JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 59, No. 8, Dec. 2024, pp. 3564–3600

doi:10.1017/S002210902300131X

News and Markets in the Time of COVID-19

Harry Mamaysky
Finance Division, Columbia Business School
hm2646@columbia.edu

Abstract

The onset of COVID-19 was characterized by voluminous, negative news. Higher narrativity news topics (measured by textual proximity to articles describing the 1987 stock market crash and textual distance from Federal Reserve communications) were systematically associated with contemporaneous market responses, which were larger on high volatility days (hypersensitivity), and with markets–news feedback. Hypersensitive news topic-market pairs were associated with next-day reversals. A test using the news–markets relationship identifies a mid-March 2020 structural break, which was knowable by the end of April. Post break, markets and news became considerably less coupled, and hypersensitivity and reversals abated.

I. Introduction

COVID-19 gained global attention starting in January of 2020, and was declared a pandemic by the World Health Organization on Mar. 11, 2020. In the early months of the pandemic, financial markets experienced almost unprecedented volatility accompanied by large price declines in risky assets. The S&P 500 fell 34% from its February peak to its March trough. Ten-year yields fell precipitously, the VIX index reached levels not seen since the Global Financial Crisis of 2008–2009, and credit markets experienced large sell-offs. Such severe market reactions had never before occurred in response to infectious disease outbreaks (Baker, Bloom, Davis, Kost, Sammon, and Viratyosin (2020)). Media coverage of the events of the day was dire. Headlines warned of "unexploded bombs" of corporate debt, of a stalling economy, and of an unprecedented recession.[1] Shiller (2019) argues that many financial and economic outcomes are shaped by narratives, and

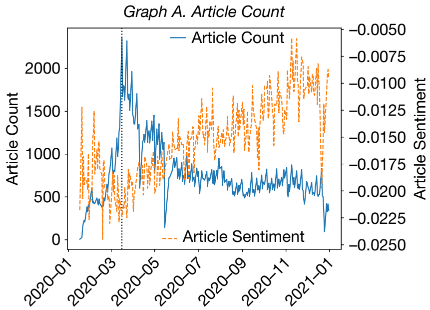
I thank Charles Calomiris, Jennifer Conrad (the editor), Paul Glasserman, Markus Schwedeler (the referee), and Ivo Welch for their helpful suggestions, as well as seminar participants from Columbia, the 2021 European Winter Meeting of the Econometric Society, the Kansas City Fed, Maryland, the Office of the Comptroller of the Currency, QWAFAxNEW, UBS, and the 2020 Wolfe Research NLP and ML Conference. Jiashu Sun and Xinran Zhang provided excellent research assistance. An earlier version of this article appeared under the title "Financial Markets and News About the Coronavirus."

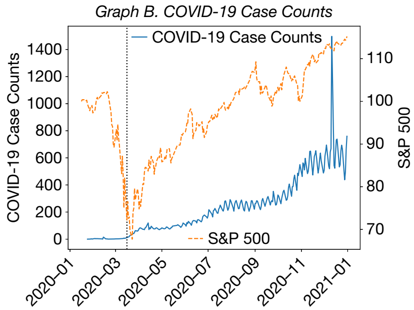
[1]"Coronavirus May Light Fuse on 'Unexploded Bomb' of Corporate Debt." *New York Times*, March 11, 2020; "Markets Plunge. Economies Stall. Panic Spreads. It All Feels Very 2008." *New York Times*, March 13, 2020; "With Unprecedented Force and Speed, a Global Recession is Likely Taking Hold." *Washington Post*, March 15, 2020.

FIGURE 1

More specifically, News Coverage and COVID-19 Case Counts

Graph A of Figure1 shows the daily count of Reuters articles mentioning "coronavirus" or "COVID-19" on the left axis, and the average sentiment (see Section III) of those articles on the right axis. Graph B shows the daily incidence of COVID-19, in thousands of cases, as reported by the Johns Hopkins Coronavirus Resource Center (see Section II) on the left axis, and the S&P 500 index scaled to start at 100 on the right axis. The vertical dotted line in both graphs shows the day of the peak in article counts.





by early-2020, a narrative had emerged that the pandemic would, in short order, lead to economic disaster.

Graph A of Figure 1 shows the typical progression of a narrative epidemic documented by Shiller (2019), with the number of articles about coronavirus peaking by late March and then slowly falling as 2020 drew to a close.[2] Graph A also shows that as the volume of articles discussing the pandemic approached its peak, the sentiment expressed in those articles became progressively more negative. The narrative epidemic was characterized by voluminous and negative press coverage. Graph B shows that the S&P 500 index bottomed just after the peak of the COVID-19 narrative epidemic. By August, the stock market was setting new all-time highs, coronavirus article counts were dropping, and the sentiment expressed in those articles was becoming less negative. In contrast, COVID-19 disease incidence, as measured by case counts obtained from the Johns Hopkins Coronavirus Resource Center, continued to rise steadily throughout all of 2020 as shown in Graph B.[3] The narrative and disease epidemics thus followed very different trajectories. Which of the epidemics had more influence on markets? Based on the anecdotal evidence of Figure 1, the answer is clear.

The economic disaster narrative that emerged in the early days of the COVID-19 pandemic had a profound impact on investors. Understanding market behavior thus requires understanding the evolution of this narrative. Goetzmann, Kim, and Shiller ((GKS) (2022)) write that "[n]arratives are a social medium through which an idea can spread rapidly through conversation and potentially affect aggregate beliefs about asset prices." GKS show that when the media carries more crash-narrative stories, investors become more concerned about market

[2]Shiller ((2019), Chapter 8) emphasizes that narrative epidemics can occur over very different time frames: 1 year in this case. The large dip in articles in May occurs because the Reuters archive contains no entries from May 15 to May 20, 2020, likely due to an error in the archive construction.

[3]The spike in this series on Dec. 11, 2020 (see dating discussion in Section II) appears in the Johns Hopkins data, and plays little role in the article's analysis.

crashes. In addition to its high volume and negative sentiment, news flow at the height of the COVID-19 market panic was also highly narrative in nature.[4]

Highly narrative news likely attract investor attention via the availability heuristic of Tversky and Kahneman (1974), in which people respond to events based on their salience, not their objective probability. Kuran and Sunstein ((1999), KS) argue that "in a broad array of contexts that call for risk judgments, individuals lack reliable or first-hand knowledge. In such contexts, they assess probabilities with the help of the availability heuristic." In the early days of the pandemic, there was certainly a lack of first-hand knowledge. Most information about pandemic's impacts came from media coverage. KS introduce the concept of an *availability cascade* to describe a social mechanism "through which expressed perceptions trigger chains of individual responses that make these perceptions appear increasingly plausible through their rising availability in public discourse." According to this logic, if more narrative news is more salient, it will elicit more individual responses (investor positioning), which renders highly narrative news yet more plausible, resulting in further media coverage and market responses. I refer to the idea that highly narrative news articles attract investor attention and thereby trigger cascade-like effects (bad news coverage leads to negative market responses which lead to more negative news coverage) as the *narrativity hypothesis.*

In this article, I use the Reuters news archive to investigate the narrativity hypothesis. The time period of my analysis covers all of 2020, prior to which media focus on COVID-19 was virtually nonexistent. By late 2020, widespread inoculations using multiple effective vaccines had begun, stock markets globally had recovered from their early-2020 lows, and investor attention turned to market bubbles, meme stocks, and cryptocurrencies. The end of 2020 thus serves as a neat bookend for any plausible time frame of the pandemic-induced market crisis.

My news data set consists of the 189,548 Reuters articles that mention "coronavirus" or "COVID-19" in 2020. Reuters is a particularly useful news source because it is comprehensive, well-followed, and nonpartisan.[5] I classify coronavirus news stories into 12 topics using a modified latent Dirichlet allocation (LDA) approach. The choice of 12 topics optimally balances the desire for more topics, to better characterize news flow, with the preference that topics be *coherent* (i.e., that topics contain words that sensibly belong together, as in Newman, Han Lau, Grieser, and Baldwin (2010)). News flow about the pandemic exhibits great variation both on a given day and across time. From early discussions of the health and market impact of the disease, the narrative shifts to a discussion of the impact of the pandemic on corporations (2 topics), of its effect on credit markets (2 topics), and of the policy responses by central banks and governments. The other topics focus on currencies, European economies, oil and commodities, and sports.

Following Calomiris and Mamaysky (2019), I construct a measure of daily topical sentiment by interacting topical frequency with sentiment, measured using the Loughran and McDonald (2011) sentiment dictionary. In addition to topical sentiment, I track aggregate sentiment, the cross-sectional standard deviation of

[4]Section III.B for a formal definition of narrativity.

[5]According to AllSide's media bias ratings, the Reuters news service is in the political center: https://www.allsides.com/media-bias/media-bias-ratings.

sentiment, and the article count series from Figure 1. I analyze how these text measures are related to the returns of 5 asset classes: the S&P 500 stock index, the VIX volatility index, the FTSE US high-yield corporate bond index, and 2- and 10-year Treasury yields. The behavior of these 5 asset classes is representative of the breadth of market responses to the pandemic.

I find that daily asset class returns respond strongly to contemporaneous news flow, and each asset class responds to distinct aspects of the news. Lagged market returns Granger cause much of next-day news flow, even when controlling for lagged news flow. Lagged news Granger cause next-day market activity, frequently reversing the contemporaneous impact of news on markets. Following Glasserman, Mamaysky, and Shen (2024), who suggest that price changes are more sensitive to information during high volatility periods, I show that news have a larger impact on markets when the lagged VIX is high. I refer to this as *hypersensitivity*, and show hypersensitivity is associated with overreaction to news.

Using the test in Andrews (1993), (2003), I show the contemporaneous relationship between news and markets undergoes a structural break in the middle of Mar. 2020. Along the dimensions discussed in the prior paragraph, news and markets become considerably less coupled post break, with less evidence of hypersensitivity and reversals. The mid-March structural break could have been identified using only data through the end of Apr. 2020. Repeating the break test in expanding windows for the rest of 2020 shows the timing of the break remains unchanged. The ability to identify, in real time, when periods of tight coupling between news and markets end should be of great practical interest.

A. Economic Narratives and Testable Hypotheses

Goetzmann et al. (2022) propose that narrativity can be measured by the similarity of news articles to a corpus with high narrative content. GKS argue that the Oct. 19, 1987 stock market collapse came to dominate media crash narratives. To establish a narrativity measure for my 12 news topics, I measure the textual distance between news topics and *Wall Street Journal* articles from Oct. 20 to Oct. 23, 1987 (as in GKS). I refer to news topics with a lower textual distance (higher similarity) to the 1987 crash corpus as having high narrativity. I then measure the textual distance between news topics and Federal Reserve (Fed) policy statements, which represent factual, as opposed to narrative, descriptions of underlying economic conditions. I refer to topics with a low distance (high similarity) to the Fed corpus as having low narrativity content. As a matter of definition, therefore, topics that are more similar to the Oct. 20–23, 1987 *Wall Street Journal* articles will be labeled high narrativity, and topics that are more similar to Fed policy statements will be labeled low narrativity.

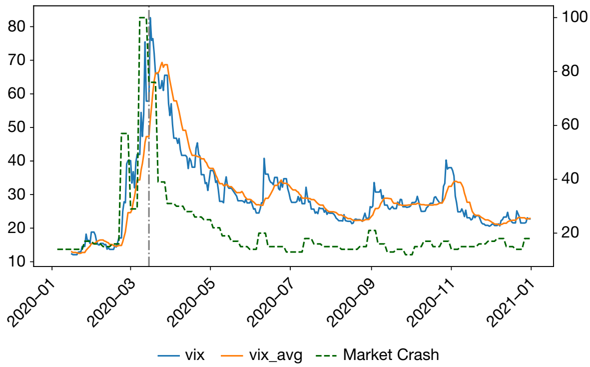
A key tenet of the narrativity hypothesis, which builds on the Kuran and Sunstein (1999) availability cascade idea, is that investor responses to news flow are not entirely rational because more highly narrative news coverage evokes larger investor responses.[6] A direct implication of larger investor responses is that markets

[6]It is implausible that a purely rational evaluation of the cash flow prospects of stocks could have undergone a sufficiently rapid decline and recovery to justify the price path of the S&P 500 shown in Figure 1. Gormsen and Koijen (2020) argue that price variation of longer-dated dividend strips was

FIGURE 2

VIX and Investor Attention

Figure 2 shows the VIX and the rolling 10 trading-day average level of the VIX (left axis). The dashed green line shows the Google search volume for the term "market crash" (right axis), obtained from trends.google.com. The dashed vertical line shows the Mar. 15, 2020, regime break explained in Section V. The correlation between the VIX (10-day rolling VIX) and the "market crash" search volume measure is 86.8% (83.6%) in the pre-Mar. 15, 2020 part of the sample.



are more likely to respond to highly narrative news series than to more factual (less narrative) ones:

Prediction 1. High narrativity topics should be systematically associated with more frequent contemporaneous market responses than factual (low narrativity) topics.

GKS argue that investor attention to crash narratives can be captured using Google search volumes for terms like "stock market crash" or "market crash." Figure 2 shows the VIX and the Google search volume for the term "market crash" during 2020. Both spike early in the sample, remain elevated into April, and then fall later on in the year. The two series also share much high frequency variation. Therefore, another interpretation of the VIX interaction in my hypersensitivity regressions is as a measure of investor attention. If narrativity operates through investor attention, then during times of high investor attention, high narrativity topics should result in larger market responses.

Prediction 2. The level of the VIX (as a proxy for investor attention) should impact market responses to high narrativity topics but not to low narrativity topics.

Starting with Jegadeesh (1990) and Lehmann (1990), an extensive literature has shown that stock returns often experience short-term reversals. In the context of news, Chan (2003), Tetlock (2010), (2011), and Griffin, Hirschey, and Kelly (2011)

---

responsible for both the market sell-off and subsequent rebound, and was likely caused by changes in anticipated discount rates because rational expectations about far-off dividends cannot have experienced such drastic changes. Though the narrativity hypothesis does not require rationality, it is consistent with at least a partially rational market sell-off if investor risk aversion responded to highly narrative, negative news coverage. I return to this point in Section VI.

https://doi.org/10.1017/S002210902300131X Published online by Cambridge University Press

have shown that stock price moves that occur on news days tend to be more persistent than stock price moves that occur on no-news or on stale-news days. If hypersensitivity (i.e., larger market responses to news) reflects increased investor attention rather than higher information content, then hypersensitivity should be associated with return reversals, as contemporaneous overreactions are corrected the next day.

