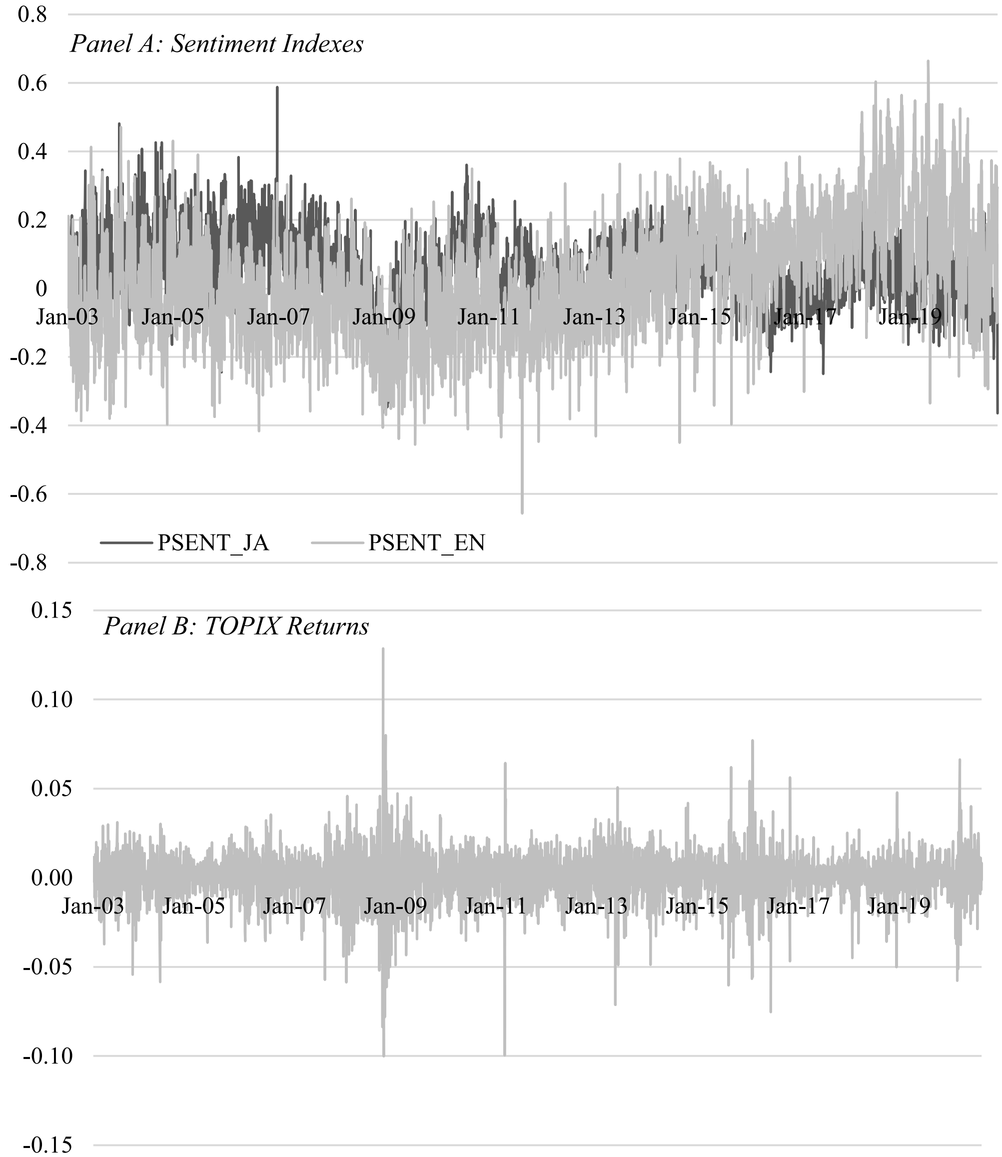

Fig. B.1. Time series of sentiment and returns.
Note: Panel A presents the time series of daily sentiment indexes based on Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles. Panel B shows the time series of daily log returns for the TOPIX index.

Table C.1
Controlling for the effect of COVID-pandemic.

	R_t	R_t	R_t	R_t	R_t
Constant	−1.131	−0.059	−0.057	−0.058	−0.054
	(1.600)	(0.187)	(0.187)	(0.186)	(0.186)
PSENT_JA	4.305***	1.385***	1.332***	1.396***	1.388***
	(1.844)	(0.219)	(0.221)	(0.219)	(0.218)
PSENT_EN	0.414***	0.574***	0.576***	0.571***	0.572***
	(0.882)	(0.126)	(0.128)	(0.125)	(0.126)
log(articles_total)	0.207	0.003	0.003	0.002	0.002
	(0.299)	(0.032)	(0.032)	(0.032)	(0.032)
COVID		0.003	0.027		
		(0.113)	(0.129)		
log(1+COVID_CASES)			0.015		
				(0.023)	
log(1+COVID_DEATHS)					0.031
					(0.068)

(*continued on next page*)

Table C.1 (*continued*).

