



# COVID risk narratives: a computational linguistic approach to the econometric identification of narrative risk during a pandemic

Yuting Chen<sup>1,2</sup> · Don Bredin<sup>1</sup> · Valerio Poti<sup>1</sup> · Roman Matkovskyy<sup>2</sup>

Received: 5 November 2021 / Accepted: 15 November 2021 / Published online: 29 November 2021  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

## Abstract

In this paper, we study the role of narratives in stock markets with a particular focus on the relationship with the ongoing COVID-19 pandemic. The pandemic represents a natural setting for the development of viral financial market narratives. We thus treat the pandemic as a natural experiment on the relation between prevailing narratives and financial markets. We adopt natural language processing (NLP) on financial news to characterize the evolution of important narratives. Doing so, we reduce the high-dimensional narrative information to few interpretable and important features while avoiding over-fitting. In addition to the common features, we consider virality as a novel feature of narratives, inspired by Shiller (Am Econ Rev 107:967–1004, 2017). Our aim is to establish whether the prevailing narratives drive or are driven by stock market conditions. Focusing on the coronavirus narratives, we document some stylized facts about its evolution around a severe event-driven stock market decline. We find the pandemic-relevant narratives are influenced by stock market conditions and act as a cellar for brewing a perennial economic narrative. We successfully identified a perennial risk narrative, whose shock is followed by a severe market drop and a long-term increase of market volatility. In the out-of-sample test, this narrative went viral since the start of the global COVID-19 pandemic, when the pandemic-relevant narratives dominate news media, show negative sentiment and were more linked to “crisis” context. Our findings encourage the use of narratives to evaluate long-term market conditions and to early warn event-driven severe market declines.

**Keywords** Narrative economics · COVID-19 · Natural language processing · Tone analysis · Early warning indicator

---

The authors gratefully acknowledge the support of Science Foundation Ireland under grant number 16/SPP/3347 and the support of H2020 COST Action CA19130, “Fintech and Artificial Intelligence in Finance—Towards a transparent financial industry”.

---

✉ Yuting Chen  
yuting.chen@ucd.ie

Extended author information available on the last page of the article

**JEL Classification** C53 · D81 · E37 · E71 · G17 · G41

## 1 Introduction

Based on the idea that “significant market events generally occur when there is similar thinking among large groups of people” (Shiller, 2015, p 101), and such similar thinking can be established or propagated via the spread of viral narratives, Shiller (2019) sheds light on the understudied area of “narrative economics”. Narratives were studied in computer science and linguistics far before than in economics and finance. Thus, narrative economics naturally draws on the established computer science/linguistics literature. Understanding the mechanisms of popular stories that influence individual and collective economic behavior can contribute to anticipating, preparing for and mitigating the damage of financial crises, recessions and other major economic events. Recent narrative papers in economics and finance started formalizing the research methods [see a survey by Grimmer and Stewart (2013)], theoretical grounds [e.g., Eliaz and Spiegler (2020)] and stylized facts [e.g., Bertsch et al. (2021) and Nyman et al. (2021)].

In 2020, along with the spread of COVID-19, coronavirus narratives went viral internationally. Since the world-wide outbreak of the pandemic (started in January 2020), “COVID-19” dominated media reports and popular commentary. The narratives soon became important, with the potential to impact aggregate financial markets. We find the presence of prevailing pessimistic, uncertain and competing narratives especially during the first few months of the pandemic (some examples are available in Appendix 2). During the same period, global financial markets (e.g., S&P 500 Index) dropped heavily and became unprecedentedly volatile. Such narratives might act an important role in the dramatic market movements. It was unclear, however, whether the spread of panic via narratives drives the market declines, or it simply reflects the past declines.

The impact of a narrative could be high as long as it went viral, regardless of its validity, that is, truth is not enough to stop viral false narratives to impact human beliefs and change their behaviors (Shiller 2019). Not only authentic information, but also rumors and misinformation regarding to COVID-19 went viral and became influential (see WHO Novel coronavirus situation report 13).<sup>1</sup> During a period with high uncertainty and large volume of blended true and false information, would the financial markets be more sensitive to narratives?

It is thus helpful to study the popular narratives and understand whether the viral narratives impact market moves during an exogenous turmoil period. This paper contributes to the “narrative economics” literature by dissecting important economic narratives and examining their interactions with the stock market. We will address two related research questions. First, do the COVID narratives drive the severe stock

---

<sup>1</sup> A new term, “infodemics”, was coined by World Health Organization (WHO), to describe the phenomenon that widely spread untrue information regarding the pandemic causes a significant portion of the population to change their behaviors, in a harmful way.

market movements? Second, are there perennial economic narratives during the pandemic that drive down stock market values and drive up volatility?

By answering the first question, we document some stylized facts about the COVID narratives circulated in the financial market, and examine whether they satisfy the necessary conditions for the narratives to drive the stock markets. We rely on multiple well-developed natural language processing techniques and procedures to extract interpretable features of narratives. First, we define “narratives” as a collection of stories with a specific topic, but we also examine whether some important narratives could be described as causal models mapping actions into consequences (Eliasz and Spiegler 2020).<sup>2</sup>

Perfectly identifying specific narratives is a technically challenging task. Given that we study the narratives via their interpretable features, however, we can bypass the step of individual narrative identification. Instead, we identify the articles with the presence of narratives of interest. The textual properties of the articles are good proxies of the properties of the narratives. Based on common “codebook functions” [a term borrowed from Egami et al. (2018)], we summarize the texts to a lower dimensional space. The narratives are located initially by both data collection and then by NLP tools or procedures. More specific narratives can be located with more constraints in data collection and with more complicated NLP process. The common features we extract using textual analysis include narrative amount (referred to as “popularity” or “intensity”), textual sentiment (Loughran and McDonald 2011), topic distribution (Blei et al. 2003), topic consensus [similar with Nyman et al. (2021)] and semantic meaning (Mikolov et al. 2013). To understand the virality of narratives, we also employ a novel virality measure based on a basic epidemiology model (Kermack and McKendrick 1927).

Our narrative measures show that, around the severe market declines, the pandemic-relevant narratives went viral and dominant, fragment into competing topics, became more pessimistic and uncertain, and were more linked to the “recession” and “crisis” context. Our VAR-based analysis, however, finds no evidence that the coronavirus narratives circulating in financial markets Granger-cause stock market conditions. Therefore, the pandemic-relevant narratives only reflect past performance.

To answer the second question, first, we creatively use Latent Dirichlet Allocation (Blei et al. 2003) to identify recognizable important economic narratives over time. The LDA algorithm is trained on general market news, instead of coronavirus news, to capture potential long-lasting common narratives. All news are collected from LexisNexis. Our method for filtering out important narratives is inspired by Baker et al. (2016) and Nyman et al. (2021), who rely on the correlation of their time series measures with standard economic and financial measures to gauge the robustness of their metrics. A high correlation indicates a high relevance of their measures to economy or finance. In our paper, we move this step to the upper stream of the workflow to help identify important LDA topics. Specifically, we rank the topics based on the correlation between the their topic intensity with the CBOE Volatility

---

<sup>2</sup> The degree of freedom of the broadness of a topic is large. A small group of narratives could be very specific, while larger group narratives might be a collection of narratives with a general topic.

Index (VIX), and we consider the topics with the highest ranking as important. The logic is that a topic highly correlated with implied volatility should necessarily contain narratives that are somehow relevant to the market's perception of risk. As we identify the important topic, we calculate its topic intensity overtime. We find that the topic became viral around the market drop in February 2020 as they did in previous historical crisis periods, indicating the existence of perennial economic narratives that potentially play a role in financial markets. The narratives are manually examined via their representative texts and named risk narratives.

With the perennial risk narratives identified, we investigate whether the narratives reflect investor sentiment or convey fundamental information or neither, along the lines of Tetlock (2007). Our results show that this perennial risk narratives Granger-cause stock market returns and volatility in the long term, implying that they convey new information regarding to both stock market value and risk.

The remainder of the paper is structured as follows. Section two reviews the key literature, with focuses on narratives and NLP methods. Section three outlines our methodologies and section four introduces our data. Section five reports the results and discussions. The final section concludes.