Prediction 3. Hypersensitivity should be associated with return reversals.

If higher narrativity content attracts higher readership, the media will be incentivized to use more narrative language to describe past market events, a primary focus of financial news coverage. Through the availability heuristic, highly narrative news should attract more investor attention and thus have a larger impact on future market prices.

Prediction 4. High narrativity news series will be more highly connected in the news–markets Granger causality network.

The alternative to the narrativity hypothesis is the *rational information hypothesis*, which posits that COVID-19 news articles carried factual information to which markets reacted rationally. Under this hypothesis, news narrativity, in and of itself, should not impact the markets–news relationship. Since investors would already follow all relevant news flow, attention proxies, like the VIX, should not lead to differential market responses to news. Furthermore, the rational information hypothesis predicts that lead–lag relationships between news and markets should not be impacted by narrativity.

Prediction 5. Predictions 1–4 do not hold under the rational information hypothesis.

I test these predictions separately in the pre- and post-break periods. The tests, detailed in Section VI, examine whether a given market m and news topic k exhibit a contemporaneous or lead–lag relationship, hypersensitivity, or overreaction, depending on topic k's narrativity. My results, which rely on a cross section of asset class returns, suggest that the impact of narrativity is conditional. It is highest in the pre-break subsample, but the rational information hypothesis better describes the post-break data.

B. Relationship to the Literature

There is a large literature on the impact of COVID-19 on firm-level and macroeconomic outcomes (Arteaga-Garavito, Croce, Farroni, and Wolfskeil (2021) and Hassan, Hollander, van Lent, Schwedeler, and Tahoun (2021) provide good overviews). The present article focuses on how text data can be used to better understand the market impact of the COVID-19 crisis. Text data is a particularly promising tool for the real-time monitoring of crises because it is both timely and able to reflect unexpected events which traditional data series may not capture.

Several articles use data from earnings calls or SEC filings. Using the Loughran and McDonald (2011) sentiment dictionary applied to earnings call transcripts, Ramelli and Wagner (2020) show that COVID-19 had a more negative return impact on internationally exposed and highly leveraged, cash-poor firms. Li, Liu, Mai, and Zhang (2021) use topic analysis applied to COVID-19 related paragraphs of earnings calls to show that firms with stronger corporate cultures outperformed those with weaker cultures when faced with similar COVID-19 exposure. Davis, Hansen, and Seminaro-Amez (2020) use firms' 10-Ks published prior to the onset of COVID-19 to show that firms' risk exposures impact how their returns are affected by COVID-19 news. Using topic analysis of earnings call discussion of COVID-19 and other historical disease outbreaks, Hassan et al. (2021) show firms are equally concerned about negative supply and demand issues during COVID-19, whereas in past disease outbreaks firms focused more on demand shortfalls.

Other work uses news or social media data. Baker et al. (2020) show that COVID-19 news explained a large portion of the February through March stock return volatility, in contrast to past infectious disease outbreaks. Arteaga-Garavito et al. (2021) use data on medical announcements and Twitter posts from major newspapers as high-frequency measures of COVID-19 news flow, and show that COVID-19 contagion risk is priced.

Most of the extant articles simply assume a subjective end date to the COVID-19 market crisis. By applying natural language processing and econometric techniques to markets and news data, I formally show that the crisis regime ended in mid-March of 2022. Karavias, Narayan, and Westerlund (2023) use a panel of 61 countries' stock index returns with weekly market and COVID-19-related (though not news-based) variables to show that a structural break took place in their specification in the first week of April. My structural break is roughly 2 weeks earlier, perhaps because shifts in the daily markets–news relationship can be detected more quickly. My finding of a stark difference in the news–markets regimes pre- and post-break, as well as the work of Karavias et al. (2023), emphasize the importance of formal break detection for analysis of crisis episodes.

While much of the COVID-19 text-based literature looks at earnings calls or SEC filings, I analyze an extensive collection of news articles, which are available at a higher frequency than earnings calls and are thus more timely. The article's focus on feedback between economic outcomes and text data in the context of COVID-19 is unique. Other articles do not explore how text measures are impacted by past economic outcomes, and focus only on correlations between text and contemporaneous or future firm-level variables.[7] But news stories, tweets, SEC filings, and earnings calls are themselves impacted by prior macroeconomic and firm-specific events. They, in turn, impact the decisions of companies and investors, and thus future economic outcomes. Such feedback linkages are first-order determinants of markets and news behavior during the COVID-19 crisis.

This article differs from Goetzmann et al. (2022) in important ways. GKS focus on the interaction of aggregate measures of news narrativity with aggregate U.S. stock market returns and implied volatility over several decades. They show that higher

[7]Arteaga-Garavito et al. (2021) analyze how social media activity responds to medical announcements, but do not study feedback between social media activity and markets.

narrativity forecasts higher investor crash attention (measured via Google searches and investor surveys) and higher implied volatility. My analysis differs by focusing on the cross section of narrativity across news topics and on its differential impact across asset classes. By focusing only on COVID-19, I show that the impact of narrativity is greater during peak crisis times, and subsides as markets and news normalize. I also document feedback between news coverage and markets, and show that this feedback is more pronounced for more narrative news series.

To the best of my knowledge, this article is the first to:[8] systematically categorize COVID-19 news flow using machine learning methods; analyze a cross section of the major asset classes and how these respond to and impact news innovations; show break tests applied to the news–markets relationship quickly identify regime shifts; show that news hypersensitivity is a characteristic of stressed, but not of typical, markets; show evidence of feedback between news and markets; and connect all these effects to the narrative content of news. The present article introduces several methodologies (such as the modified LDA coherence approach and structural break tests applied to news–markets pairs) which should prove useful in the analysis of future episodes of market stress. The data series used from this paper are available as Supplementary Material to this article or at sites.google.com/view/hmamaysky.

The rest of the article proceeds as follows: Section II describes the data. Section III discusses the text analytics and the news-based series. Section IV discusses features of the news-returns relationship in the pre-break part of the sample. Section V details the structural break tests and analyzes the markets–news relationship post break. Section VI discusses tests of the narrativity and rational information hypotheses. Section VII presents multiple robustness checks. Section VIII concludes and suggests areas for future work.

II. Data

The articles used in this study are obtained from the Thomson Reuters news archive, and include all English-language articles from 2019 to the end of Dec. 2020 that mention the (case insensitive) terms "coronavirus" or "COVID-19."[9] The first 2020 article mentioning coronavirus is on Jan. 8, and coverage begins in earnest on Jan. 17. My analysis starts on this date. Articles released on day t count as day t articles if their timestamps are prior or equal to 4PM EST; post-4PM articles count as day $t+1$ articles. Post-4PM Friday and weekend articles are classified for the subsequent Monday.[10] Articles occurring after 4PM EST on Dec. 31, 2020 are excluded from the analysis. Reuters identifies a collection of related articles (the initial one and revisions) via the Primary News Access Code (PNAC). When there are multiple articles in a PNAC chain, the last one is selected. The final corpus contains 189,548 articles.

I use two additional corpora to measure the narrativity content of the Reuters COVID-19 news coverage. The 1987 crash corpus consists of 164 *Wall Street*

[8]The first version of this article appeared in Apr. 2020. Despite many revisions, the main conclusions (structural break timing, hypersensitivity, lesser markets–news feedback post break) have not changed.

[9]The only matching article in 2019 (Feb. 27) describes the Coalition for Epidemic Preparedness Innovations presciently partnering with a German firm to speed up anti-pandemic vaccine production.

[10]Figure A1 in the Supplementary Material shows the distribution of hours of article publication.

Journal articles from Oct. 20 to Oct. 23, 1987, obtained from Factiva. The Fed corpus consists of 155 Federal Open Market Committee statements from Feb. 4, 1994 to May 4, 2022.

I collect daily price data on the S&P 500 index, the VIX index, the FTSE US High-Yield Market index (HY), which tracks the performance of high-yield corporate bonds, and US 2- and 10-year Treasury yields, labeled, respectively, GT2 and GT10. These 5 asset classes capture a broad range of market responses to COVID-19 news flow. The data are from Bloomberg, and run from Jan. through the end of Dec. 2020.

I obtain data on global confirmed COVID-19 cases from the Johns Hopkins Coronavirus Resource Center. These data are updated daily, and start on Jan. 22, 2020. According to the GitHub page that stores the data (time_series_covid19_confirmed_global.csv from github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series), the update frequency is: “Once a day around 23:59 (UTC).” For example, the data for Apr. 1, 2020 would be released at 23:59 UTC, which is 7:59PM EST, and is after the market close for Treasuries, the S&P 500, VIX, and high-yield bond indexes.[11] For these markets, the reaction to the Apr. 1, 2020 COVID-19 case counts would not happen until the next trading day’s close, or Apr. 2, 2020. For this reason, I use the day t increase in the global COVID-19 case count as the day $t+1$ value of my *corona* series. This aligns the case counts with the days on which markets would have reacted. For Mondays, the 1-day increase comes from Sunday case counts. Assigning to Mondays, the cumulative case count increase from Friday, Saturday, and Sunday would introduce a day-of-week effect, which using only the Sunday increase avoids.

Figure 1 shows the *corona* series, reported in thousands of cases. The daily COVID-19 case counts start relatively low, and experience exponential growth starting in mid-March. Table 1 reports summary statistics for the markets and *corona* series. For the table, the HY and S&P 500 indexes are normalized to a value of 100 on Jan. 16, 2020, the day prior to the start of the analysis. The *corona* series, though it represents the daily increase in case counts, is highly persistent with an AR(1) coefficient of around 0.91. The markets series’ AR(1) coefficients measure the autocorrelation of either returns (labeled with r) or differences (labeled with d). The AR(1) coefficient for daily S&P 500 returns, for example, is −0.35, suggesting a very large degree of mean-reversion. For context, the AR(1) coefficient for daily S&P 500 returns in 2019 was −0.09. The VIX index is also strongly mean-reverting over this sample, while HY index returns are strongly positively auto-correlated. Interestingly, Treasury yield changes have relatively low autocorrelations.

III. Text Analysis

Article j’s sentiment is given by

$$\text{SENT}_j = \frac{\text{POS}_j - \text{NEG}_j}{\text{TOTAL}_j}, \tag{1}$$

[11] According to the *FTSE Fixed Income Index Guide*, indexes have pre-6PM closes.

https://doi.org/10.1017/S002210902300131X Published online by Cambridge University Press

TABLE 1

Summary Statistics for Daily Variables

Table 1 displays summary statistics for the variables used in the analysis. All data are daily. The labels SENT, SENT_SD, and COUNT refer, respectively, to the daily mean and standard deviation of article sentiment, as well as the daily number of articles. The F_[TOPIC] and S_[TOPIC] series refer to topical frequency and sentiment, respectively. The coronavirus case counts, labeled CORONA, are obtained from the Johns Hopkins Coronavirus Resource Center, and are reported in thousands. No. of Obs. shows the numbers of days for which there are observations. The AR(1) (first-order autocorrelation) coefficient is shown by default for the level of each series, but if marked with *r* (*d*) it is shown for 1-day returns (differences). The markets series are labeled as follows: SP500 is the S&P 500 stock index; VIX is the VIX volatility index; GT2 and GT10 are the 2- and 10-year Treasury yields, respectively; and HY refers to the FTSE US high-yield index. The series VIX_AVG refers to the average level of VIX over the prior 10 trading days.

Variable	No. of Obs.	Mean	Std. Dev.	5%	25%	Median	75%	95%	AR(1)
SENT	247	−0.015	0.004	−0.021	−0.018	−0.015	−0.012	−0.009	0.790
SENT_SD	247	0.022	0.001	0.020	0.021	0.022	0.023	0.024	0.517
ART_COUNT	247	767.154	361.641	375.800	555.500	679.000	832.500	1,472.200	0.837
F_SPORTS	247	0.063	0.026	0.020	0.048	0.062	0.075	0.112	0.533
F_CENTRAL_BANK	247	0.092	0.027	0.034	0.080	0.098	0.111	0.126	0.833
F_MARKETS	247	0.074	0.024	0.048	0.060	0.070	0.080	0.120	0.812
F_HEALTH	247	0.195	0.071	0.142	0.159	0.174	0.214	0.288	0.547
F_EUROPE	247	0.063	0.011	0.050	0.058	0.063	0.068	0.080	0.346
F_OIL_&_COMM	247	0.048	0.012	0.035	0.042	0.048	0.054	0.065	0.683
F_CURRENCY	247	0.076	0.027	0.040	0.059	0.073	0.088	0.130	0.715
F_CREDIT	247	0.023	0.011	0.004	0.017	0.025	0.031	0.039	0.702
F_CORP_&_GOVT_US	247	0.117	0.033	0.057	0.102	0.118	0.135	0.172	0.780
F_CORP_ACTUAL	247	0.103	0.053	0.041	0.063	0.089	0.135	0.211	0.830
F_CORP_FUTURE	247	0.104	0.031	0.060	0.084	0.101	0.119	0.162	0.686
F_CREDIT1	247	0.041	0.020	0.007	0.027	0.043	0.054	0.073	0.720
S_SPORTS	247	−0.001	0.000	−0.002	−0.001	−0.001	−0.001	−0.000	0.542
S_CENTRAL_BANK	247	−0.001	0.001	−0.002	−0.002	−0.001	−0.001	−0.001	0.833
S_MARKETS	247	−0.001	0.001	−0.002	−0.001	−0.001	−0.001	−0.001	0.813
S_HEALTH	247	−0.003	0.002	−0.006	−0.003	−0.003	−0.002	−0.001	0.648
S_EUROPE	247	−0.001	0.000	−0.002	−0.001	−0.001	−0.001	−0.001	0.675
S_OIL_&_COMM	247	−0.001	0.000	−0.001	−0.001	−0.001	−0.001	−0.000	0.789
S_CURRENCY	247	−0.001	0.001	−0.002	−0.001	−0.001	−0.001	−0.001	0.726
S_CREDIT	247	−0.000	0.000	−0.001	−0.000	−0.000	−0.000	−0.000	0.708
S_CORP_&_GOVT_US	247	−0.002	0.000	−0.002	−0.002	−0.002	−0.001	−0.001	0.576
S_CORP_ACTUAL	247	−0.002	0.001	−0.004	−0.002	−0.001	−0.001	−0.000	0.881
S_CORP_FUTURE	247	−0.002	0.001	−0.003	−0.002	−0.001	−0.001	−0.001	0.701
S_CREDIT1	247	−0.001	0.000	−0.001	−0.001	−0.001	−0.000	−0.000	0.626
CORONA	242	240.953	217.541	1.887	76.688	212.188	318.997	648.339	0.906
SP500	240	97.996	10.270	76.909	92.172	100.011	104.703	113.183	r −0.350
VIX	240	29.929	12.202	15.036	22.860	26.920	33.433	57.117	d −0.335
HY	247	97.160	5.678	85.509	94.882	99.286	100.352	104.607	r 0.311
GT2	248	0.346	0.420	0.121	0.145	0.161	0.226	1.438	d 0.074
GT10	248	0.852	0.310	0.581	0.662	0.728	0.908	1.606	d −0.071
VIX_AVG	248	29.586	11.633	14.677	22.840	27.183	32.698	57.342	d 0.846
FACTOR	247	0.003	0.353	−0.606	−0.056	0.062	0.177	0.341	0.455
FACTOR_ALL	247	−0.000	0.297	−0.656	−0.098	0.065	0.186	0.329	0.560

where POS_j and NEG_j are the number of positive and negative Loughran and McDonald (2011) words appearing in the article, and TOTAL_j is the total number of words after excluding stop words.[12] I employ the Das and Chen (2007) algorithm to mark negated words,[13] and negated sentiment words are ignored in the above counts. Daily aggregate sentiment SENT_t is the equal-weighted average SENT_j of all articles classified as day *t* articles. This series is shown in Figure 1. The average value of daily sentiment in the corpus is −0.015 (Table 1). Glasserman, Li, and Mamaysky (2024) show that, in a corpus of approximately 1.4 million articles about S&P 500 firms from 1996 to 2018, the average article sentiment using

[12]The measure in (1) is standard in the literature, but improving sentiment measurement is an active research area (Garcia, Hu, and Rohrer (2023), Ke, Kelly, and Xiu (2022)). Future work can explore whether these more sophisticated tools yield greater insights into the COVID-19 market crisis.