	R_t	R_t	R_t	R_t	R_t
COVID*PSENT_JA			2.979* (1.619)		
COVID_PSENT_EN			0.187 (0.690)		
R_{t-1}	0.164 (0.082)	0.008 (0.015)	0.008 (0.015)	0.008 (0.015)	0.008 (0.015)
Day Control	YES	YES	YES	YES	YES
D-B Control	YES	YES	YES	YES	YES
No. Obs.	153	4313	4313	4313	4313
Adj. R^2	0.043	0.014	0.014	0.014	0.014
F-stat	1.77	6.99	6.11	7.00	6.98

Note: This table reports the estimated coefficients for a regression of the form specified in Eq. (2) The dependent variable is the daily return on the TOPIX index (R_t). The explanatory variables are the probability weighted measures of sentiment derived from Japanese (*PSENT_JA*) and English (*PSENT_EN*) language articles, and the natural log of the number of articles published in both languages (*articles_total*). Dummy variables are used to control for day of week and Dekansho-bushi calendar effects. The COVID pandemic is controlled for using a dummy variable indicating the presence of *COVID* cases in Japan (from 16 January onwards), the number of new daily cases (*COVID_CASES*) and deaths (*COVID_DEATHS*) reported by the World Health Organization. Newey–West standard errors are shown in parentheses.
***, **, and * indicates statistical significance at the 1%, 5%, and 10% level respectively.
Sample period: January 2003–September 2020. COVID period: January 2020–September 2020.