## 2 Literature review

With years of pioneering contribution, Robert Shiller (2017) officially proposes the term *Narrative Economics* to refer to “the study of the spread and dynamics of popular narratives, the stories, particularly those of human interest and emotion, and how these change through time, to understand economic fluctuations“(Shiller, 2017, p. 967). Under the framework of narrative economics, the central question is how popular stories change through time affect economic outcomes. Based on many stylized facts from the history, Shiller (2019) claims that viral economic stories move markets and affect economic outcomes through impacting individual and collective decisions. The hypothesis is that ideas embedded in popular narratives influence or co-influence outcomes via changing individual or collective economic behaviors. Shiller emphasizes the effect of narrative contagion on economic events, which is a novel dimension of narratives in addition to the common ones NLP can detect. He also proposes the use of epidemiological models to examine the spread of narratives.

There are some recent developments of the narrative economics literature such as Eliaz and Spiegler (2020), Nyman et al. (2021), Larsen et al. (2021) and Bertsch et al. (2021). Shiller (2019) does not provide explicit models for the role of narratives in economic contexts, but Eliaz and Spiegler (2020) use directed acyclic graph (DAGs) to describe a narrative as a causal model that maps actions into consequences. Larsen et al. (2021) investigates the role of media narratives in the formation process of inflation expectations, and he finds that the news topics in media are good predictors of inflation expectations. Nyman et al. (2021) study the role of narrative sentiment in financial system developments and find a formation of exuberance prior to the global financial crisis. Bertsch et al. (2021) characterize the properties of narratives over business cycles and find that narratives tend to consolidate

during expansion periods and tend to fragment during contractions. All the empirical analysis on narratives mentioned here rely on natural language processing tools.

It worth noting that when using “narrative”, researchers have different definitions. Shiller (2019) provides a general definition of “narrative”<sup>3</sup> and highlights that “narratives” could be referring specific stories with actionable intelligence, or simply referring articles with certain keywords or terms. Nyman et al. (2021) also define a narrative as a collection of articles. Larsen et al. (2021) and ter Ellen et al. (2019) treat topics as narratives. By contrast, Eliaz and Spiegler (2020) define a narrative as a causal model that maps actions into consequences.

Although the literature strand of narrative economics was formally initiated only recently, narratives were studied in computer science and linguistics fields far before its applications in economics and finance. A richer and older literature of narratives studying different angles (dimensions) of various types of narratives naturally nourish the narrative economics literature. As textual data are in fact the container of narratives, many research in textual analysis are also narrative studies. The literature covers methodologies and applications of common processing on narratives, including simple counts [such as Fang and Peress (2009) and Baker et al. (2016)], textual sentiment analysis [such as Tetlock (2007), Loughran and McDonald (2011) and Nyman et al. (2021)], topic modeling [such as Blei et al. (2003), Bertsch et al. (2021) and Larsen et al. (2021)], semantic meaning analysis [such as Mikolov et al. (2013) and Reimers and Gurevych (2019)] and commonsense understanding [such as Mostafazadeh et al. (2016)]. Each of the processing methods can generate some characteristics of the narratives in a certain feature dimensions. As claimed by Egami et al. (2018), there is no single correct “codebook function” that compress high-dimensional, complicated and sparse text to a low-dimensional measure, only if the functions can be interpretable, tractable and produce measures with high fidelity and of theoretical interest.

Textual sentiment is one of the most popular measures been widely used in finance research [see a survey by Kearney and Liu (2014)]. Among many financial textual sentiment analysis methods, one state-of-the-art method based on bag-of-words (dictionary-based) approach is the Loughran and McDonald (2011) word lists (LM, henceforth). The portion of the words belonging to a certain word list proxies the intensity of textual sentiment/tone that word list is representing. Their negative word list is one of the most widely used lists for detecting positive/negative tone in financial texts. With the same fashion of the bag-of-words method, Baker et al. (2016) created the economic policy uncertainty (EPU) index based on the frequency of articles containing words in a short word list. In this case, the word list represents economy and policy-specific risk sentiment. Recent literature using classic bag-of-words methods include Liu and Han (2020) and Ardia et al. (2019).

Useful features other than sentiment can be extracted, too, thanks to the development of multiple NLP tools, especially those rely on machine learning. Mikolov

---

<sup>3</sup> “The word narrative is often synonymous with story. But my use of the term reflects a particular modern meaning given in the Oxford English Dictionary: “a story or representation used to give an explanatory or justificatory account of a society, period, etc.”, [p xi, Shiller (2019)].

et al. (2013) propose word embedding to vectorize tokenized words. Through unsupervised learning from a large amount of textual documents, word embedding reduces the dimensionality of word tokens, according to their contextual information. Each word token will be represented by a vector with a fixed number of dimensions. Semantic similarity between any two words can be calculated using cosine similarity, that is, the cosine of the angle between two vectors in the semantic space. With the trained vectors, one can conduct many tasks such as sentiment analysis and topic classifications. Latent Dirichlet Allocation, proposed by Blei et al. (2003), is also an unsupervised learning algorithm based on the words–document relationship.<sup>4</sup> LDA can train an algorithm to estimate the probability distribution among  $K$  topics for a piece of textual content. It identifies latent clusters of contents, which could be interpreted as meaningful topics after careful manual examination.

In addition to the state-of-the-art NLP methods, Shiller (2019) proposes the use of epidemiological models to study the spread of narratives. The purpose of this application is to measure the contagion parameters (as in epidemic studies) and forecast the narrative spread. Cinelli et al. (2020) fit the spreading of COVID-19 narratives on social medias with a basic epidemiological model to characterize the evolution. The model requires only a time series of the narrative intensity index as the input. There are other methods of modeling information contagion developed in the information diffusion literature [such as Romero et al. (2011)], but time series measures of virality using those methods are not well established.

Epidemiologists have long pursued multiple mathematical avenues to model the spread of disease. One classic model, which Shiller (2019) suggests and Cinelli et al. (2020) borrow, to describe the spread of narratives, is the basic Kermack–McKendrick Susceptible–Infective–Recovered (SIR) model (Kermack and McKendrick 1927). A key hypothesis is that a narrative epidemic resembles a disease epidemic in the contagion and recovery dynamics. More details of the basic SIR model can be found in Appendix 1.

In testing the role of market narratives, in the form of media news, Tetlock (2007) is influential, with his focus on the textual sentiment feature. Tetlock (2007) measures the intertemporal links between news tones and the stock market using basic vector auto-regressions. The hypothesis under examination is that high media pessimism is associated with low investor sentiment, resulting in downward pressure on prices with a reversion in the long term. Alternatively, it is possible that news pessimism is a proxy of negative information about the fundamental asset value, that has not yet been priced in, hence, predicting price decrease without reversion. Another alternative hypothesis is that the pessimistic media tone does not contain any new information, implying no return predictability. Only the textual sentiment dimension of market narratives was examined in this research.

---

<sup>4</sup>  $K$  stands for the number of topics one wants the model to detect.

### 3 Methodology

This section describes the details of our identification method for both coronavirus narratives and a perennial risk narrative; narrative feature characterization on intensity, topic consensus, sentiment, semantic meaning and virality, and the methods we use to study the role of the important narratives in the stock market.<sup>5</sup>

#### 3.1 Narrative identification

The first step of our narrative analysis is the identification of narratives. Using a combination of data selection and NLP techniques, we identify the texts containing certain narrative clusters of interest with an reasonable accuracy.<sup>6</sup>

The first narrative cluster we identify is the popular coronavirus narratives relevant to economy and finance. This identification can be achieved by solely data selection. News medias formally produce and deliver stories to capture the public attention. The stories reported on major media channels are highly correlated with those mostly discussed in the society. As argued by Nimark and Pitschner (2019), media news could be viewed as the filtered popular stories by editors. Media news thus naturally satisfies the “popular” property of our target narratives. We rely on LexisNexis, a textual database platform, to search and download the media news. This platform provides convenient searching functions for designing queries with complex conditions and constrains. Using keywords of “COVID-19” and “coronavirus”, the articles containing coronavirus stories can be returned. The risk of over or under identification is low for two reasons. First, the two keywords were rarely used before the coronavirus pandemic, especially in banking and finance news. Second, news from major media tend to report with formal language, the two keywords are always mentioned when the topic is clearly relevant to coronavirus. To limit the collected news to be economy and finance related, we add the “banking and finance” tag in the queries.