[13]Implemented via the `mark_negation` function of Python's `NLTK` package.

equation (1) is −0.011. The sentiment in the present corpus is lower, though not dramatically so. I also calculate the daily standard deviation of article-level sentiment; this series, labeled *sent_sd*, proxies for the dispersion of daily news flow.

I assign topics to articles by running LDA. To run LDA, I first stem all words in the corpus, drop stop words and several other commonly occurring words,[14] and retain all words occurring more than 10 times in any month's set of coronavirus articles. I then construct a document-term matrix $\mathcal{D}$, the jth row of which contains the count of all 21,010 retained words in article j. LDA works by randomly assigning a topic allocation to the nth word in the jth document in proportion to the probability of observing that topic in the document and the probability of observing that word in the topic. The latter two probabilities are determined from the topic allocations of all the other words in the corpus. The process is repeated until the vector of topic allocations over all words settles into its steady state distribution. LDA was introduced in Blei, Ng, and Jordan (2003) and is nicely summarized in Steyvers and Griffiths (2007).

The LDA model is estimated using only the 72,263 Thomson Reuters articles that mention "coronavirus" or "COVID-19" through the end of Apr. 2020, so D is of dimension 72,263 × 21,010. This April end date is chosen to ensure clean structural break tests (so the break is not driven by a changing topic model) as explained in Section A3.3 of the Supplementary Material. Section A3.4 of the Supplementary Material shows the article's results are unchanged if the topic model is estimated using the full sample. The output of the LDA process is a collection of 12 topics, each of which is a probability distribution over the retained words. The model is then used to calculate the topic loadings (probabilities over topics) for all 189,548 articles in the corpus. I refer to the jth document's loading on the kth topic as $f_{j,k}$.

Ke, Montiel Olea, and Nesbit (2021) show that different LDA model can lead to the same log likelihood for a given corpus. Because of this, I estimate 80 different topic models, across eight choices of number of topics (3, 6, …, 24) with 10 runs (each starting at a random seed) for each topic number. I then select the topic model (the 12-topic model used in the article) with a desirable *coherence* score across the 80 runs. Coherence, introduced by Newman et al. (2010), measures the extent to which the top words in the model's topics tend to co-occur in the corpus. Section A1 of the Supplementary Material details the procedure. Importantly, the 12-topic model was selected *prior* to running any of the analysis in the article. While the main conclusions of the article are robust to the choice of topic model (see Section A3.4 of the Supplementary Material), a model with more coherent topics increases the interpretability of the results. The takeaway is that topic models should be chosen with care. The usual approach of running a single LDA with a guess as to the appropriate number of topics is unlikely to yield the best-possible model.

Figure 3 shows word clouds for two topics (*central bank* and *markets*), with each word's size drawn in proportion to its probability weight. Topic labels reflect my qualitative judgment about the topic theme. For example, the *central bank* topic contains words such *bank*, *govern*, and *billion*; the *markets* topic contains words

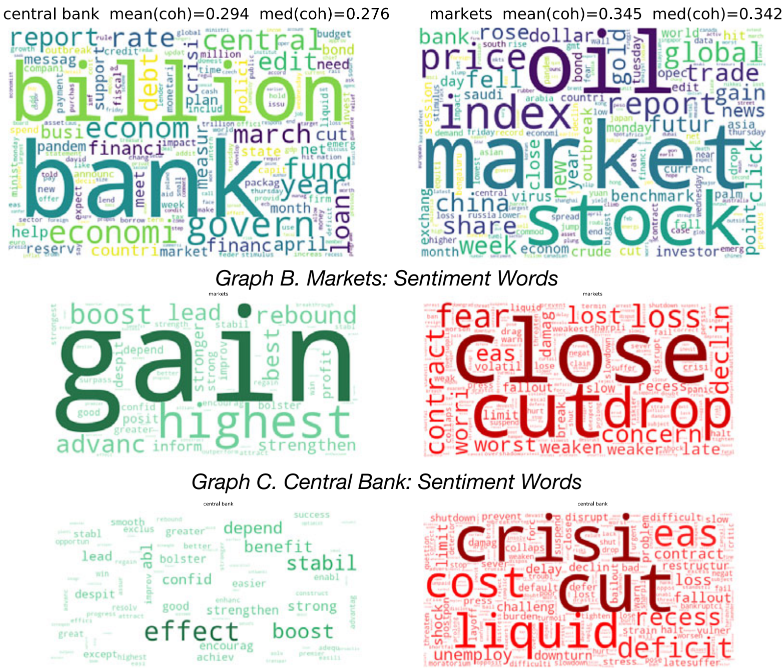
[14]These are: *said, thomsonreut, https, tmsnrt, www, reuter, coronavirus, com, nl.* Because it occurs less frequently in the corpus than *coronavirus* and because of its lower cosine similarity with other frequently occurring words, the word *covid* was not dropped.

FIGURE 3

Word Clouds for Markets and Central Bank Topics

The word clouds in Figure 3 describe topics associated with articles that mention "coronavirus" or "COVID-19." Each word cloud is labeled with the topic name, as well as the mean and median coherence for the topic (as described in Section A1 of the Supplementary Material). The coherence measures are evaluated using all 189,548 articles in the entire sample. The topic model was selected using data, including the scaled coherence score, only through the end of Apr. 2020. The sentiment word clouds show the incidence of positive and negative sentiment Loughran and McDonald (2011) words in each of the topics. Larger words indicate higher incidence. The positive and negative words are being shown on the same scale.

Graph A. Wordclouds for Select Topics and Sentiment Words



like *index*, *market*, *stock*, and *oil.* The remaining 10 topics consist of: two credit topics *credit* (about corporate credit) and *credit1* (about structured credit); three corporate topics *corp & govt US*, *corp future*, and *corp actual* (about earnings); and *health*, *currency*, and *europe*, *oil & comm*, and *sports.* Word clouds for these topics are shown in Figure A4 in the Supplementary Material.

Table 2 reports headlines of high and low sentiment articles representative of each topic. Representative articles for topic k are those that have a topic allocation to k that is above the 90th percentile (across all documents) loading on k and that have allocations to all other topics that are below the 80th percentile for the non-k topics. Most representative articles appear appropriate as their headlines match their topic classification and sentiment.[15] For example, the two *credit* headlines are about a negative and positive ratings action, respectively. The two *europe* headlines are about a positive and negative news story having to do with Europe. The topics are sorted from the most Fed-like to the most 1987-like, based on a textual proximity

[15]Table A1 in the Supplementary Material shows a larger set of sample headlines.

TABLE 2

Headlines for Representative Articles by Topic

Table 2 shows headlines of extreme sentiment (second highest and lowest for each topic) news stories whose weight for topic k is above topic k's 90th percentile value, and whose topic loadings $j \neq k$ are below the 80th percentile for j. The topics are sorted by their narrativity rank given by the residual ε_k from regression (5) in Section III.B. The 1987 and Fed columns show the Euclidean distance from (4) between these two corpora and each topic.

Label	1987	Fed	Headline	Sent	Date
CREDIT	0.094	0.114	Fitch Takes Action on 14 Italian Banking Groups on Coronavirus Disruption	−0.042	03–24
			Fitch Rates Texas Instruments' $750 Million of Five-Year Senior Notes 'A+'	0.011	03–03
CENTRAL_BANK	0.075	0.106	India extends suspension of bankruptcy filings	−0.111	09–24
			S.Korea fin min says to boost loans to developing countries fighting coronavirus	0.042	04–26
CURRENCY	0.064	0.103	RPT-BUZZ-Replay-EUR/USD doubts, sterling exposed, yen setback	−0.126	06–24
			BUZZ-EUR/USD-Sold on virus led broad risk 'off' USD strength	0.052	12–20
CREDIT1	0.102	0.120	Fitch Ratings: USPF Housing Defines Coronavirus Scenarios for Loan Program Models	−0.040	04–30
			Fitch to Rate BANK 2020-BNK30; Presale Issued	0.019	12–08
CORP_FUTURE	0.059	0.105	BUZZ-Hershey: Falls on Q1 profit miss, massive sales decline in China	−0.092	04–23
			Australia's Fortescue sees strong steel demand in 2021	0.060	12–08
SPORTS	0.086	0.119	Rugby-Champions Cup, Challenge Cup quarter-finals postponed due to coronavirus	−0.119	03–16
			Soccer-Five positive in latest Premier League COVID-19 tests	0.091	10–12
CORP_&_GOVT_US	0.066	0.112	McDonald's accused of firing worker who sued over COVID-19 claims – Bloomberg	−0.123	06–19
			Pelosi says bipartisan talks on COVID-19 relief making 'great progress'	0.062	12–10
HEALTH	0.082	0.120	Zimbabwe police arrest critics ahead of anti-government protests	−0.144	07–20
			Britain making good progress with antibody tests – junior minister	0.069	04–27
OIL_&_COMM	0.091	0.126	UPDATE 1-LNG tanker diverted from China in sign of weaker demand	−0.063	02–04
			CBOT wheat closes firm on strong demand	0.029	03–20
MARKETS	0.064	0.115	Indian stocks suffer worst day in history as coronavirus shuts businesses, cities	−0.078	03–23
			BUZZ-Australian financials extend gains to fifth day on hopes of economic rebound	0.081	11–10
EUROPE	0.076	0.121	Refinitiv Newscasts – Confusion at Heathrow as UK cut off from Europe	−0.116	12–21
			Refinitiv Newscasts – BioNTech confident vaccine can beat new mutation	0.045	12–22
CORP_ACTUAL	0.094	0.134	BRIEF-Tritech Group Updates on Business Disruptions due to COVID-19	−0.119	04–07
			BRIEF-H2O Innovation Presents Update on COVID-19 and Ensures Continuity of its Operations for its Customers	0.082	03–17

measure (the residual from equation (5)) that is explained in Section III.B. The more Fed-like headlines in Table 2 (top of table) discuss credit markets, currency behavior, and regulatory actions, while the more 1987-like headlines (toward the bottom) focus on market behavior, subjective assessments of European news, and corporate earnings.

Figure 3 also shows the prevalence of Loughran and McDonald (2011) sentiment words in the *markets* and *central bank* topics. The most prevalent negative sentiment words (largest font) in the *markets* topic are *close*, *cut*, *drop*, and *fear*, while the most prevalent positive sentiment words are *gain*, *highest*, and *boost*. Font sizes in the positive and negative word lists are on the same scale, so *gain* is the most frequently used sentiment word in either list in the *markets* topic. For the *central bank* topic the most prevalent negative sentiment words are *crisis*, *cut*, *liquid*, *cost*, and *deficit*, while the most common positive words are *effect*, *boost*, and *stabil(ity)*. The positive words are less frequent than the negative ones, as indicated by their smaller font size in the word clouds, suggesting that central bank news coverage tended to focus on negative developments. There are more unique negative than positive words across these two topics, which reflects the composition of the Loughran–McDonald (LM) dictionary: 2,355 negative and 354 positive words. Figures A5–A7 in the Supplementary Material show the LM word incidence of the other 10 topics.

A. Aggregation to Daily Level

I aggregate the topic distributions of all day t articles into a measure of *topic frequency*:

$$f_{t,k} = \frac{1}{N_t} \sum_{j \in \{\text{day } t \text{ articles}\}} f_{j,k}, \tag{2}$$

where N_t is the number of articles in day t and $f_{j,k}$ is the jth article's loading on topic k. Figure A8 in the Supplementary Material shows topic frequencies over time. These series are labeled F_[topic] in subsequent analysis. The figure also shows daily aggregate sentiment (SENT), standard deviation of article-level sentiment (SENT_SD), article counts (ART_COUNT), and coronavirus new case incidence (CORONA). Table 1 presents summary statistics. The text series are quite persistent as measured by their daily AR(1) coefficients.

To capture the sentiment associated with each topic, I follow the approach of Calomiris and Mamaysky (2019) (who find that topical sentiment forecasts country-level stock returns) and define *topical sentiment* as the product of topic frequency and daily sentiment:

$$\text{SENT}_{t,k} = \text{SENT}_t \times f_{t,k}. \tag{3}$$

This measures the extent to which daily negative or positive news flow concentrates in specific topics. For example, if day t has very negative sentiment, and has articles mainly about *central bank* and *credit* topics, the topical sentiment of those two topics would be very negative, whereas the topical sentiment of the non-prevalent topics would be close to 0. Topical sentiment series are labeled S_[topic] and are summarized in Table 1.