References

Allen, D.E., McAleer, M., Singh, A.K., 2019. Daily market news sentiment and stock prices. Appl. Econ. 51, 3212–3235.
Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66, 47–78.
Bai, J., Perron, P., 2003. Computation and analysis of multiple structural change models. J. Appl. Econometrics 18, 1–22.
Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. J. Finance 61, 1645–1680.
Boudoukh, J., Feldman, R., Richardson, M., 2013. Which News Moves Stock Prices? A Textual Analysis. NBER Working Paper 18725.
Brown, G.W., Cliff, M.T., 2005. Investor sentiment and asset valuation. J. Bus. 78, 405–440.
Calomiris, C., Mamaysky, H., 2019. How news and its context drive risk and returns around the world. J. Financ. Econ. 133, 299–336.
Cathcart, L., Gotthelf, N.M., Shi, Y., 2020. News sentiment and sovereign credit risk. Eur. J. Financ. Manag. 26, 261–287.
Chan, F., Durand, R.B., Khuu, J., Smales, L.A., 2017. The validity of investor sentiment proxies. Int. Rev. Financ. 17, 473–477.
Chiah, M., Hu, X., Zhong, A., 2022. Photo sentiment and stock returns around the world. Finance Res. Lett. 46 (Part B), 102417.
Chow, G.C., 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica 28, 591–605.
Clements, A.E., Todorova, N., 2016. Information flow, trading activity and commodity futures volatility. J. Futures Mark. 36, 88–104.
Da, Z., Engelberg, J., Gao, P., 2015. The sum of all FEARS: Investor sentiment and asset prices. Rev. Financ. Stud. 28, 1–32.
Dang, T.L., Moshirian, F., Zhang, B., 2015. Commonality in news around the world. J. Financ. Econ. 116, 82–110.
Du, Z., Huang, A.G., Wu, W., 2022. Language and domain specificity: A Chinese financial sentiment dictionary. Rev. Financ. 26, 673–719.
Dzielinski, M., 2011. News Sensitivity and the Cross-Section of Stock Returns. NCCR Finrisk Working Paper No. 719.
Gotthelf, N., Uhl, M.W., 2019. News sentiment: A new yield curve factor. J. Behav. Financ. 20, 31–41.
Greene, W.H., 2012. Econometric Analysis, seventh ed. Pearson Education, Edinburgh.
Groß-Klußmann, A., Hautsch, N., 2011. When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. J. Empir. Financ. 18, 321–340.
Gu, C., Kurov, A., 2020. Information role of social media: Evidence from Twitter sentiment. J. Bank. Financ. 121, 105969.
Gupta, K., Banerjee, R., 2019. Does OPEC news sentiment influence stock returns of energy firms in the United States? Energy Econ. 77, 34–45.
Hamao, Y., Masulis, R.W., Ng, V., 1990. Correlations in price changes and volatility across international stock markets. Rev. Financ. Stud. 3, 281–307.
Harris, M., Raviv, A., 1993. Differences of opinion make a horse race. Rev. Financ. Stud. 6, 473–506.
Hendershott, T., Livdan, D., Schürhoff, N., 2015. Are institutions informed about news? J. Financ. Econ. 117, 249–287.
Heston, S.L., Sinha, N.R., 2017. News vs. sentiment: Predicting stock returns from news stories. Financ. Anal. J. 73, 67–83.
Hirshleifer, D., Shumway, T., 2003. Good day sunshine: Stock returns and the weather. J. Finance 58, 1009–1032.
Iwatsubo, K., Watkins, C., 2021. The changing role of foreign investors in Tokyo stock price formation. Pac.-Basin Finance J. 67, 101548.
Kamstra, M.J., Kramer, L.A., Levi, M.D., 2003. Winter blues: A SAD stock market cycle. Amer. Econ. Rev. 93, 324–343.
Kandel, E., Pearson, N.D., 1995. Differential interpretation of public signals and trade in speculative markets. J. Polit. Econ. 103, 831–872.
Kearney, C., Liu, S., 2014. Textual sentiment in finance: A survey of methods and models. Int. Rev. Financ. Anal. 33, 171–185.
Khuu, J., Durand, R.B., Smales, L.A., 2016. Melancholia and Japanese stock returns – 2003 to 2012. Pac.-Basin Finance J. 40, 424–437.
Lee, W.Y., Jiang, C.X., Indro, D.C., 2002. Stock market volatility, excess returns, and the role of investor sentiment. J. Bank. Financ. 26, 2277–2299.
Leinweber, D., Sisk, J., 2011. Relating news analytics to stock returns. In: The Handbook of News Analytics in Finance. Wiley & Sons, Hoboken NJ (Chapter 6).
Lemmon, M., Portniaguina, E., 2006. Consumer confidence and asset prices: Some empirical evidence. Rev. Financ. Stud. 19, 1499–1529.
Loughran, T., McDonald, B., 2016. Textual analysis in accounting and finance: A survey. J. Account. Res. 54, 1187–1230.
Loughran, T., McDonald, B., 2020. Textual analysis in finance. Annu. Rev. Finan. Econ. 12, 357–375.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 59, 817–858.
Obaid, K., Pukthuanthong, K., 2022. A picture is worth a thousand words: measuring investor sentiment by combining machine learning and photos from news. J. Financ. Econ. 144, 273–297.
Riordan, R., Storkenmaier, A., Zhang, S.S., 2013. Public information arrival: Price discovery and liquidity in electronic limit order markets. J. Bank. Financ. 37, 1148–1159.
Sakakibara, S., Yamasaki, T., Okada, K., 2013. The calendar structure of the Japanese stock market: The 'sell in May effect' versus the 'Dekansho-bushi effect'. Int. Rev. Financ. 13, 161–185.
Shirai, S., 2018. The Bank of Japan's Super-Easy Monetary Policy from 2013–2018. Asian Development Bank Institute Working Paper Series, No. 896.
Singh, P., Kumar, B., Pandey, A., 2010. Price and volatility spillovers across North American, European and Asian stock markets. Int. Rev. Financ. Anal. 19, 55–64.
Smales, L.A., 2014. News sentiment in the gold futures market. J. Bank. Financ. 49, 275–286.
Smales, L.A., 2015. Time-variation in the impact of news sentiment. Int. Rev. Financ. Anal. 37, 40–50.

Smales, L.A., 2016. News sentiment and bank credit risk. J. Empir. Financ. 38, 37–61.
Tetlock, P.C., 2007. Giving content to investor sentiment: the role of media in the stock market. J. Finance 62, 1139–1168.
Uhl, M.W., 2014. Reuters sentiment and stock returns. J. Behav. Financ. 15, 287–298.
Von Beschwitz, B., Keim, D.B., Massa, M., 2020. First to read the news: News analytics and algorithmic trading. Rev. Asset Pricing Stud. 10, 122–178.
Wang, C., 2003. Investor sentiment, market timing, and futures returns. Appl. Financial Econ. 13, 891–898.

Internet Resource

Anon, 2013. Transcript of the Federal Open Market Committee meeting on April 30–May 1 2013. https://www.federalreserve.gov/monetarypolicy/files/FOMC20130501meeting.pdf.