The second narrative cluster we identify is popular “market narratives”. We identify them with the same method. Instead of the coronavirus keywords, “market” is used in the searching function. We place the details in the data section.

As the news documents are collected, we tag the texts into more specific categories, namely topic specific narratives. We use LDA (Blei et al., 2003) to gauge their relevance of the latent topics. The most classic LDA algorithm is used for the sake of replicability and comparability. We train the LDA algorithms on both text sets. First, using the “coronavirus narrative” documents, we identify latent topics that are idiosyncratic to coronavirus. Second, using the “market narrative” documents from 1980 to 2019, we aim for a successful identification of common economic narratives that are constantly important in the history. After running this unsupervised learning algorithm, we have a number of latent topics

<sup>5</sup> Project codes are available at: [https://github.com/YTRoBot/COVID\\_Risk\\_Narratives\\_for\\_DFIN](https://github.com/YTRoBot/COVID_Risk_Narratives_for_DFIN).

<sup>6</sup> “Topics” is one way of narrative clustering.

represented by bags of highest probability words with different weights of relative importance.

With some assistance of human intelligence, we recognize the narratives that the latent topics should represent. Viewing the bags of words gives us an intuition about the latent topics. However, like many other machine learning classifiers, LDA cannot guarantee interpretability of the detected clusters. When the representative words cannot provide a clear indication of an interpretable topic, we read a selection of documents with the highest probability of this topic to better understand of the covered narratives. Unlike common classification algorithms (Decision Trees, k-Nearest Neighbour Classifier, etc.), which allocate an object to a single category or multiple categories with certainty, LDA assigns a probability of 100% to a set of topics for each document. That means one document could be treated as a mix of a set of topics with different weights. We aggregate the topic probabilities across documents on a daily, weekly or monthly basis as a measure of topic prevalence (we formally name it as narrative intensity in the next subsection), and its trends and patterns are used to assist the interpretation.

Given our goal of identifying economic narratives that might interact with market events, we check the correlation between the narrative measures and market conditions, with a focus around significant market events such as recessions, expansions, bull and bear markets. For the historical “market narratives”, we calculate the correlation between the topic prevalence series and the VIX index. A high correlation with the implied volatility index indicates that the latent topic represents narratives regarding market volatility and economy uncertainty. We rank the topics by this correlation and investigate the top ones.

For robustness-check, we run the same analysis with a list of  $K$ s (from 10 to 90 with an interval of 10), that is, the number of topics detected by LDA. A larger  $K$  can generally result in more specific latent topics, it can also lead to too fine topics (Nimark & Pitschner, 2019). Although there are some measures, such as coherence values, developed for the purpose of LDA “performance evaluation”, no consensus was achieved about to what degree should we rely on such measures. Topic coherence proposed by Mimno et al. (2011) is one of the measures used for  $K$  tuning. Such measures, however, cannot guarantee that the winner has the true number of topics [see Roberts et al. (2019)]. After all, there are infinite criteria for text classification or clustering, the LDA algorithm is specifically a pure data-driven method, based on the word-document relationship. As there is no correct answer for the value of  $K$ , it is feasible to set the number of topics arbitrarily with a reasonable value, depending on the purpose of analysis. For example, Nimark and Pitschner (2019) arbitrarily set  $K$  as 10. In this paper, we aim to identify interesting long-lasting latent topics. The “ $K$ -tuning” is not a priority. We only need a random selection of latent topics to be well-identified, that is, interpretable and stable. In our application, we train the LDA models with  $K$  ranging from 10 to 90 with an interval of 10. The coherence values are calculated only as a reference, which are close to 0.42 across all 9 models. For any interesting topics

identified and interpreted, to assess the “stability” metric, we cross check their presence and patterns in all LDA models with different  $K$ s.<sup>7</sup>

### 3.2 Narrative features

In the second step, after the narrative identification, is to extract the features from the texts, namely topic consensus, semantic meaning of some key words of interest, narrative intensity, textual sentiment and narrative virality. The features are used to characterize the narratives at different points of time, and represent potential channels through which narratives correlate with the financial markets.

First, under the hypothesis that the sentiment in coronavirus narratives drive the market tumbles, we expect an increase of negative and uncertain tones before the event. Second, according to Nimark and Pitschner (2019), major events shift the news focus and make coverage more homogeneous, we expect that the coronavirus narratives dominate the banking and finance news, with a high narrative intensity measure. Third, based on Shiller (2017), we expect the virality measure of the coronavirus narratives to be high, especially at the early stage. Bertsch et al. (2021) find that narratives tend to fragment into competing explanations during contractions. Thus, fourth, we expect the narrative consensus to decrease around the severe market declines. Lastly, under the hypothesis that narratives contribute to the stock market declines via linking the pandemic with a possible recession, we expect the semantic relationship between “coronavirus” and “recession” to be high around the market collapse.

#### 3.2.1 Topic consensus

The literature finds that topic consensus or homogeneity is an informative measure in major market events (Nimark & Pitschner, 2019), business cycles (Bertsch et al., 2021) and financial system developments (Nyman et al., 2021). With the latent narratives identified by LDA and manual inspection, we measure the homogeneity/consensus of narratives in the cross-section of narrative documents. The measure reflects if, at a point of time, some narratives become dominant. Bertsch et al. (2021) find that narratives tend to fragment into competing explanations during contractions and Nyman et al. (2021) find that narratives increase consensus around the strongly positive narrative prior to the crisis. We examine the patterns during the pandemic, which was accompanied by a financial crisis. We follow Nyman et al. (2021) and use Shannon entropy (Shannon, 1948). Although there are variations of consensus measure in the literature [e.g., Nimark and Pitschner (2019), Nyman et al. (2021), and Bertsch et al. (2021)], they are not essentially different for our purpose.

Our entropy of the topic probability distribution is calculated as the logarithmically weighted sum of the topic probabilities,

---

<sup>7</sup> We also test the algorithm with different random seeds to ensure the stability of the important topic.

$$H_d = - \sum_{k=1}^K p_{k,t} \log p_{k,t}, \quad (1)$$

where  $p_{k,t}$  is the average probability of topic  $k$  over all documents published on day  $t$ . The entropy is then lower bounded by zero, which is achieved when all 100% of the topic probability is allocated on one topic. It has an upper bound of  $\log K$ , which is achieved when  $p_k = 1/K$  for all topics. A smaller entropy value indicates a higher narrative consensus.

### 3.2.2 Semantic meaning

Semantic meaning of word tokens is useful not only by itself but also in assisting other NLP tasks. Semantic meaning of a narrative keyword changes when the context of usage is changed. We use the conventional word embedding tool [see Mikolov et al. (2013)], which converts word tokens into vectors and locates them in a high-dimensional space based on their contextual relationships. A common use of embedding is to reveal the contextual relationship between two words, which can be estimated by calculating the cosine similarity between the positions of two words, in the high-dimensional semantic meaning space.

We first examine the semantic meaning of keyword “coronavirus” via its closest words over the pandemic period. We then extract the semantic meaning relationships between the COVID-19 keyword and some keywords of economic narratives. To capture the changes in time, we fit the embedding algorithm on the coronavirus narrative documents on a rolling basis, with a fixed window length of 50 days and a pace of 1 day. With the trained models, we collect the most similar words with “coronavirus” and their cosine similarity at all estimation days to determine the pattern of variation. By plotting the cosine similarity between “coronavirus” with relevant keywords in economic narratives, e.g., “recession” and “crisis”, we can understand how the coronavirus narratives relate with economic narratives at different stages of the pandemic.

### 3.2.3 Narrative intensity

We refer to the popularity of narratives as narrative intensity. It has straightforward meaning and is easy to compute. For example, the widely used macro indicator, “EPU” index (Baker et al., 2016), is an intensity measure of economy and policy uncertainty narratives, calculated by the number of articles. We generate two narrative intensity indices. The first index is the count of news tagged as coronavirus narrative, normalized by the count of banking and finance news. We select this denominator because we limit the narratives of interest to be banking and finance news. Both numbers are counted on the news from LexisNexis. The second index is the aggregate probability of the LDA topic that represents important economic narratives. The aggregate probability of a topic over the documents proxies the public attention on the narratives the topic represent. We have two versions of intensity measures because the narrative identification method are different, that is, our

COVID-19 narratives can be identified in the data collection step, while the economic narratives need to be estimated with the assistance of LDA.