The 12-topic model captures a rich heterogeneity of news flow about the crisis, from early articles dealing with the health impact of coronavirus to later articles dealing with central bank interventions and credit impacts. The news measures exhibit substantial time series and cross-sectional variation which will prove useful when relating these series to market activity across multiple asset classes. Section A1.1 of the Supplementary Material gives more detail on the evolution of topic frequencies in the corpus.

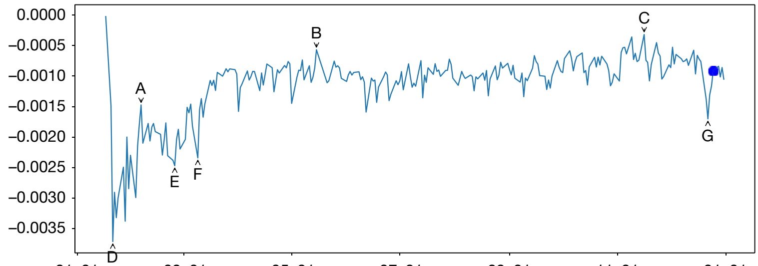
Figure 4 shows the evolution of topical sentiment, $\text{SENT}_{t,k}$ for the *markets* (Graph A) and *central bank* (Graph B) topics. Each graph shows several (subjectively identified) local peaks and troughs of topical sentiment. For each labeled point, the headline of the most positive (peaks) or negative (troughs) article

FIGURE 4

Annotated Topical Sentiment

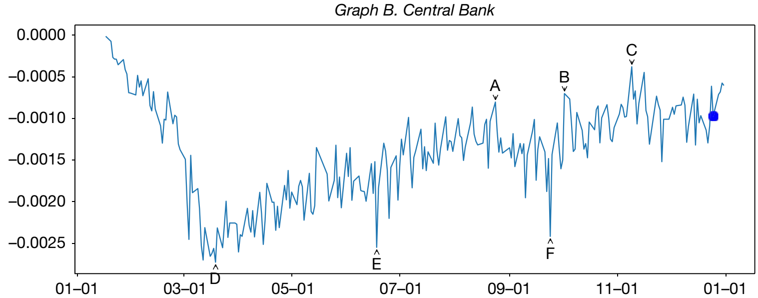
The graphs in Figure 4 show the annotated daily topical sentiment from (3). Extreme points are subjectively identified to highlight spikes in each series. Each labeled point for topic k is associated with a headline of an article that has a topic k loading above the 90th percentile across all article-topic observations for topic k and a loading on any other topic $j \neq k$ that is lower than the 80th percentile topic j loading across all article-topic observations. The headline shown for low (high) sentiment points is for the lowest (highest) sentiment article satisfying the 90/80 filter for topic k day t. The blue dot indicates Christmas of 2020, a time of relatively low news coverage. *No articles* means no articles satisfying the 90–80 filter were found on that day.

Graph A. Markets



A: BUZZ-Aussie energy index sees best day in over a month as oil prices jump (0.06)
B: No articles
C: MIDEAST STOCKS-Most major Gulf markets gain; Saudi index flat (0.03)
D: EMERGING MARKETS-Latam assets keel as China virus sours risk appetite (–0.04)
E: EMERGING MARKETS-Latam currencies trade lower as coronavirus concerns rise (–0.07)
F: CANADA STOCKS-Toronto stocks eye worst day since 2008 as energy sector reels from oil crash (–0.05)
G: RUBBER-Japan futures tumble on Shanghai plunge, virus fears (–0.06)

Graph B. Central Bank



A: UPDATE 1-Hungary's 2020 budget deficit could reach 7–9% of GDP –ministry (–0.03)
B: Brazil's central bank extends reserve requirements reduction to April 2021 (0.00)
C: Swedish govt extends COVID measures to support companies (–0.01)
D: Tunisia lets firms defer loans in face of coronavirus (–0.08)
E: BRIEF-If a second wave of coronavirus was severe like first wave, impact on Canadian economy could be serious – Bank of Canada deputy governor (–0.06)
F: India extends suspension of bankruptcy filings (–0.11)

associated with that topic on day t is shown at the bottom of each graph. An article is associated with a topic using the same 90–80 percentile rule used in Table 2. For example, point D in Graph A was the most negative sentiment day for the *markets* topic. The most negative *markets* article on that day warned about keeling Latam assets on the back of the "China virus." In Graph B, point C is the most positive late-sample sentiment day for the *central bank* topic. The most positive *central bank* headline on that day talks about the Swedish government "extend[ing] COVID measures to support companies." Figures A9–A11 in the Supplementary Material repeat this analysis for all 12 topics.

B. Topic Narrativity

To test the narrativity and rational information hypotheses of Section I.A, I first determine the narrative closeness of my 12 topics to the 1987 and Fed corpora.[16] The 1987 corpus, consisting of 164 *Wall Street Journal* articles in the days following that year's stock market crash, is meant to capture narrativity or story-telling quality in the sense of Goetzmann et al. (2022), who argue that more narrative news coverage elicits greater next-day investor crash attention and higher next-day implied volatility. The Fed corpus, consisting of 155 Federal Reserve statements since 1994, focuses largely on economic and market conditions, and the associated monetary policy decisions. I use closeness to the Fed corpus as a measure of fact-based, as opposed to narrative, reporting.

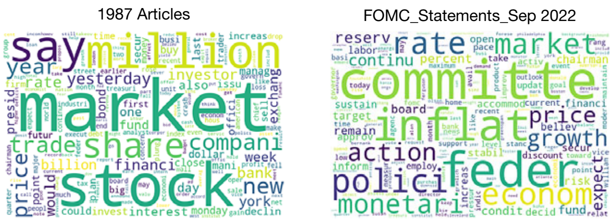
Graph A of Figure 5 shows word clouds for the 1987 and Fed corpora, after stop words are excluded. The font size of each word corresponds to that word's relative frequency in the corpus. Among the most frequent words in the 1987 corpus are *market*, *stock*, *million* (indicative of magnitudes of crash impacts), *share*, *trade*,

FIGURE 5

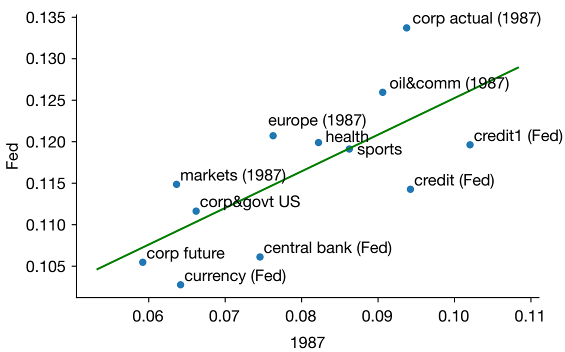
Narrativity Measures

Graph A of Figure 5 shows word clouds for word frequency, after excluding stop words, of the 1987 and Fed corpora that are described in Section II. For the 12 model, Graph B shows a regression of each topic's Fed Euclidean distance against each topic's 1987-corpus Euclidean distance. The Euclidean distance calculation is explained in Section III.

Graph A. Narrativity Measures



Graph B. Fed vs 1987 Narrativity Distance



[16]I thank an anonymous referee for pointing the analysis in the article in this direction.

and *price*. All these words would intuitively be used to describe an event like the 1987 stock market crash. On the other hand, the most frequent words in the Fed corpus are *committee*, *inflat(ion)*, *feder(al)*, *polic(y)*, *econom(y)*, and *monetar(y)*. Both corpora relate to economic outcomes and conditions, but clearly, they emphasize very different aspects of these phenomena.

I measure the proximity of each of my 12 topics to the two corpora by using the negative of their Euclidean distance,[17] which for topic k and corpus C is

$$C_k \equiv \left(\sum_{w=1}^{W} (\tau_k(w) - \tau_C(w))^2 \right)^{1/2}, \tag{4}$$

where W is the number of words common to both topic k and corpus $C \in \{1987, \text{FED}\}$, $\tau_k(w)$ is the loading on word w of topic k and $\tau_C(w)$ is the fraction of corpus C represented by word w (among the W common words).[18] The 1987 and Fed columns of Table 2 present the value of the Euclidean distance between each of my 12 topics and these corpora. Graph B of Figure 5 shows a scatter plot of the 12 topics with each topics' 1987 distance measure 1987_k on the x-axis, and its Fed distance measure FED_k on the y-axis.

As is clear from the scatter plot, there is a positive correlation between FED_k and 1987_k, as some topics are closer to both corpora, while others are further away. The green line in the figure is the regression fit of topics' Fed-distance on their 1987-distance:

$$\text{FED}_k = a + b \times 1987_k + \varepsilon_k, \tag{5}$$

where FED_k and 1987_k are given by (4). The residuals from this regression measure relative proximity to the 1987 and the Fed corpora. The points lying above the line (e.g., *markets*, *europe*, *oil&comm*, and *corp actual*) are those that are *furthest* from the Fed corpus given their distance to the 1987 corpus; alternatively, these four topics are closed to the 1987 corpus given their Fed distance. On the other hand, the points below the regression line are topics that are closer to the Fed corpus given their 1987 distance; these topics are *currency*, *central bank*, *credit*, and *credit1*.

Keep in mind that the topic word distributions are obtained from my 72,263 (April and earlier) Reuters news articles, while the distance measures are to two

[17] Results that use cosine similarity as the proximity measure are similar.

[18] Much of the recent literature (Tetlock (2011), Hoberg and Phillips (2016), and Cohen, Malloy, and Nguyen (2020)) uses a bag-of-words approach to assess the similarity of two documents based on the proximity of their word count vectors. In contrast, I analyze the similarity of news topics (groups of naturally co-occurring words) to corpora with a known narrative slant. This refinement of the standard methodology (i.e., comparing topic similarity rather than document similarity) is well suited to the article's focus on understanding news flow via topic incidence. GKS measure similarity with neural network based Doc2Vec embeddings (which convert the text of documents into high-dimensional vectors). However, Doc2Vec is not well suited for analyzing the narrative content of news topics that, because they ignore word order, do not easily lend themselves to characterization via embeddings. There has been recent work on combining neural network language models with topic models (see Churchill and Singh (2022)) which may yield promising insights for future work.

corpora that are largely independent of the events of COVID-19 (the Fed corpus has a small overlap). I also assigned topic names prior to performing the 1987 and Fed proximity analysis. Despite this, the ranking of topic proximity to the two corpora is highly intuitive. The sample headlines in Table 2 are shown for topics sorted by their residuals ε_k from (5). As already mentioned in Section III, the sample headlines line up well with topical proximity to the Fed (top of table) and 1987 (bottom) corpora, respectively.

To better understand how narrativity of the COVID-19 news flow evolved over 2020, I use the daily topic incidence measures from (2) combined with the residual narrativity rankings from (5) to construct a daily measure of narrativity, v_t, as follows:

$$v_t = \sum_{k=1}^{12} f_{t,k} \times \varepsilon_k. \tag{6}$$

Note that $f_{t,k}$ changes daily while ε_k is fixed for the whole sample. The daily narrativity measure v_t will be high on days with relatively higher incidence of 1987-like topics (which have positive residuals in equation (5)) and will be low on days with relatively high incidence of Fed-like topics. Graph A of Figure 6 shows the evolution v_t. This frequency-weighted average of narrativity starts out high in early 2020, peaks sometime in March, and then settles down into a lower range for the rest of 2020. Because, by definition, $\sum_{k=1}^{12} f_{t,k} = 1$, this pattern is not mechanically driven by high early news volume (as in Figure 1). Early news coverage of the pandemic was thus highly narrative in nature.[19]

Graph B of Figure 6 multiplies v_t by daily sentiment SENT_t. This highlights that highly narrative (i.e., proximate to the 1987 crash corpus) days were also days of particularly negative sentiment. Figure 6 emphasizes that high narrativity was related to the onset of the COVID-19 pandemic and was prevalent in the early part of the sample. Furthermore, higher narrativity was associated with more negative sentiment. In Section VI, I use the cross section of the 1987- and Fed-proximity measures across topics to conduct tests of the narrativity and rational information hypotheses.

IV. Behavior Before the Break

Assuming investor expectations are linear in the observable state variables and using the Campbell and Shiller (1988) and Campbell (1991) log linearization of returns, I show in Section A2.1 of the Supplementary Material that the contemporaneous relationship between news and returns is given by

$$h_{t+1} = c + b^{\top} w_{t+1} + e_{t+1}, \tag{7}$$

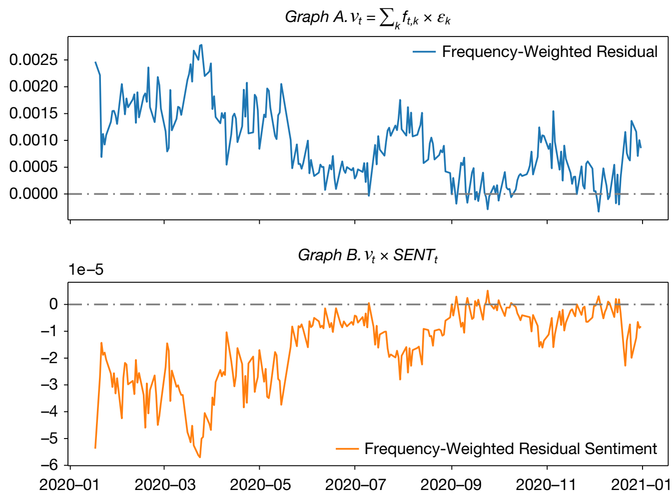
where h_{t+1} is the asset return from day t to $t+1$, c is a constant, w_{t+1} is a $k \times 1$ vector of news flow and other information available to the econometrician at time $t+1$, b is

[19]The seasonal peaks in the series are due to *corp actual* having the highest residual in (5) and to its focus on corporate earnings which have a strong seasonal pattern (Figure A8 in the Supplementary Material).

FIGURE 6

1987 Versus Fed Narrativity

Using the narrativity residuals from (5) (shown in Figure 5), Graph A of Figure 6 shows the daily frequency-weighted residual $v_t = \sum_k f_{t,k} \times \varepsilon_k$ from (6), where the sum is over the 12 model topics (note that $\sum_k f_{t,k} = 1$ at every time t). Higher levels of v_t indicate daily narrative proximity to the 1987 corpus, and lower values of v_t indicate daily narrative proximity to the Fed corpus. Graph B shows the frequency-weighted residual times the daily sentiment (i.e.,) $v_t \times \text{SENT}_t$.



a $k \times 1$ coefficient vector, and e_{t+1} consists of unobservable information. This contemporaneous markets–news specification has been used in Tetlock, Saar-Tsechansky, and Macskassy (2008) and Glasserman, Li, and Mamaysky (2024) for returns and in Boudoukh, Feldman, Kogan, and Richardson (2018) for squared returns.