More specifically, we aggregate the document-level topic probabilities by mean or sum on a given date. The following equation shows the construction method of the second narrative intensity index aggregated by mean. The sum version is constructed by removing the denominator.

$$\text{Intensity}_{k,t} = \frac{\sum_{d=1}^{D_t} P_{d,k,t}}{D_t}, \text{ s.t. } \sum_{k=1}^K P_{d,k,t} = 1, \forall d, t. \quad (2)$$

Here,  $\text{Intensity}_{k,t}$  stands for the intensity index of topic  $k$  at time  $t$ , and  $P_{d,k,t}$  represents document  $d$ 's probability for topic  $k$  at time  $t$ . Each document has a single time stamp  $t$ .  $D_t$  refers to the total number of documents with time stamp of  $t$ . With  $K$  topics, we will have  $K$  time series of intensity indices. The  $K$  topics are the LDA output, formatted in a vector of estimated probabilities of topics for an article, with a sum of 100%.

### 3.2.4 Textual sentiment

Textual sentiment is a popular measure that has been employed in economics and finance research with fruitful findings [see Kearney and Liu (2014)]. We extract textual sentiment as an one of the important features of the narratives. We measure textual sentiment using a state-of-the-art bag-of-words tone analysis method.<sup>8</sup> Specifically, we use the Loughran and McDonald (2011) and Bodnaruk et al. (2015) word lists to estimate the tones of the textual contents.<sup>9</sup> We are interested in only the optimism/pessimism and uncertainty dimensions of in this study. The pessimism tone is chosen following Tetlock (2007), and we choose the uncertainty tone because the COVID-19 pandemic causes a high level of uncertainty on economics and financial markets. “negative”, “uncertainty” and “weak modal” word lists are used to estimate the two tones.<sup>10</sup>

The intensity score of each tone is calculated using following equation:

$$\text{Tone}_{s,t} = \frac{\sum_{d=1}^{D_t} \text{Count}_{d,s}}{\sum_{d=1}^{D_t} \text{Count}_{d,\text{total}}}. \quad (3)$$

In this equation,  $\text{Tone}_{s,t}$  represents the score of tone  $s$  on day  $t$ ,  $\text{Count}_{d,s}$  represents the count of words in document  $d$  that are listed in the word list of tone  $s$ ,  $\text{Count}_{d,\text{total}}$  represents the count of words in document  $d$  that are included in the Loughran and

<sup>8</sup> Although there is frequent reference to “sentiment” in NLP applications, most finance studies define investor sentiment as the level of noise traders’ beliefs relative to Bayesian beliefs.

<sup>9</sup> Namely “positive”, “negative”, “uncertainty” and “weak/moderate/strong modal”.

<sup>10</sup> “weak modal” verbs represent the uncertainty of actions.

McDonald (2011) master dictionary, and  $D_t$  refers to the total number of documents at day  $t$ .<sup>11</sup>

We calculate the average tones for each day from the narrative documents. Given that the coronavirus narratives is identified directly from the data collection step, the tones of the documents proxy the tones of the coronavirus narratives.

### 3.2.5 Virality

Drawing on Shiller (2017), we adopt the SIR model (Kermack & McKendrick, 1927) on the narrative intensity series to estimate the latent contagion parameters. A key assumption when applying the model is that the narratives resembles real epidemics in terms of the spreading dynamics. In epidemiology, the susceptible, infected and recovered individuals are always observable variables. In a narrative setting, however, it is not easy nor explicit to distinguish an susceptible individual from those who have “immunity” to a narrative (or a set of narratives). In our study, we can only observe the narrative intensity series, which is correlated with the aggregate “infected” population of narratives. Fortunately, SIR can also be used with only the “infected” population data. By “infected”, we mean that one is interested in a narrative. When a large portion of news articles are discussing a topic, a significant portion of the population is interested in the narratives contained.

We apply the SIR model on the COVID-19 narratives to estimate the contagion and recovery dynamics of the narrative epidemics. The input of the SIR model is the daily intensity measure of the coronavirus narratives. The outputs of the SIR model estimation are the optimized values of  $\beta$  and  $\gamma$ , the contagion rate and recovery rate.<sup>12</sup> The basic reproduction number (denoted as  $R_0$ ), which is calculated as  $\frac{\beta}{\gamma}$ , is a measure of contagion ability. A larger  $R_0$  means a faster spreading of the narrative. By multiplying  $R_0$  by  $S(t)$ , the susceptible population at time  $t$ , we have the effective reproductive number, denoted as  $R_t$ .<sup>13</sup> At the start of an epidemic,  $R_t$  equals  $R_0$  because  $S(0)$  equals one. Along with time, as the susceptible population decrease, the speed of spreading naturally decrease, which is captured by a decreasing  $R_t$ . With the estimated parameters and modeled  $S(t)$ , we calculate the effective reproductive number as our measure of virality. Using the trust region reflective algorithm, we find the parameters that minimize fitting errors.<sup>14</sup> We set lower bounds as 0 and upper bounds as 1 given the nature of the two variables. To prevent the estimated basic reproduction number to be extremely large, for example, when the  $\gamma$  (recovery rate) is estimated to be very small, we further set a lower bound for  $\gamma$  as 0.005, which allows the maximum possible  $R_0$  to be 200.

In the tests with an expanding window, we find that “pseudo real-time” estimates for the parameters are unstable during the early stage of a narrative epidemic. With

<sup>11</sup> The 2018 version from <https://sraf.nd.edu/textual-analysis/code/>.

<sup>12</sup> See Appendix 1 for more details.

<sup>13</sup> To avoid confusion, we denote the ratio of recovered population at time  $t$  as  $R(t)$  and denote the effective reproductive number as  $R_t$ .

<sup>14</sup> See [https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.least\\_squares.html](https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.least_squares.html).

enough data points, during the later stage, the estimated parameters become stable, but it is not necessarily due to convergence. Instead, from our simulation (in Appendix 1), we find that a real time estimation of the parameters will become less sensitive to the variations with longer window. The narrative epidemics, unlike normal epidemics, “re-occur” easily. It can be explained by the fact that people tend to be not immune to a narrative permanently. When certain events happen, a narrative could become contagious again, possibly with new  $\beta$  and  $\gamma$  and a portion of recovered population become susceptible again. Modification to account for this feature is expected to better reflect the reality. It would be preferred to conduct time varying parameter estimation, however, given that the only observable input is  $I$ , the infected population, it is not computationally feasible. Our contingency is to utilize the turning points as an indicator of “narrative mutation” and adjust the parameters and susceptible population (part of the recovered population become susceptible again) for the “post-mutation stage”.

We smooth the  $I$  series using Hodrick–Prescott filter and find the turning points based on whether the smoothed  $I$  change from a downward sloping trend to an upward sloping trend. We posit the scale of the  $S$  (susceptible population) adjustment depends on the sharpness of the turn. We estimate the adjustment based on a simple function as follows:

$$\Delta R(t) = \hat{R}(t) \left( \frac{1}{1 + e^{-\text{angle}}} \right) + \epsilon_t \tag{4}$$

$$\text{angle} = (|g_{\text{post}}| + |g_{\text{pre}}|) / 2. \tag{5}$$

The adjustment of  $S$  (or equivalently,  $R$ ) around turning points depends on the current observed  $I(t)$ , the estimated  $\hat{R}(t)$  and the scale of the raw change of the growth rate of  $I$  between time  $t + 1$  and  $t$ , that is, the difference between the growth rate before the turning point,  $g_{\text{pre}}$ , and the growth rate after the turning point,  $g_{\text{post}}$ .<sup>15</sup> Specifically,  $\hat{S}(t + 1) = (1 - I(t)) \left( \frac{\hat{S}(t) + \Delta R(t)}{\hat{S}(t) + \hat{R}(t)} \right)$ . The angle is then estimated as the average of absolute slopes before and after each point in the sample period. We use a sigmoid function to evaluate the scale of turns, so that the maximum adjustment does not exceed the current  $R$ . It implies that when there are extreme narrative mutations making the narrative very attractive again, large portion of the recovered population become susceptible again. With the sigmoid function, we also arbitrarily restrict the adjustment of recovered population to be at least half of the current recovered population. This mechanism mimics the re-occurring patterns in narratives. In detail, the modification method is shown in a pseudo-program in Appendix 1, in which we also provide a simulation to show the advantages of the adjusted model.