As I show in Section V, a structural break in (7) took place around the middle of March of 2020. In the remainder of this section, I analyze the pre-break period (i.e., dates prior to Mar. 15, 2020 (a Sunday)). I then show in Section A3.2 of the Supplementary Material that these results are robust to the choice of Mar. 15, 2020 as the break date.

A. Contemporaneous Relationships

I specialize (7) as follows:

$$h_{t+1} = c + b_1 h_t + b_2 h_{t-1} + b_3 N_{t+1} + b_4 N_{t+1}\left(\text{VIX}_t^{10} - \bar{\text{VIX}}^{10}\right) + b_5 \text{VIX}_t^{10} + e_{t+1}. \tag{8}$$

Standard errors use Newey–West with three lags. The observable information includes two lagged returns, h_t and h_{t-1}, which control for the dependence of current news flow on past returns, and for the auto-correlation properties of the dependent variable. In addition w_{t+1} contains N_{t+1}, which can be one of: aggregate daily sentiment SENT_t, the daily standard deviation of article-level sentiment SENT_SD, the daily article count ART_COUNT, one of the 12 topical sentiment series $\text{SENT}_{t,k}$, or the COVID-19 case count series CORONA.

In some specifications, N_{t+1} is given by the linear combination of the other news series with the highest explanatory power for the contemporaneous returns of the 5 asset classes. Similar to Cochrane and Piazzesi (2005), I run the following pooled regression:

$$\tilde{h}_t^{(m)} = a + b^{\top}\mathcal{N}_t + \varepsilon_t^{(m)}, \tag{9}$$

where $\tilde{h}^{(m)}$ is a normalized version of the return series m and $\mathcal{N}_t \in \mathbb{R}^{14}$ is a vector containing the 12 topical sentiment series, article count, and SENT_SD on day t (SENT_t is not needed because it spanned by the topical sentiment series). Each market return is normalized to have zero mean and unit standard deviation, and the VIX change series is scaled by -1 (which makes all scaled series positively correlated). The news factor is then defined as $\text{FACTOR}_t \equiv \hat{a} + \hat{b}^{\top}\mathcal{N}_t$. I calculate two versions of this factor: the first (FACTOR) estimates (9) over pre- and post-Mar. 15, 2020 subsamples separately and concatenates the two series; the second (FACTOR_ALL) estimates (9) over the entire sample. Coefficient estimates for FACTOR_ALL are shown in Table A3 in the Supplementary Material.

The model in Glasserman, Mamaysky, and Shen (2024) suggests information shocks (e.g., media focus on the economic impacts of COVID-19) can push the economy into a high-information production, high-volatility regime, where asset prices become both depressed and extremely sensitive to new information. Because the VIX is very negatively autocorrelated in the sample (Table 1), I use a 10-day average of the VIX, VIX_t^{10}, to smooth out high-frequency variation. The $N_{t+1}\left(\text{VIX}_t^{10} - \overline{\text{VIX}}^{10}\right)$ term in (8) tests if the effect of N_{t+1} on returns differs depending on whether the level of volatility is currently high or low. $\overline{\text{VIX}}^{10}$ is the average of VIX_t^{10} in the time period over which the regression is being estimated. To avoid endogeneity issues, I use VIX_t^{10} on day t (not $t+1$). GKS propose measuring investor attention to crash narratives via Google search volumes for terms like “market crash.” Figure 2 shows that the VIX and VIX_t^{10} are highly correlated the GKS-proposed Google search measure, so the interaction term in (8) can also be thought of as a proxy for investor crash attention.

For the S&P 500 and HY indexes, h_{t+1} equals the day t to day $t+1$ total return. For the VIX, h_{t+1} is the day-over-day difference in the VIX index, which is highly correlated with the daily return from investing in a VIX futures contract. For 2- and 10-year Treasuries, h_{t+1} is the day-over-day change in yields. Given the relatively low level of rates in 2020, Treasury duration did not change materially during the sample period meaning Treasury returns are, to first-order, linear in yield changes. Therefore, for all asset classes, returns either equal h_{t+1} or are approximately linear in h_{t+1}. For each of the five dependent variables, I run 18 different versions of (8), one for each of the possible N_{t+1}s (12 topics, SENT_t, SENT_SD, ART_COUNT, CORONA, FACTOR, and FACTOR_ALL). The benefit of having multiple news topics and markets is the ability to explore a rich heterogeneity of markets–news relationships, which is especially important for the hypothesis tests in Section VI.

Table 3 summarizes the results of estimating (8) in the pre-Mar. 15, 2020, part of the sample. The column groupings correspond to a particular market variable, and the rows correspond to one of the 18 different news variables. b_3 (labeled EV,

https://doi.org/10.1017/S002210902300131X Published online by Cambridge University Press

TABLE 3

Summary of Analysis for the All-Article Corpus: Early (Pre-Break) Subsample

Table 3 is a summary of the contemporaneous and lead–lag results for regressions of daily market returns on text-based series. The column groupings correspond to different market variables and the rows correspond to the text-based explanatory variables. The first two entries for every market variable show the b_3 (EV, for explanatory variable) and b_4 (EV × VIX) coefficients in (8) that are significant at the 10% level or better. The EV column shows the impact of a 1-standard-deviation change in the explanatory variable in units of standard deviation of the market variable. The EV × VIXL1 column shows the impact of a unit increase in VIX[10] on the value of EV (which is in units of standard deviation). The last entry for each market variable indicates the c_3 (EVL1) coefficients from (10) that are significant at the 10% level or better. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The summary statistics underneath the table are: the mean absolute b_3 among significant b_3 coefficients; SIG is the number of significant b_3 coefficients; JNT_SIG (joint significance) is the number of specifications where b_3 and b_4 are both significant; HYPER (hypersensitivity) is the number of times that both b_3 and b_4 are significant and have the same sign; LEAD (lead–lag relationship) is the number of times that c_3 is significant; UNDER (underreaction) is the number of times b_3 and c_3 are both significant and have the same sign; OVER (overreaction or reversal) is the number of times b_3 and c_3 are both significant have opposite signs; and OVER + HYPER (overreaction and hypersensitivity) is the number of time b_3 and c_3 have opposite signs while b_3 and b_4 have the same sign.

	SP500			VIX			HY			GT2			GT10		
	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1
SENT	0.761***	0.085***		−0.701***	−0.081***		0.603***	0.104***	−0.347**	0.616***	0.093**		0.574***	0.092*	−0.218*
SENT_SD	−0.742**	−0.106**	1.137***	0.857***	0.117***	−1.098***	−1.100**	−0.180***	1.181***			0.906***			1.231***
FACTOR	0.773***	0.078***		−0.705***	−0.066***		0.607***	0.101***	−0.493***	0.826***	0.081***		0.725***	0.100***	−0.368**
FACTOR_ALL	0.522*			−0.471*					−0.360*	0.943***	0.087***		0.703***		
ART_COUNT	−1.240**			1.485***	0.034*	−0.578*	−0.792**	−0.045**	1.506***	−0.957*					
CORONA		−0.046**	0.868***		0.046**	−0.823***		−0.041**	0.222*			0.401**	0.426**	0.067**	0.505***
SPORTS	0.940**		−0.613*	−1.041***			0.645**	0.035*	−1.333***	1.129**			0.891*		
CENTRAL_BANK										0.615*					
MARKETS	0.857***	0.075***		−0.798***	−0.069***		1.079***	0.110***	−0.442*	0.426*	0.078***	−0.892***	0.960***	0.139***	−1.077**
HEALTH	1.053***	0.104**		−1.099***	−0.113***		1.245***	0.171***	−0.892***				0.940*		
EUROPE	0.985***	0.078*		−0.829***			0.676***	0.101***							−0.617**
OIL_&_COMM	0.663***	0.215***		−0.460***	−0.187***		1.073***	0.223***		0.817***	0.169***	−0.808***	1.110***	0.232***	−0.863***
CURRENCY	0.771**			−0.681**						1.262***	0.148***		1.127***	0.139***	
CREDIT										0.576*	0.061*				
CORP_&_GOVT_US							0.338**		0.379*					−0.077**	
CORP_ACTUAL			0.297**			−0.403**									
CORP_FUTURE										0.531**	0.089***		0.457***	0.095***	
CREDIT1															

Mean $|b_3|$ = 0.815, SIG = 53, JNT_SIG = 38, HYPER = 38, LEAD = 29, UNDER = 2, OVER = 17, OVER + HYPER = 16.

for *explanatory variable*) is normalized to report the effect of a 1-standard-deviation change in the news variable in units of standard deviations of the market variable; b_4 (EV × VIXL1) shows the change in the normalized b_3 for a one-unit increase in the lagged VIX^{10}. For example, the value of -0.701 for the effect of aggregate sentiment on the VIX means when aggregate sentiment increases by 1-standard-deviation, the VIX experiences a contemporaneous decline of 0.701 standard deviations of daily changes. The average absolute value of b_3 is 0.815 (shown at bottom of the table), suggesting that the contemporaneous effect from a 1-standard-deviation news shock results in an almost 1-standard-deviation market response, on average. Non-significant, at the 10% level, b_3 and b_4 coefficients are left blank. Tables A4–A8 in the Supplementary Material detail the regression estimates of the model in (8) for any specification with a significant b_3 or b_4 coefficient.[20]

All b_3 coefficients for sentiment series in specifications involving the S&P 500, HY, and 2- and 10-year Treasury yields that are significant have positive signs. So negative news as conveyed by either aggregate sentiment or any topical sentiment series is associated with contemporaneous declines in the S&P 500 and HY indexes, and with contemporaneous decreases in 2- and 10-year Treasury yields. For the S&P 500, HY, and 2-year Treasury yields, the signs of b_3 for SENT_SD and ART_COUNT are negative, when significant, suggesing a contemporaneous negative return or drop in yields associated with a higher standard deviation of article-level sentiment or higher article count. For VIX, all signs are reversed. Bad news, as proxied by low aggregate or topical sentiment, and high news dispersion (SENT_SD) or article count, are associated with a contemporaneous VIX increase.

In total there are 53 significant instances of b_3. A striking feature of Table 3 is that in all 38 cases when both b_3 and b_4 are significant, they have the same sign. On high-volatility (or high crash attention) days, as measured by an elevated level of the 1-day lagged VIX_t^{10}, the effects of news are larger than they are during normal-volatility days. I call this phenomenon *hypersensitivity* to contemporaneous news. A portion of the volatility that asset markets experience in high-volatility states is not due to an increased volatility of news flow, but instead is due to an increased sensitivity of markets to similarly volatile news, consistent with the model in Glasserman, Mamaysky, and Shen (2024).

COVID-19 Case Incidence and the Aggregate News Factor

As mentioned earlier, day t case counts (the *corona* series) come out after day t's market close, and thus markets can react to these only with a 1-day lag. Even adjusting for this lag, the effect of *corona* is only significant for 10-year Treasury yields. This corroborates the observation in Figure 1 that S&P 500 returns are more related to measures of news flow than they are to the actual incidence of COVID-19 infections.

The FACTOR series has highly significant b_3 and b_4 coefficients (largely by construction) for all market returns in the early sample. The FACTOR series also exhibits hypersensitivity for all markets, suggesting that hypersensitivity is a

[20]The Chg R^2 column in Tables A4–A8 shows that the incremental contribution of the VIX interaction term in (8) is large.

general phenomenon and not simply a feature of a particular markets–news pair. Also FACTOR_ALL exerts a significant influence on contemporaneous returns, but less so than FACTOR, which points to a structural break in the data (since the FACTOR composition changes from the early- to the late-subsample) as will be discussed in Section V.

B. Reversals

I next analyze the tendency of markets–news pairs to exhibit reversals during the pre-break sample by conducting a series of Granger causality tests. The general specification of these tests mirrors the contemporaneous regression in (8):

$$(10)\qquad \rho_{t+1} = c_0 + c_1\rho_t + c_2\rho_{t-1} + c_3\tau_t + c_4\tau_t\left(\mathrm{VIX}_t^{10} - \bar{\mathrm{VIX}}^{10}\right) + c_5\mathrm{VIX}_t^{10} + e_{t+1}.$$

Here, ρ_{t+1} is the day $t+1$ response variable, and τ_t is the day t test variable. The two lags of the response variable control for contemporaneous correlations with the test variable, and for the possibility that the test variable is itself Granger caused by the lagged response variable. I say that τ Granger causes ρ if the c_3 coefficient above is significant at the 10% level or better. Standard errors are calculated using Newey–West with three lags.