<sup>15</sup> Please note that  $R(t)$  here means the recovered population ratio, instead of the effective reproduction number.

### 3.3 Intertemporal links between narratives and the stock market

We use the Tetlock (2007) method to evaluate whether the narratives contain new information or reflect investor sentiment to the stock market. After testing for stationarity, the narrative features' level or first difference time series are put in a VAR model, to examine the causality relationship with stock market measures including the detrended log volume and the S&P 500 index return.<sup>16</sup> Following Tetlock, we include five dummy variables for day-of-the-week and January to control for potential return anomalies.<sup>17</sup> Five lags of VIX index are also included to control for past volatility. The three major regressions of the VAR model are as follows:

$$\mathbf{Y}_t = \alpha + \beta \cdot \mathbf{L5}(\mathbf{Y}_t) + \lambda \cdot \mathbf{Exog}_t + \epsilon_t \quad (6)$$

Here,  $\mathbf{Y}_t$  represents the vector of variables at time  $t$ , including the S&P 500 market return, the detrended log volume, and one of the narrative features.<sup>18</sup>  $\mathbf{Exog}_t$  represents the five exogenous variables.  $\mathbf{L5}$  is a lag operator transforming the variable attached to a row vector of its five lags.  $\beta$  is the coefficient matrix for the endogenous variables within 5 lags.

As explained by Tetlock, if the tones measures can predict the market return without reversion in the long term, the text contents are likely to contain fundamental information that has not been incorporated into prices; if the tones measure can predict short-term return with an reversion in the long term, the text contents might serve as a proxy for low past or future investor sentiment; if no predictability was found, the content might contain no new information about past, present, and future dividends. To reveal the relationships, Tetlock looks at the coefficients and conducts Wald tests to see whether the first lag, the sum of the second to the fifth lag, and the sum of all five lags' coefficients are significant. However, an assumption of this method is that, in the returns' regression, the first lag of tones should be significant. The absence of a short-term predictability also implies the lack of information content. Furthermore, the permanent effect is determined using only five lags' coefficients, losing the possibility of a longer term effect. For the reasons, we also rely on impulse response functions to reveal the long-term relationships.

In this paper, we examine other narrative features in addition to tones to examine whether those variables contain information or sentiment. We also consider the change rate of VIX as one of the endogenous instead of exogenous variable as one robustness test. Besides the basic VAR method, we also consider pure Granger causality tests because of small sample issue, given that we only have about one and half years' data for the coronavirus narratives.

<sup>16</sup> We mainly use ADF test and Phillips–Perron unit root tests.

<sup>17</sup> Only four dummy variables for weekdays excluding Friday were used to prevent multicollinearity.

<sup>18</sup> The detrended log volume is calculated following Tetlock (2007).

### 3.4 Perennial risk narratives

In the context of a pandemic, perceivable uncertain and pessimistic sentiments are prevailing in media news and investors' daily conversations. Interestingly, similar patterns manifest around all crisis periods and volatile times. We are interested in examining whether there are narratives that characterize bad times being present around the "COVID market tumbles" from February to March 2020. If there are, we further study their role in the financial market.

To find such risk narratives, we utilize LDA to learn from a set of textual data in the history. LDA learns data-driven clusters of words from texts. It means that the clusters are not guaranteed to represent interpretable topics. However, if similar narratives do went viral around bad times, they are likely to be captured by the data-driven process as latent topics. The existence of such topics is sufficient to imply that perennial narratives exist and have economic meaning.

With the combination of data selection, topic modeling and the relationship with economic measures, we maximize the possibility of identifying recognizable important economic narratives. Here, we train LDA algorithms on general market news over 40 years, instead of the coronavirus news (that we used for coronavirus narratives' analysis), to capture potential long-lasting common narratives. The general banking and finance news, characterized by at least 5 mentions of "market", necessarily contain narratives circulating in the economy and financial markets. It was mentioned in Baker et al. (2016) and Nyman et al. (2021) that the correlation of new measures with some standard economic and financial measures is useful in revealing whether there are similar information contained. We check the correlation between the LDA topic prevalence indices with the CBOE Volatility Index (VIX). A topic, whose public attention is highly correlated with the implied volatility, should necessarily contain narrative contents related to the market's perception of risk. We train LDA with a set of  $K$ s, on historical texts from 1980 to the end of 2019. The topics will be ranked by the correlation between their topic prevalence and VIX. We check the rankings across the LDA models with different  $K$  s, a topic that is always on the top of the ranking and that being present in most LDA models is preferred, for its robustness and stability. The texts in 2020 is treated as the out-of-sample testing set, which covers the first year of the COVID-19 pandemic. As we identified a topic satisfying the criteria, we fit the LDA algorithm on the testing data and focus on its topic intensity around the COVID market tumbles in February and March. If the topic became viral around the pandemic market collapse, we have some evidence of the existence of perennial economic narratives that reflect investors' perception of risk. We interpret the topic based on mainly three information, namely its highest probability words, the high correlation with VIX and economic events, and the human understanding of the representative texts.

The topic filtered through the criteria has high correlation with VIX, we further use Granger-causality to test the causal relationships between the text-based measure and the financial measure. If there are stationary or co-integrated, we can rely on the Granger-causality test results of level data. A Granger-causal relationship is useful as a stylized fact and is potentially useful in forecasting. For example, Nyman

et al. (2021) find that their text-based measure Grange-causes VIX. They claim that their measure could be included in an indicator dashboard to monitor systemic risk.

## 4 Data

The textual data we use to identify popular narratives are media news from LexisNexis. We set some constraints in the searching queries to locate the articles that contain the topics of interest. The first group of narratives we study is the coronavirus narratives circulated in the economy. Through the “Power Search” function on LexisNexis, we apply a subject tag of “Banking & Finance” and a source tag of “major world newspapers” and set the searching terms as “at least one mention of “coronavirus” or “COVID””.<sup>19</sup> The subject tag, along with the keywords, ensure that the retrieved news are both finance and coronavirus related. On one hand, because “coronavirus” and “COVID” primarily indicate topics related to coronavirus and were not likely to be mentioned in banking and finance news before the pandemic, we are confident that the selected news all contain narratives about the COVID-19 pandemic. On the other hand, because major world newspapers report news with formal professional languages, “coronavirus” or “COVID” are always mentioned when the article is focused on coronavirus, the risk of under identification is low. We download all the COVID-19 news published from 1st January 2020 to 25th April 2021.

Our second group of narratives is the market narratives. We download another set of news with at least five mentions of “market” and published by major U.S. newspapers. We download all the market news published within the period from 1st January 1980 to 31st December 2020.

The most frequent words of the cleaned early-stage coronavirus narratives can be seen in Fig. 1.<sup>20</sup> Some key words clearly reflect the general topics in the data set. For example, “China” and “US” are the two countries attracted the spotlights during the early stage of the pandemic; “health”, “viru” (the stem of virus) and “pandemic” are clear indicators of the COVID-19 narratives. Words such as “stock”, “market” and “economy” indicate that the narratives are focusing on the stock market. Interestingly, “crisi” (the stem of “crisis”) is one of the top 50 frequent words during the early stage of the pandemic, indicating the presence of negative-sentiment narratives. This preview of our data shows that “data selection” alone is sufficient to locate the collection of popular coronavirus narratives.

We clean the textual contents following the pre-processing routines.<sup>21</sup> For each piece of content, we find its publish time from the texts using regular expressions.

<sup>19</sup> see <https://advance.lexis.com>.

<sup>20</sup> We generate this word cloud using only the texts before May 2020, which is the early stage of the pandemic. The reason is that the topics are more complicated and diversified along with time, but our purpose is only to show that the COVID narratives are well captured.

<sup>21</sup> We removed redundant characters such as emojis, multiple lines/spaces, punctuations, numbers, html tags, tokenized the texts and removed stopwords.