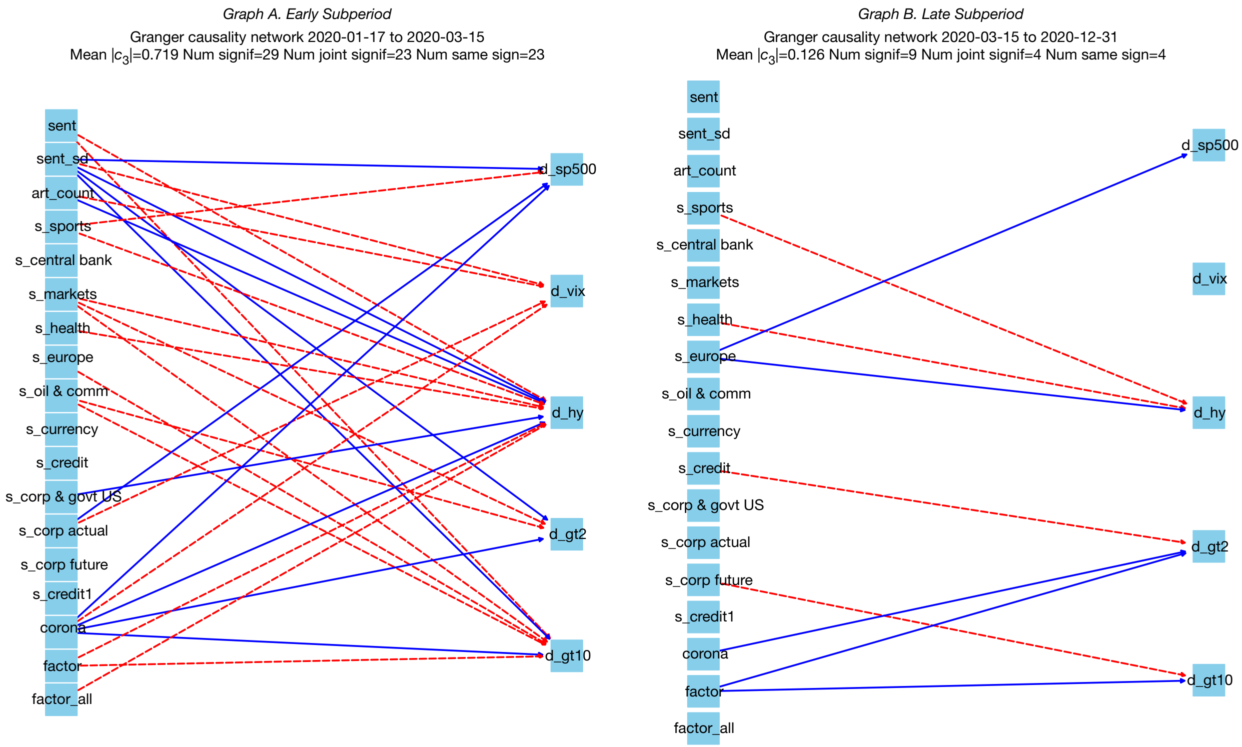
I estimate (10) in the early subsample with markets as the response variables, and the news series as the test variables. For each markets–news pair, the EVL1 column of Table 3 reports all significant c_3 coefficients from the Granger causality regressions of future markets on current news. There are 29 cases of news variables Granger causing next-day market variables, as indicated by a significant entry in the EVL1 column. Table A14 in the Supplementary Material shows the details of these regressions. Graph A of Figure 7 shows a graph representation of the Granger causality network. The blue solid arrows represent significant positive links from lagged news to future markets, and red dashed arrows represent significant negative links. The network is relatively dense with 29 significant connections, as many news series Granger cause next day market returns. Also worth noting is that there are 23 instances of c_3 and c_4 coefficients that are both significant; in all 23 cases they have the same sign suggesting that hypersensitivity is a prominent feature of how news Granger cause next day market returns.

Overreaction, or price reversal, occurs when the contemporaneous news coefficient, b_3 from (8), and the lagged news coefficient, c_3 from (10), are both significant and have opposite signs. This indicates that the time t market effect of a news variable is partially or fully reversed at time $t+1$. Note the c_3 coefficient reported in column EVL1 of Table 3 is scaled in the same way as coefficients b_3 and b_4 in the EV and EV × VIXL1 columns; it is reported in standard deviations of the market variable per unit of standard deviation of the news variable. Thus the magnitudes of coefficients in the EV (contemporaneous regression) and EVL1 (lagged news regression) columns of Table 3 are directly comparable. For example, in the early subsample, SENT_SD's effect on S&P 500 contemporaneous returns is a decrease of 0.742 standard deviations of S&P 500 returns for a 1-standard-deviation increase in SENT_SD. The coefficient of lagged SENT_SD for future S&P 500 returns suggests a 1-standard-deviation increase in SENT_SD

FIGURE 7

Granger Causality Network: News to Markets

Figure 7 displays Granger causality tests, using Newey–West standard errors with three lags. The D_[MARKET] variables refer to daily returns or changes in the particular market series. A link is shown if the lagged test variable is significant at the 10% or better level in equation (10). A blue, solid (red, dashed) line indicates the coefficient c_3 from the test to the response variable in (10) is positive (negative). The text above each graph shows the average $|c_3|$ coefficient among all significant markets–news pairs, the number of links in the graph (Num signif), the number of links that are associated with a significant c_4 coefficient from equation (10) (Num joint signif), and the number of times when both c_3 when c_4 are significant and have the same sign (Num same sign).



forecasts a next-day positive S&P 500 return of 1.137 standard deviations. This completely reverses the contemporaneous S&P 500 reaction to the standard deviation of article-level sentiment. The average value of $|c_3|$ for significant c_3s in (10) (shown in Graph A of Figure 7) is 0.72, suggesting that during the pre-break period lagged news have a large effect on 1-day ahead market returns.

Counting instances of b_3 and c_3 coefficients with opposite signs indicates that this occurs 17 times (see Table 3) in the early subsample (out of 29 cases of significant c_3s). Of these 17 instances of overreaction, 16 occur in the presence of hypersensitivity in the contemporaneous markets–news relationship, suggesting that hypersensitivity is associated with contemporaneous overreaction, and thus next day reversals.

C. News Tone Response to Markets

I next check whether markets Granger cause news. I estimate (10) with daily changes in the five market series as the test variables, and the next day's changes in the 18 news series as the response variables. Graph A of Figure 8 shows the results as a Granger causality graph, and Table A15 in the Supplementary Material shows the detailed regression results. Every significant link from a time t market variable to a time $t+1$ news variable is shown as an arrow. The arrow is blue and solid (red and dashed) when the c_3 coefficient in (10) is significant and positive (negative).[21] The direction of Granger causality is intuitive. Positive time t returns for S&P 500 and HY lead to more positive time $t+1$ news coverage and fewer time $t+1$ articles; the same is true of increases in 2- or 10-year Treasury yields. A positive move in the VIX leads to more negative sentiment, higher article counts, and higher standard deviation of sentiment across articles (SENT_SD) on day $t+1$. There are 49 markets–news pairs with significant c_3 coefficients, suggesting a very dense Granger causality network (the maximum number of links is 90). The c_3 coefficient is normalized to report standard deviation changes in the news variable for a 1-standard-deviation change in the market variable. The average value of $|c_3|$ for significant markets–news pairs is 0.40, suggesting that the economic magnitude of the effect is large.

V. Structural Break

Section IV showed news and markets are tightly coupled in the pre-break sample. In this section, I discuss the structural break tests for the 90 markets–news pairs (5 asset classes and 18 news series) under consideration. The post-break markets–news behavior is notable for how starkly it contrasts with the pre-break behavior.

A structural break in (8) at a known break point $t=\pi T$, where $\pi \in [0,1]$ and T indicates the end of the sample, can be detected using the Chow test statistic:

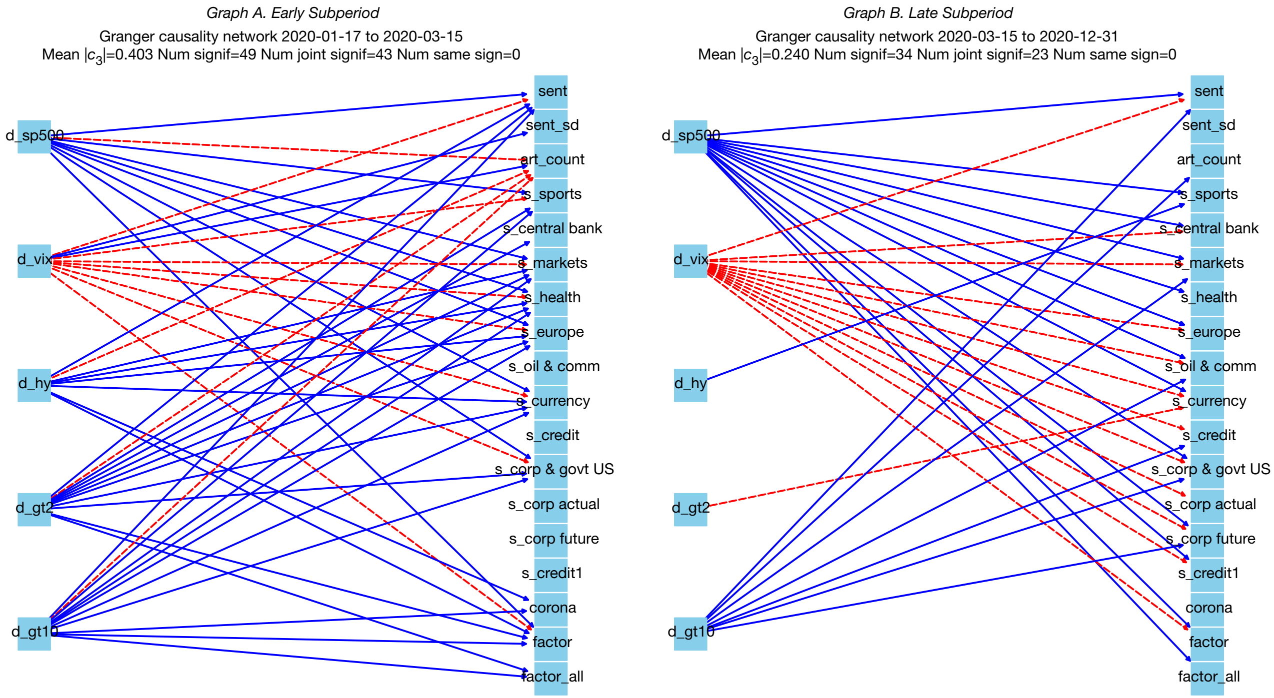
$$\phi(t)=\frac{(\mathrm{SSR}-\mathrm{SSR}_e-\mathrm{SSR}_l)/k}{(\mathrm{SSR}_e+\mathrm{SSR}_l)/(N_e+N_l-2k)}, \quad (11)$$

[21]Since (10) controls for time t and time $t-1$ news, c_3 captures the impact on $t+1$ news of the portion of day t market prices that is orthogonal to contemporaneous and lagged news.

FIGURE 8

Granger Causality Network: Markets to News

Figure 8 shows Granger causality tests, using Newey–West standard errors with three lags. The D_[MARKET] variables refer to daily returns or changes in the particular market series. A link is shown if the lagged test variable is significant at the 10% or better level in equation (10). A blue, solid (red, dashed) line indicates the coefficient c_3 from the test to the response variable in (10) is positive (negative). The text above each graph shows the average $|c_3|$ coefficient among all significant markets–news pairs, the number of links in the graph (Num signif), the number of links that are associated with a significant c_4 coefficient from equation (10) (Num joint signif), and the number of times when both c_3 when c_4 are significant and have the same sign (Num same sign).



https://doi.org/10.1017/S002210902300131X Published online by Cambridge University Press

where e (early) refers to dates prior to or equal to t and l (late) refers to those dates after t, SSR refers to the sum of squared residuals over the entire sample, SSR_e (SSR_l) refers to the sum of squared residuals over the early (late) part of the sample, N_e (N_l) refers to the number of observations in the early (late) part of the sample, and k refers to the number of regressors. As the number of observations grows, $k\phi(t)$ approaches a χ^2 distribution with k degrees of freedom.

When t is not known, the t^* that maximizes $\phi(t)$ over the entire sample can be chosen. However, since t^* is chosen to maximize the Chow statistic in (11), $k\phi(t^*)$ will no longer be asymptotically χ^2. Andrews (1993), (2003) tabulates the distribution of $k\phi(t^*)$ when t^* is selected over all possible t's in some interval of the data. For an unknown break point π which is drawn from an interval $[\pi_0, 1-\pi_0] \subset (0,1)$ of the data and the null hypothesis of no structural break, Andrews (1993), (2003) tabulates the distribution of:

$$\sup_{\pi \in [\pi_0, 1-\pi_0]} k\phi(\pi T). \tag{12}$$

The tabulated distribution depends on k, which equals 6 for the specification in (8), and on π_0 which I set to 0.151.[22] The test is not sensitive to the choice of π_0. I run the break test using the Apr. 2020 topic model and the full data sample.

For each of the 5 market variables, Figure 9 shows a histogram of the optimal break points $t^* = \pi^* T$ that maximize (12) for each of the 18 text series. For every break point, I also tabulate the number of markets–news pairs that are significant at least at the 10% level according to the distribution tables in Andrews (2003). Graph F shows the distribution of break points for all 90 tests conducted, as well as the number of break points that are significant.[23]

All 5 asset classes exhibit strong evidence of a regime break in March. For each asset class, all 18 markets–news pairs show evidence of a significant break. The results are consistent across all markets–news pairs. The break dates are concentrated on Mar. 9, Mar. 16 and 17, and Mar. 23. March 16 and 17 are the Monday and Tuesday following an emergency rate cut announced by the Fed on Sunday, Mar. 15. And the Mar. 23 break date follows another emergency Fed meeting on Sunday, Mar. 22 when the Fed removed quantitative guidance from its announced emergency programs and vowed to simply purchase Treasury and mortgage-backed securities "in the amounts needed." Markets reacted favorably to the Fed's, and other central banks' announcements, which were perceived to be expansive and bold (see Hartley and Rebucci (2020)). Section VI connects the timing of the breaks with the narrativity and information hypotheses.

I use Sunday, Mar. 15, 2020 as the cutoff date between the early and late parts of the sample. Section A3.2 of the Supplementary Material shows the article's results hold if break dates are set to the optimal date for each markets–news pair or

[22]The value of $\bar{\text{VIX}}^{10}$ affects only the b_3 coefficient in (8) but leaves the residuals of the regression unchanged, so $\phi(t)$ in (11) is unaffected by $\bar{\text{VIX}}^{10}$, which I set to 0 for the break tests. For the early (Section IV) and late (Sections V.A and V.B) subsample analysis, I measure $\bar{\text{VIX}}^{10}$ in each interval separately.

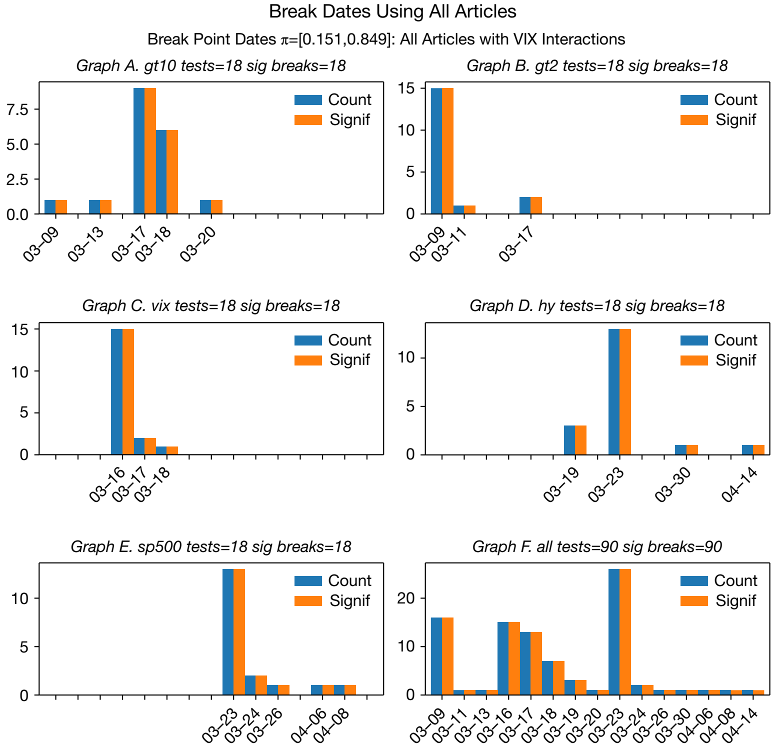
[23]Tables A4–A8 in the Supplementary Material list the break dates corresponding to t^* and the associated $k\phi(t^*)$ value in square brackets.

https://doi.org/10.1017/S002210902300131X Published online by Cambridge University Press

FIGURE 9

Break Dates Using All Articles

For each market variable shown in Figure 9, there are 18 regressions which are tested for a break (12 topical sentiments, overall sentiment, standard deviation of sentiment, article count, FACTOR, FACTOR_ALL, and COVID-19 case counts). The data start on Jan. 17, 2020. The starts show number of break points that are significant at the 10% level or better using the Andrews (2003) distribution for the maximal Chow statistic with $\pi_0 = 0.151$.



to the average break date for all pairs. Section A3.3 of the Supplementary Material shows the March structural break could have been identified as early as the end of April, motivates the use of the April topic model, and shows that tests using additional data as 2020 progressed still pointed to a mid-March structural break.

A. Contemporaneous Relationships

Table 4 explores the contemporaneous markets–news relationship in (8) in the post-break sample; only coefficients significant at the 10% level or better are shown. Tables A9–A13 in the Supplementary Material show the full regression results for all explanatory variables where at least one of the b_3 and b_4 coefficients is significant. There are now only 32 markets–news pairs associated with significant b_3 coefficients, and 18 with significant b_3 and b_4 coefficients. These compare, respectively, to 53 and 38 significant markets–news pairs documented in the pre-break subsample in Table 3. More strikingly, the average $|b_3|$ is now 0.253, compared to 0.815 in the pre-break subsample. Not only are there fewer markets–news pairs that are significant in (8), but when they are significant the

TABLE 4

Summary of Analysis for the All-Article Corpus: Late (Post-Break) Subsample

Table 4 is a summary of the contemporaneous and lead–lag results for regressions of daily market returns on text-based series. The column groupings correspond to different market variables and the rows correspond to the text-based explanatory variables. The first two entries for every market variable show the b_3 (EV, for explanatory variable) and b_4 (EV × VIX) coefficients in (8) that are significant at the 10% level or better. The EV column shows the impact of a 1-standard-deviation change in the explanatory variable in units of standard deviation of the market variable. The EV × VIXL1 column shows the impact of a unit increase in VIX[10] on the value of EV (which is in units of standard deviation). The last entry for each market variable indicates the c_3 (EVL1) coefficients from (10) that are significant at the 10% level or better. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The summary statistics underneath the table are: the mean absolute b_3 among significant b_3 coefficients; SIG is the number of significant b_3 coefficients; JNT_SIG (joint significance) is the number of specifications where b_3 and b_4 are both significant; HYPER (hypersensitivity) is the number of times that both b_3 and b_4 are significant and have the same sign; LEAD (lead–lag relationship) is the number of times that c_3 is significant; UNDER (underreaction) is the number of times b_3 and c_3 are both significant and have the same sign; OVER (overreaction or reversal) is the number of times b_3 and c_3 are both significant have opposite signs; and OVER + HYPER (overreaction and hypersensitivity) is the number of time b_3 and c_3 have opposite signs while b_3 and b_4 have the same sign.

	SP500			VIX			HY			GT2			GT10		
	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1	EV	EV × VIXL1	EVL1
SENT	0.375***	0.027**					0.472***	0.035***			0.014**				
SENT_SD				0.169**											
FACTOR	0.154***	0.018***		−0.221**			0.209***	0.023***		0.147**		0.073*	0.119*		0.076*
FACTOR_ALL	0.166**			−0.242**			0.241***								
ART_COUNT							−0.354*				−0.015***				
CORONA				−0.230*							0.028***	0.149*			
SPORTS							0.288***	0.045***	−0.204**						
CENTRAL_BANK	0.337***	0.023**					0.444***	0.028**			0.013**				
MARKETS	0.281***	0.031***		−0.140**	−0.012*		0.249***	0.033***							
HEALTH	0.232**	0.022***					0.288***	0.023***	−0.108*						
EUROPE	0.434***	0.033***	0.194**	−0.330**			0.456***	0.037***	0.176*						
OIL_&_COMM							0.203**								
CURRENCY	0.230***			−0.203**	−0.023*		0.244***	0.030***							
CREDIT												−0.083*			
CORP_&_GOVT_US					0.016*		0.201***	0.026***			−0.026**				
CORP_ACTUAL							0.162*				0.014**				
CORP_FUTURE	0.139**						0.124**	0.018**							−0.073*
CREDIT1															

Mean $|b_3|$ = 0.253, SIG = 32, JNT_SIG = 18, HYPER = 18, LEAD = 9, UNDER = 4, OVER = 2, OVER + HYPER = 2.

magnitude of the effect is more than three times lower! As before, in all cases where b_3 and b_4 are jointly significant, they have the same sign, indicating evidence of hypersensitivity even in the post-break part of the sample. However, the prevalence of hypersensitivity in the late subsample is far lower than in the pre-break subsample (18 vs. 38), suggesting hypersensitivity is a feature of stressed markets, but is less so during more typical market conditions. Overall, contemporaneous markets–news relationships in the post-break sample are much weaker than in the pre-break sample.

The majority of markets–news relationships in the late subsample involve the S&P 500 and HY index. These are also the two markets with a 7:1 ratio of discount rate to cash flow belief variation in a Campbell (1991) decomposition, as explained in Section VI. Apparently, markets with a higher proportion of variability from changing beliefs about discount rates are more responsive to news. The VIX is also fairly responsive to news, though with little evidence of hypersensitivity. Two- and 10-year Treasuries are mostly unaffected by contemporaneous news in the late subsample, with the exception of FACTOR, which is not surprising. The direction of the impact is largely the same as in the early subsample, with positive news associated with contemporaneous increases in the S&P 500, HY, and the two rates series, and a drop in the VIX. As before increased dispersion of intraday news, SENT_SD, is associated with an increased VIX.

B. Lead–Lag Relationships

The right graph of Figure 7 shows the Granger causality network which results from estimating (10) in the later part of the sample where markets are the response variables and news series are the test variables. Table A16 in the Supplementary Material shows the details of these regressions. Each significant positive (negative) c_3 coefficient is shown as a solid blue (dashed red) arrow. Relative to the network prevailing during the pre-break subsample in the left graph, the network is now much sparser. The average absolute value of the significant c_3s in the post-break sample is 0.126, versus 0.719 in the pre-break period. The number of significant relationships from lagged news to future markets is now 9, from 29 previously. Post-break, markets become much less responsive to lagged news. As can be seen from Tables 3 and 4, there are only two markets–news pairs with reversals in the post-break subsample, compared to 17 in the early subsample. Diminished market responsiveness to news results in many fewer reversals.

Graph A of Figure 8 shows the results from estimating (10) in the post-Mar. 15, 2020 subsample with news as the response and lagged markets as the test variables. Table A17 in the Supplementary Material details the estimation results. The Granger causality network from day t markets to day $t+1$ news is more sparse than its early subsample counterpart in Graph A of Figure 8. There are now 34 significant c_3 coefficients, whose mean absolute value is 0.240; this compares to 49 significant coefficients with a mean absolute value of 0.403 in the early subsample. Post-break, markets and news become less tightly coupled. The more pronounced response of future news to lagged markets in the pre-break period complements Garcia's (2018) finding that *Wall Street Journal* and *New York Times* news respond more to negative lagged market returns than to positive lagged returns; my results suggest this relationship has important time variation.

VI. Interpretation

The key assumption underlying the narrativity hypothesis is that investors respond to elements of media coverage based solely on narrative content rather than on factual information. Unless investors are completely irrational, narrative content can impact risk premia (e.g., by making investors more risk averse due to a very colorful Great Depression analogy), but narrative, non-factual news should not cause (even partially) rational investors to systematically revise their cash flow expectations. This suggests that during times of high narrativity, asset price fluctuations should be driven more by discount rate, rather than cash flow, news. A finding that cash flow news is the dominant driver of price fluctuations would support the rational information hypothesis.

A variance decomposition (Campbell and Shiller (1988), Campbell (1991)) can thus serve as a sanity check for the narrativity hypothesis. I run this analysis using returns of ETFs that track my 5 asset classes: S&P 500, the VIX index, high-yield bonds, and short- and medium-dated Treasuries.[24] I use daily data from Jan. 17, 2020 to Dec. 31, 2020. The return forecasting variables include lagged returns, dividend yields, the VIX index, interest rates, and the FACTOR and FACTOR_-ALL series. Details are in Section A2.2 of the Supplementary Material. Of the 5 asset classes, only 2-year Treasuries (whose dividends are highly predictable) have variation in cash flow beliefs that exceeds the variation in beliefs about discount rates. For the other assets classes, discount rate belief variation is higher than the variation of cash flow beliefs, and for the S&P 500 and HY ETFs, this ratio is roughly 7:1. Across the major asset classes, time variation in risk premia played a crucial role during the COVID-19 market crisis.[25]

The structural break tests in Figure 9 suggest the markets–news relationships prevailing in the early part of the crisis changed abruptly in the middle of March, with most markets–news pairs experiencing breaks in the 2-week period from Mar. 9 to Mar. 23. The majority of breaks occur just after the Fed's emergency rate cut on Sunday, Mar. 15 or just after its removal of quantitative guidance on Mar. 22. These actions' immediate impact on financial markets was likely through the discount rate channel, since cash flows were not directly affected or targeted.[26]

Another important event in Mar. 2022 was the introduction of the Coronavirus Aid, Relief, and Economic Security (CARES) Act, which cleared the Senate on

[24]The standard variance decomposition starts with the log-linearization of the return equation $R_{t+1} = (P_{t+1} + D_{t+1})/P(t) = (P_{t+1}/D_{t+1} + 1) \times D_{t+1}/D_t \times D_t/P_t$ around the long run mean of $\log D_t/P_t$. The decomposition of returns into changes in beliefs about cash flows and discount rates follows directly from this approximation. This decomposition can be applied any long-lived security which pays dividends. Fixed income ETFs which maintain a fixed duration profile, or VIX ETFs that roll VIX futures, fit the model as much as stock ETFs do. More bond-specific decompositions are possible, e.g., Campbell and Ammer (1993), but these are not needed for the present analysis.

[25]There are numerous models with rational investors, e.g., habit, long-run risk, etc., capable of generating large time-variation in discount rates. See Campbell (2018) for an overview.

[26]See Bhattarai and Neely (2022) for a survey of the channels of unconventional monetary policy.

https://doi.org/10.1017/S002210902300131X Published online by Cambridge University Press

Mar. 25 and was signed into law by the president on Mar. 27. The CARES act was provided fiscal stimulus by funneling \$2.2 trillion to individual consumers, to small and large corporations, and to local governments. As such, it likely impacted investor beliefs about firms' near-term cash flows. The timing of break dates soon after the Mar. 15 and Mar. 22 Fed announcements suggests that it was the Fed's discount rate related action that triggered a transition from the more- to the less-tightly coupled regime. Though a clean separation of the break causes is challenging (the passage of CARES may have been anticipated) the break timing supports the variance decomposition finding that major asset class returns were driven by changing discount rate, as opposed to cash flow, beliefs.[27]

A. Hypothesis Tests: Early Subsample

To test Predictions 1–5 of Section I.A, I exploit the cross section of asset returns and study the combined news–markets relationships of Sections IV and V. The workhorse model for hypothesis testing is

$$O_{m,k} = a_m + b_{\text{Fed}} \times \text{FED}_k + b_{1987} \times 1987_k + \varepsilon_{m,k}, \tag{13}$$

where $O_{m,k}$ is an outcome variable for the market m and news topic k relationship, a_m is an asset-class-level fixed effect, which allows for the possibility that asset classes have systematically different levels of the outcome variable, and FED_k and 1987_k are scaled textual similarity measures. FED_k equals 1 for the news topic whose Euclidean distance, defined in (4) and shown in Table 2, is the lowest relative to the Fed corpus and equals 1/12 for the topic whose Euclidean distance to the Fed corpus is the highest. 1987_k is similarly defined relative to the 1987 corpus. Because both the Fed and 1987 similarity measures are included in (13), there is no need to use the residuals from (5) in this specification. With 12 topics and 5 markets, the regression in (13) has 60 observations.

Prediction 1 states that high narrativity topics should be systematically associated with more frequent contemporaneous market responses than low narrativity (more factual) topics. To test this, I set $O_{m,k}$ in (13) to equal 1 if a given markets–news pair is characterized by a significant b_3 coefficient in (8) (Table 3 summarizes the results of these regressions), and to 0 otherwise. The CONT-E column of Table 5 shows this analysis for the pre-break part of the sample. The topic closest to the 1987 corpus is 75% (11/12 × 81.5%) more likely to be characterized by a significant, contemporaneous markets–news relationship relative to the topic that is furthest away from the 1987 corpus. The effect is highly significant. Similarly, the topic closest to the Fed corpus is 11/12 × 58.9% *less* likely to be characterized by a significant markets–news relationship in (8), though this effect is significant only at the 11.5% level.

Prediction 2 holds that the market impact of high narrativity topics should depend on the level of investor attention (proxied for by the lagged, average VIX) but this effect should be less pronounced for low narrativity topics. To test this, I set $O_{m,k}$ in (13) to 1 if a given news–markets pair in (8) is characterized by hypersensitivity, that is, b_3 and b_4 are significant and have the same sign. As column HYP-E

[27] I thank Jennifer Conrad, the editor, for pointing out this interpretation.