**Fig. 1** Word cloud of the news corpus. Notes: the word cloud was generated based on term frequency. The top 50 frequent words are visualized, with larger font size representing higher frequency. Stopwords were removed for visualization

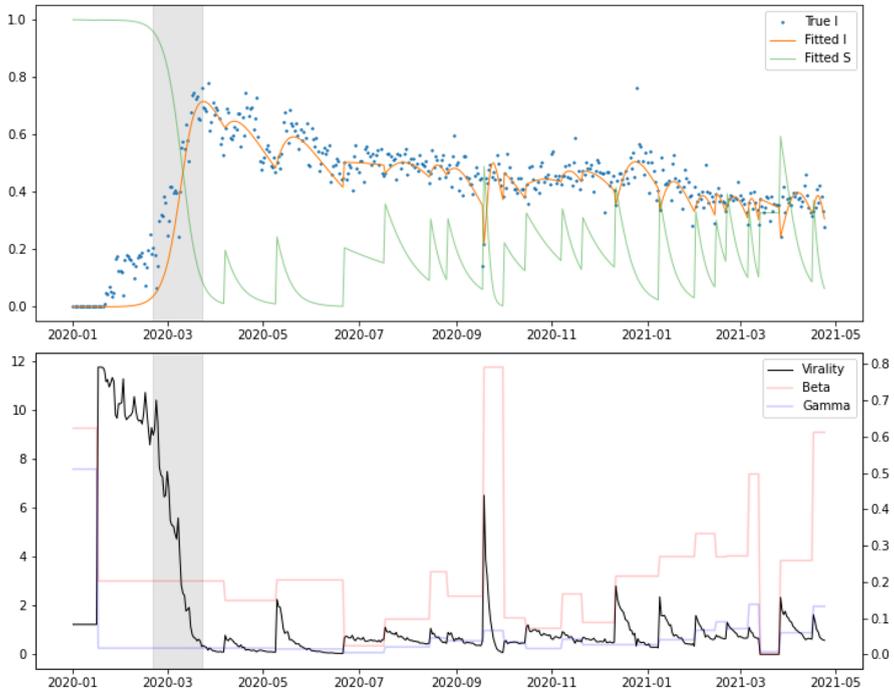
## 5 Results

With the narrative features we reduce the high-dimensional texts to a set of interpretable dimensions. We first show the evolution of the COVID-19 narratives on the narrative features, with a focus around the market tumbles in February and March 2020. We find that the features of coronavirus narratives change in a way consistent with our hypothesizes, implying a potential role of narratives in the stock market. We then report the results of the intertemporal links between narratives and the stock market. Lastly, we provide the evidences of a perennial risk narrative and its role in the stock market.

### 5.1 Evolution of coronavirus narratives

#### 5.1.1 Narrative intensity and virality

First, for the convenience of discussion, we report the intensity measure along with the virality measure in Fig. 2. The upper panel plots the narrative intensity index (blue dots) with the fitted values (orange line) using the adjusted SIR model. Here, the intensity index is constructed as the ratio of the count of “coronavirus narrative documents” over the count of banking and finance news, both are published by major world newspapers and stored in LexisNexis. We show that coronavirus narratives dominated finance narratives since the end of February 2020 and is still dominating at the end of our sample window. There are multiple small waves and a long term decreasing trend can be observed after the peak at 27th March 2020, with 78% of the banking and finance news covering coronavirus narratives. The coverage rate is about 35% at the end of April 2021. The dominance of one topic around the severe market decline is consistent with Nimark and Pitschner (2019). The COVID-19 pandemic shifts the news focus and make the media coverage more homogeneous. The high intensity index shows that coronavirus narratives is a rare case in the history



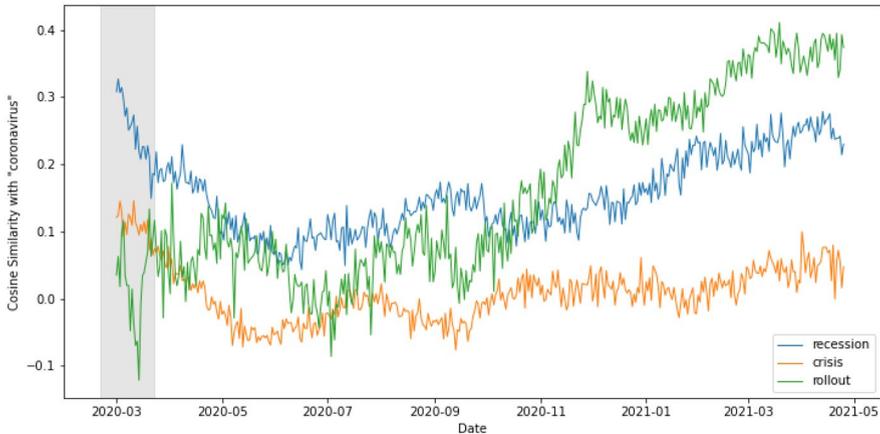
**Fig. 2** COVID-19 narrative intensity and virality. Notes: this figure plots the evolution of intensity (upper panel) and virality (lower panel) indices of the COVID narratives. On the upper panel, the blue dots represent the true observations of the proxy of infected population  $I$ . The orange line represents the modeled curve of  $I$  and the green line represents the modeled curve of the susceptible population  $S$ . On the lower panel, the dark line represents the estimated virality of the coronavirus narratives. The light red and blue lines represent the SIR coefficients  $\beta$  and  $\gamma$ , respectively, with values shown on the right y axis. The grey-shadowed area indicates the severe decline period of S&P 500 Index from February 21st, 2020 to March 23rd, 2020

that has both large and long-lasting public attention among investors. Projecting this pattern in an epidemiological model, one would expect a high contagion rate ( $\gamma$ ) and a low recovery rate ( $\beta$ ), leading to a high basic reproduction number.<sup>22</sup>

The lower panel shows the narrative virality index (dark line). The estimated virality index, measured as the effective reproduction number, was high at the start of the narrative epidemic with the sudden increase of “ $I$ ” in only a few weeks. The virality index decreases when a large portion of the population is already infected or recovered, only mutations can relocate part of the recovered population to susceptible population, and inflate the virality measure. Our epidemiological model successfully identified multiple turning points, at which we believe the narratives mutate

<sup>22</sup> According to the SIR model, if both the contagion rate and the recovery rate are small, the epidemic curve will be flat; if the contagion rate is large and the recovery rate is small, the epidemic curve will be both high and wide; if both the contagion rate and the recovery rate are high, the epidemic curve will be high and narrow; if the contagion rate is low and the recovery rate is high, the disease will die out soon.





**Fig. 4** Cosine similarity with “coronavirus”. Notes: this figure displays the evolution of cosine similarity of “coronavirus” with some keywords. The cosine similarity on each day was calculated from the embedding vectors, trained with coronavirus narrative documents published within 50 days. The grey-shaded area indicates the severe decline period of S&P 500 Index from February 21st, 2020 to March 23rd, 2020

### 5.1.3 Topics and topic consensus

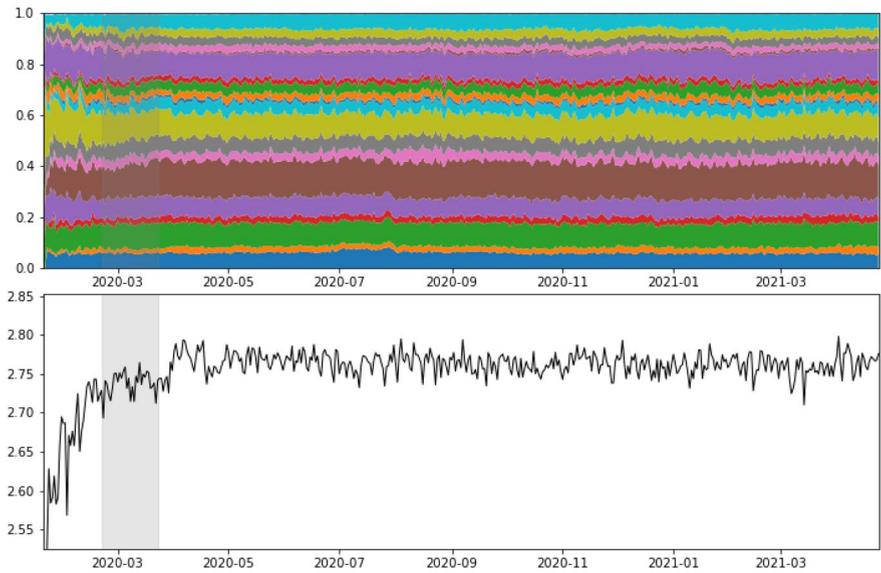
Third, we discuss the results from topic modeling. The changes of 20 topics are visualized in Fig. 5 with the consensus measure.<sup>23</sup> The difference of the two sub-graphs is that the latent topics on Fig. 5a were estimated using the LDA algorithm trained on coronavirus narrative documents, and those on Fig. 5b were estimated using the LDA algorithm trained on general market narrative documents. Both graphs show the same pattern. The coronavirus narratives became diversified before the market collapse. The pattern is consistent with the empirical findings of Bertsch et al. (2021) and Nyman et al. (2021). Our finding is different from those in the literature by applying topic consensus inside a specific set of narratives—the event relevant narratives. Our result demonstrates a stylized fact that the pandemic narratives circulating in the economy, diversify into competing topics around the severe pandemic-driven market declines. More topics were covered in the coronavirus narratives along with the development of the COVID-19 pandemic. This happens when the event and its impact become more complicated.