TABLE 5

Information Tests for 12-Topic, End-of-April Model

Table 5 reports results for the specification in (13), which regresses markets–news pair outcomes on a market fixed effect and on the Fed and 1987 narrativity measures for a given topic (defined in Section III.B). These results use the 12-topic, end-of-April model. The left part of the table (–E) corresponds to the early (pre-break) subsample. The right part of the table (–L) corresponds to the late (post-break) subsample. The columns in each panel correspond to the following outcome variables: CONT is an indicator variable for whether a given markets–news pair exhibits a significant contemporaneous relationship (b_3 from equation (8) is significant); HYP is an indicator for hypersensitivity (b_3 and b_4 in equation (8) are both significant and have the same sign); REV is an indicator variable for reversals (b_3 from equation (8) and c_3 from equation (10) are significant and have opposite signs); GRNG is a measure of markets–news Granger connectivity (equal to 2 if both the news series and the market series Granger cause each other, to 1 if only one Granger causes the other, and 0 otherwise); and M2T is set to 1 if the market series Granger causes the news series. Significance is defined at the 10% level or better. Standard errors are clustered by market. *P*-values are shown in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	CONT-E	HYP-E	REV-E	GRNG-E	M2T-E	CONT-L	HYP-L	REV-L	GRNG-L	M2T-L
1987	0.815***	0.834***	0.115	1.493***	1.063***	0.747**	0.589**	−0.079	0.268	0.198
	(0.000)	(0.000)	(0.270)	(0.000)	(0.000)	(0.012)	(0.020)	(0.394)	(0.367)	(0.195)
FED	−0.589	−0.751**	−0.198	−1.133***	−0.476***	−0.294*	−0.136	−0.031	0.041	0.152*
	(0.115)	(0.044)	(0.285)	(0.000)	(0.000)	(0.052)	(0.381)	(0.396)	(0.892)	(0.098)
HYP			0.260**					0.129		
			(0.038)					(0.287)		
No. of obs.	60	60	60	60	60	60	60	60	60	60
R^2	0.135	0.174	0.192	0.278	0.245	0.264	0.188	0.064	0.027	0.062

of Table 5 shows, hypersensitivity is strongly positively associated with narrative proximity to the 1987 corpus ($b_{1987} = 0.834$), and strongly negatively associated with narrative proximity to the Fed corpus ($b_{\text{FED}} = -0.751$). These effects are large: the difference in the probability of hypersensitivity increases by $11/12 \times b_{1987}$ when going from the least- to the most-proximate topic to the 1987 corpus; the decrease for the most Fed-proximate to the least Fed-proximate topic is $11/12 \times b_{\text{FED}}$.

The REV-E column of Table 5 tests Prediction 3, which states that hypersensitive markets–news pairs should be systematically associated with reversals. A reversal is defined as a significant b_3 coefficient in (8) and a significant c_3 coefficient in (10) which has the opposite sign to b_3. $O_{m,k}$ in (13) is thus set to 1 for markets–news pairs characterized by reversals. The presence of reversals is not associated with narrative proximity to either the Fed or the 1987 corpus, but is positively and statistically significantly associated with hypersensitivity of a particular markets–news pair $\{m,k\}$. The presence of hypersensitivity increases the probability of reversals by 26%, an economically large impact.

Prediction 4 maintains that high narrativity news series will be more highly connected in the Granger causality networks shown in Figures 7 and 8. A test variable is said to Granger cause a response variable if the c_3 coefficient in (10) is significant. The outcome variable $O_{m,k}$ for the Granger connectivity tests equals 2 if there is a topic to market ($k \rightarrow m$) link in the news to markets (Figure 7) network and an $m \rightarrow k$ link in the markets to news (Figure 8) network, equals 1 if there is only one such link, and is 0 otherwise. The results of this test for the early subsample are in the GRNG-E column of Table 5. As predicted by the narrativity hypothesis, markets–news pairs where the news topic is closer to the 1987 corpus are much more likely to be linked in the Granger causality network, and pairs where the news topic is closer to the Fed corpus are much less likely to be connected. Both effects are highly statistically significant. The M2T-E column sets the outcome variable to 1 if there is an $m \rightarrow k$ (markets to topic) link, and 0 otherwise. The b_{1987} and b_{FED}

coefficients in this version of (13) are statistically significant and of the same sign as in the Granger connectivity test in GRNG-E, but the magnitude of the coefficients is smaller. This suggests that a portion of the Granger connectivity effect comes form news to markets linkages, and a portion comes from markets to news linkages. In both cases, more highly narrative topics are more tightly connected.

To summarize, in the early subsample, there is strong evidence in favor of the narrativity hypothesis (Predictions 1–4): more narrative topics (more 1987- and less Fed-like) are more associated with contemporaneous relationships between markets and news, are more often associated with hypersensitivity, which is more often associated with reversals, and are more tightly linked in the Granger causality networks of Figures 7 and 8. This evidence is consistent with the narrativity and feedback hypotheses of Goetzmann et al. (2022), the availability heuristic of Tversky and Kahneman (1974), and the availability cascade hypothesis of Kuran and Sunstein (1999).

B. Hypothesis Tests: Late Subsample

The *-l* columns of Table 5 show that the early-sample results are either much weaker, or fully absent, in the post-break period. There is some weak evidence in favor of Predictions 1 (contemporaneous relationship) and 2 (hypersensitivity). The hypersensitivity–reversal relationship (Prediction 3) disappears. And the impact of narrativity on the connectivity of the Granger network (Prediction 4) is no longer there. The post-break evidence (with much weaker narrativity and feedback characteristics) is thus more consistent with the rational information hypothesis (Prediction 5).

VII. Robustness

Section A3 of the Supplementary Material contains many robustness checks. Section A3.1 of the Supplementary Material argues that endogeneity of news is not problematic in (8). Section A3.2 of the Supplementary Material shows the article's results are robust to other choices of break dates besides Mar. 15, 2020. Section A3.3 of the Supplementary Material shows that the break dates identified using the full sample could have been identified as early as Apr. 2020, and would have stayed roughly constant as more information came in during the rest of the year. Section A3.4 of the Supplementary Material shows that the results are robust to using a topic model estimated using all data in 2020 instead of only the articles as of April. This full sample topic model contains 24 topics. Importantly, Table A31 in the Supplementary Material shows the hypothesis test results of Section VI are robust when using the 24-topic model of Section A3.4 of the Supplementary Material, in which case the regressions in (13) have 120 observations.

VIII. Conclusion

The early part of the COVID-19 pandemic was characterized by negative feedback between news and markets. Markets reacted more to high narrativity news than to low narrativity news, and high narrativity news topics were more frequently associated with hypersensitivity, which in turn was systematically related to reversals. Markets Granger caused news, and news Granger caused market responses. High narrativity news topics were more likely to be present in the markets–news Granger causality network. The markets–news feedback loop was broken around the time of the Fed's market interventions in March of 2020. Subsequent to this break, news and markets became considerably less coupled and news hypersensitivity became much less pronounced. Markets and news started to behave more in line with the rational information hypothesis in the post-break subsample. The stark difference in the markets–news relationship in the pre- and post-break periods is one of the key results of this article.

My finding of news–markets feedback, and of the impact of news narrativity, is novel to the COVID-19 literature. It further raises the questions of whether past crises (the crash of 1987, the dot-com crash of 2000–2001, the Global Financial Crisis of 2008–2009, or the European sovereign debt crisis) were also characterized by news–markets feedback and of the role that narrativity played. Are feedback loops a hallmark of market crises in general? Is such feedback largely absent during non-crisis times? Do regime break tests focused on the markets–news relationship help identify crisis to non-crisis transition points in financial markets in real time? What is the role of news narrativity in such crisis episodes? My finding that the effect of economic narratives was less pronounced post-break raises the question of whether such conditionality exists in other crisis periods. I hope the framework introduced in this article will be useful for answering these questions.

Supplementary Material

To view supplementary material for this article, please visit http://doi.org/10.1017/S002210902300131X.

References

Andrews, D. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica*, 61 (1993), 821–856.

Andrews, D. "Tests for Parameter Instability and Structural Change with Unknown Change Point: A Corrigendum." *Econometrica*, 71 (2003), 395–397.

Arteaga-Garavito, M.; M. Croce; P. Farroni; and I. Wolfskeil. "When the Markets Get CO.V.I.D: COntagion, Viruses, and Information Diffusion." CEPR Discussion Papers 14674 (2021).

Baker, S.; N. Bloom; S. Davis; K. Kost; M. Sammon; and T. Viratyosin. "The Unprecedented Stock Market Reaction to COVID-19." *Review of Asset Pricing Studies*, 10 (2020), 742–758.

Bhattarai, S., and C. Neely. "An Analysis of the Literature on International Unconventional Monetary Policy." *Journal of Economic Literature*, 60 (2022), 527–597.

Blei, D.; A. Ng; and M. Jordan. "Latent Dirichlet Allocation." *Journal of Machine Learning Research*, 3 (2003), 993–1022.

Boudoukh, J.; R. Feldman; S. Kogan; and M. Richardson. "Information, Trading, and Volatility: Evidence from Firm-Specific News." *Review of Financial Studies*, 32 (2018), 992–1033.

Calomiris, C., and H. Mamaysky. "How News and Its Context Drive Risk and Returns Around the World." *Journal of Financial Economics*, 133 (2019), 299–336.

Campbell, J. Y. "A Variance Decomposition for Stock Returns." *Economic Journal*, 101 (1991), 157–179.

Campbell, J. Y. *Financial Decisions and Markets*. Princeton, NJ: Princeton University Press (2018).

Campbell, J. Y., and J. Ammer. "What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns." *Journal of Finance*, 48 (1993), 3–37.

Campbell, J. Y., and R. Shiller. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1 (1988), 195–228.

Chan, W. "Stock Price Reaction to News and No-News: Drift and Reversal After Headlines." *Journal of Financial Economics*, 70 (2003), 223–260.

Churchill, R., and L. Singh. "The Evolution of Topic Modeling." *ACM Computing Surveys*, 54 (2022), 1–35.

Cochrane, J., and M. Piazzesi. "Bond Risk Premia." *American Economic Review*, 95 (2005), 138–160.

Cohen, L.; C. Malloy; and Q. Nguyen. "Lazy Prices." *Journal of Finance*, 75 (2020), 1371–1414.

Das, S., and M, Chen. "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web." *Management Science*, 53 (2007), 1375–1388.

Davis, S.; S. Hansen; and C. Seminaro-Amez. "Firm-Level Risk Exposures and Stock Returns in the Wake of COVID-19." NBER Working Paper No. 27867 (2020).

Garcia, D. "The Kinks of Financial Journalism." SSRN Working Paper 2515791 (2018).

Garcia, D.; X. Hu; and M. Rohrer. "The Colour of Finance Works." *Journal of Financial Economics*, 147 (2023), 525–549.

Glasserman, P.; F. Li; and H. Mamaysky. "Time Variation in the News-Returns Relationship." *Journal of Financial and Quantitative Analysis*, forthcoming (2024).

Glasserman, P.; H. Mamaysky; and Y. Shen. "Dynamic Information Regimes in Financial Markets." *Management Science*, forthcoming (2024).

Goetzmann, W.; D. Kim; and R. Shiller. "Crash Narratives." NBER Working Paper No. 30195 (2022).

Gormsen, N., and R. Koijen. "Coronavirus: Impact on Stock Prices and Growth Expectations." *Review of Asset Pricing Studies*, 10 (2020), 574–597.

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

Hartley, J., and A. Rebucci. "An Event Study of COVID-19 Central Bank Quantitative Easing in Advanced and Emerging Economies." Working Paper No. 27339, NBER (2020).

Hassan, T.; S. Hollander; L. van Lent; M. Schwedeler; and A. Tahoun. "Firm-Level Exposure to Epidemic Disease: COVID-19, SARS, and H1N1." NBER Working Paper No. 26971 (2021).

Hoberg, G., and G. Phillips. "Text-Based Network Industries and Endogenous Product Differentiation." *Journal of Political Economy*, 124 (2016), 1423–1465.

Jegadeesh, N. "Evidence of Predictable Behavior of Security Returns." *Journal of Finance*, 45 (1990), 881–898.

Karavias, Y.; P. Narayan; and J. Westerlund. "Structural Breaks in Interactive Effects Panels and the Stock Market Reaction to COVID-19." *Journal of Business & Economic Statistics*, 41 (2023), 653–666.

Ke, S.; J. L. Montiel Olea; and J. Nesbit. "Robust Machine Learning Algorithms for Text Analysis." Working Paper, available at https://jmcgnesbit.com/files/projects/rmlata/rmlata_paper.pdf (2021).

Ke, Z.; B. Kelly; and D. Xiu. "Predicting Returns with Text Data." SSRN Working Paper 3389884 (2022).

Kuran, T., and C. Sunstein. "Availability Cascades and Risk Regulation." *Stanford Law Review*, 51 (1999), 683–768.

Lehmann, B. "Fads, Martingales, and Market Efficiency." *Quarterly Journal of Economics*, 105 (1990), 1–28.

Li, K.; X. Liu; F. Mai; and T. Zhang. "The Role of Corporate Culture in Bad Times: Evidence from the COVID-19 Pandemic." *Journal of Financial and Quantitative Analysis*, 56 (2021), 2545–2583.

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

Newman, D.; J. Han Lau; K. Grieser; and T. Baldwin. "*Automatic Evaluation of Topic Coherence.*" *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL*, (2010), 100–108.

Ramelli, S., and A. Wagner. "Feverish Stock Price Reactions to COVID-19." *Review of Corporate Finance Studies*, 9 (2020), 622–655.

Shiller, R. *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*. Princeton, NJ: Princeton University Press (2019).

Steyvers, M., and T. Griffiths. "Probabilistic Topic Models." In *Handbook of Latent Semantic Analysis*, T. K. Landauer, D. S. McNamara, S. Dennis, and W. Kintsch, eds. Mahwah, NJ: Lawrence Erlbaum Associates Publishers (2007), 427–448.

Tetlock, P. "Does Public Financial News Resolve Asymmetric Information?" *Review of Financial Studies*, 23 (2010), 3520–3557.

Tetlock, P. "All the News that's Fit to Reprint: Do Investors React to Stale Information?" *Review of Financial Studies*, 24 (2011), 1481–1512.

Tetlock, P.; M. Saar-Tsechansky; and S. Macskassy. "More than Words: Quantifying Language to Measure Firms' Fundamentals." *Journal of Finance*, 63 (2008), 1437–1467.

Tversky, A., and D. Kahneman. "Judgment Under Uncertainty: Heuristics and Biases." *Science*, 185 (1974), 1124–1131.