We report the recognized topics ( $K$  equals 20) on Table 1 and visualize the evolution of the coronavirus-specific topics, namely “COVID updates” and “vaccine”, in Fig. 6. The topics’ evolution shows that “COVID updates” was an important topic in the coronavirus narratives only at the beginning of our sample period. It indicates that during the early stage of the pandemic, investors care about the updates, possibly for their evaluation of the risk. However, this topic soon lost its importance,

<sup>23</sup> We use the version with  $K$  equals 20 for a clean visualization, while the topic consensus pattern holds with alternative values of  $K$ .



(a) Coronavirus topics



(b) Economic topics

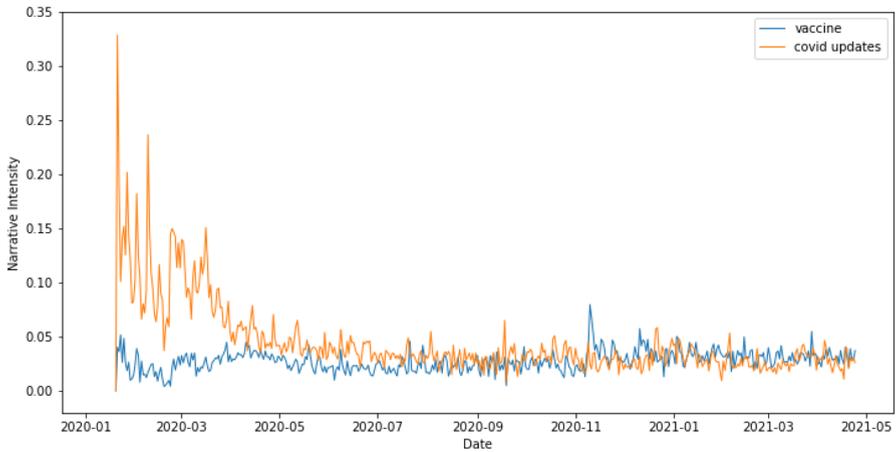
**Fig. 5** Narrative consensus over coronavirus and economic topics. Notes: the two graphs show the evolution of LDA topics and the topic entropy.  $K$  is selected as 20. The LDA of the first graph was trained on coronavirus narratives, while the LDA of the second graph was trained on general market narratives. The grey-shadowed area indicates the collapse period of S&P 500 Index from February 21st, 2020 to March 23rd, 2020

**Table 1** Names assigned to the LDA topics

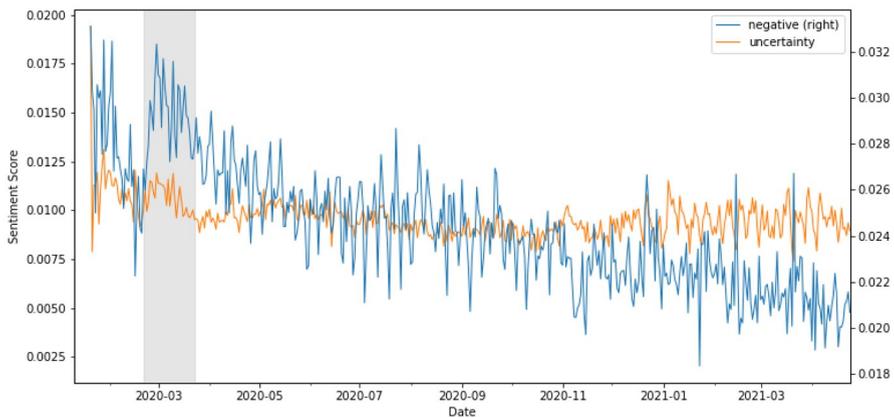
	Training data	
	Coronavirus narratives	Market narratives
Topic1	Australian government	Agriculture
Topic2	Vaccine	Energy
Topic3	Airlines and sports	US and China
Topic4	Banking	Stock trading
Topic5	Economy	Economy
Topic6	Company investment	Life
Topic7	Life	Bonds
Topic8	Oil	Oil
Topic9	Bond market	Funds
Topic10	COVID updates	Company fundamentals
Topic11	UK and EU	Living
Topic12	Stock market	Company leaders
Topic13	UK government	Government plans
Topic14	Politics	Logistics
Topic15	Company business	Mortgage and loans
Topic16	China	Technology
Topic17	Housing	Housing
Topic18	Company operations	Stock market
Topic19	US presidential election	Business deals
Topic20	Monetary policy	Economy and Risk

This table presents the topic names manually assigned to the LDA topics identified from both coronavirus narratives and market narratives. We assign names via analyzing the representative words of each latent topic and the articles with the highest topic probability

leaving space for other competing topics. Although not as important as the “COVID updates” topic, the “vaccine” topic did increase its intensity since November 2020, which is consistent with the increase of the semantic similarity between “coronavirus” and “rollout” (of vaccines). There are some common topics that identified by LDA algorithms trained on coronavirus narratives and general market narratives, such as “economy”, “stock market” and “housing”. This is because we have limited the topics via data selection for both, and the data-driven approach successfully captured them. Coronavirus-specific topics, namely “COVID updates” and “vaccine”, were identified from the coronavirus narratives only. From the evolution of narrative intensity of the two topics reported in Fig. 6, we find that the coronavirus updates was a dominant topic in the coronavirus narratives only before May 2020, while the vaccine topic increased its weight in the coronavirus narratives since November 2020, consistent with our finding from Fig. 4.



**Fig. 6** “COVID updates” and “vaccine” narratives. Notes: this figure presents the evolution of narrative intensity indices for “vaccine” and “COVID updates” topics identified by LDA. The LDA was trained on coronavirus narratives



**Fig. 7** Uncertainty and negative tones of the COVID narratives. Notes: this figure plots the negative and uncertain tone series for the COVID narratives. The tones are calculated using the Loughran and McDonald (2011) word lists. The blue line represents the negative tone series (values on the right vertical axis) and the orange line represents the uncertainty tone series. The grey-shadowed area indicates the collapse period of S&P 500 Index from February 21st 2020 to March 23rd 2020

### 5.1.4 Textual sentiment

Lastly, we report the textual sentiment scores of the coronavirus narratives on Fig. 7. The negative and uncertainty tones were the highest in the beginning of the global pandemic. A second wave of pessimism and uncertainty in the coronavirus narratives was found starting from February 16th 2020 and reaching the peak around the beginning of March 2020. This period coincides with the severe decline in the stock

market between mid-February to mid-March. Both negative and uncertain sentiment decreased since around 10th March of 2020, while the S&P 500 index increased since the mid of March. It shows that sentiment in coronavirus narratives co-move with the stock market index.

### 5.1.5 Summary of coronavirus narratives

In summary, around the period of the market collapse, the pandemic narratives goes viral and dominant, fragment into competing topics, become more pessimistic and uncertain, and are more linked to the “recession” and “crisis” context. All the patterns are reasonable and consistent with literature and our hypothesis that narratives could play an important role in the stock market. Although the causal relationship between the narratives and the severe stock market decline is not identified so far, the patterns could be used as warning signals of severe market reactions in a future crisis.

## 5.2 Do coronavirus narratives contain information?

With the stylized facts above, we now consider whether coronavirus narratives have the potential to play a role in the stock market during the pandemic. This section reports the results of the investigation on the intertemporal links between the important narratives and the stock market.

After running the vector auto-regressions, we find no evidence of predictability from coronavirus narrative measures to market return and trading volume. We only report the results of narrative tones in Tables 2 and 3. Although the signs of the coefficients are consistent with those in Tetlock (2007), they are not significantly different from zero. This implies that none of the measures contain new information. In addition, the market variables cannot predict the coronavirus narratives. An alternative justification for no predictability, is the potentially mixed information in the coronavirus narratives. The coronavirus news, unlike the single column news studied by Tetlock and Garcia (2013), do not have a strong “tag along” flavor specifically to the past asset price movements.<sup>24</sup> Instead, as seen in Table 1, there are many topics covered in our sample. When we run Granger-causality tests (see Table 4), we find that market return Granger-causes the negative tone of the coronavirus narratives. We only reject the null hypothesis of no Granger-causality from negative tone to returns with one lag. Similar results are found regarding trading volume. We also find that VIX can Granger-cause both negative tones and narrative intensity. While negative tones Granger-cause VIX, there is no Granger-causality from narrative intensity to VIX.

The pure Granger-causality results provide some evidence of intertemporal links between the coronavirus narratives and market conditions. In general, the narrative

---

<sup>24</sup> By “tag along”, we mean the behavior of following the last-day performance and trying to explain the reasons.

**Table 2** Estimated coefficients of COVID news tones in predicting the return and detrended log volume

Tone	Regressand: S&P 500 return			Regressand: volume		
	Negative	Uncertainty	Weak modal	Negative	Uncertainty	Weak modal
Lag 1	-2.289 (2.739)	-6.045 (7.200)	-17.671 (11.006)	1.981 (12.564)	-2.205 (33.062)	-3.150 (51.399)
Lag 2	0.428 (2.558)	-2.364 (7.396)	2.498 (9.550)	-0.203 (11.733)	-19.284 (33.958)	-3.803 (44.601)
Lag 3	-0.579 (2.507)	5.005 (7.372)	13.090 (10.203)	9.686 (11.499)	17.644 (33.851)	-25.949 (47.649)
Lag 4	-0.151 (2.289)	3.372 (7.078)	11.469 (9.118)	-9.061 (10.499)	-41.330 (32.498)	-52.046 (42.582)
Lag 5	-1.184 (2.039)	0.320 (4.487)	-3.037 (7.039)	-7.223 (9.355)	9.633 (20.601)	19.742 (32.872)
$\chi^2(5)$ [Joint]	1.091	1.795	4.892	1.940	2.517	3.257
p-value	0.955	0.877	0.429	0.857	0.774	0.660

This table presents the estimated OLS coefficients and Wald test results of the five lags of tones in the return equation

**Table 3** Estimated coefficients of returns and detrended log volume in predicting COVID narrative tones

	Regressand: negative		Regressand: uncertainty		Regressand: weak modal	
	S&P 500 return	Volume	S&P 500 return	Volume	S&P 500 return	Volume
Lag 1	-0.0149 (0.0078)	<b>0.0026</b> (0.0012)	-0.0012 (0.0030)	0.0007 (0.0005)	0.0016 (0.0019)	0.0005 (0.0003)
Lag 2	-0.0084 (0.0081)	-0.0010 (0.0014)	-0.0013 (0.0030)	-0.0002 (0.0005)	0.0006 (0.0019)	0.0001 (0.0003)
Lag 3	-0.0004 (0.0082)	0.0006 (0.0014)	0.0048 (0.0030)	0.0007 (0.0005)	0.0026 (0.0020)	0.0003 (0.0003)
Lag 4	-0.0028 (0.0081)	0.0013 (0.0013)	0.0035 (0.0031)	<b>-0.0010</b> (0.0005)	0.0033 (0.0020)	<b>-0.0011</b> (0.0003)
Lag 5	-0.0021 (0.0057)	0.0012 (0.0012)	-0.0002 (0.0022)	0.0006 (0.0005)	<i>0.0026</i> (0.0014)	0.0001 (0.0003)
$\chi^2(5)$ [Joint]	5.2657	14.1475	4.2704	10.6441	7.6634	14.2056
p-value	0.3843	0.0147	0.5112	0.0589	0.1758	0.0144

This table presents the estimated coefficients and Wald test results of five lags of returns and detrended log volume in the news tone equation

Italic font indicates significance at the 10% level; bold font indicates significance at the 5% level; and italic and bold font indicates significance at the 1% level

**Table 4** Wald test *p*-values of Granger-causality between coronavirus narratives and market indices

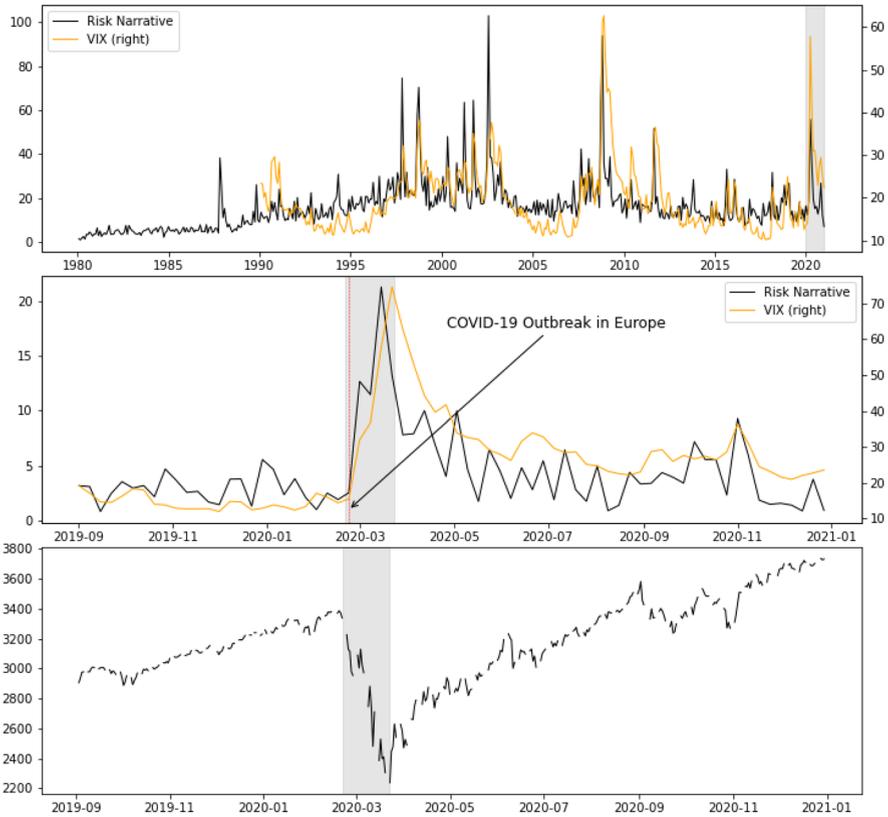
Direction	Number of Lags					
	1	2	3	4	5	6
Negative tone → Return	<b>0.0138</b>	0.1453	0.3878	0.2849	0.4604	0.3104
Return → Negative tone	<b>0.0006</b>	<b>0.0033</b>	<b>0.0094</b>	<b>0.0174</b>	<b>0.0111</b>	<b>0.0200</b>
Negative tone → Volume	<b>0.0348</b>	0.1684	0.5584	0.7246	0.5094	0.3396
Volume → Negative tone	<b>0.0011</b>	<b>0.0060</b>	0.0610	0.1142	0.0977	<b>0.0063</b>
Uncertainty tone → Return	0.0903	0.2772	0.4933	0.4248	0.4686	0.4762
Return → Uncertainty tone	<b>0.0200</b>	<b>0.0415</b>	<b>0.0263</b>	0.0839	0.1291	0.1250
Uncertainty tone → Volume	0.1236	0.2032	0.5260	0.6723	0.8204	0.9061
Volume → Uncertainty tone	<b>0.0094</b>	0.0896	0.4776	0.5747	0.4681	0.5209
Negative tone → VIX	<b>0.0011</b>	<b>0.0080</b>	<b>0.0256</b>	0.0784	0.1377	0.2563
VIX → Negative tone	<b>0.0058</b>	<b>0.0012</b>	<b>0.0025</b>	<b>0.0056</b>	<b>0.0059</b>	<b>0.0018</b>
Uncertainty tone → VIX	0.1229	0.1603	0.3452	0.7200	0.6809	0.5040
VIX → Uncertainty tone	0.7683	0.2078	0.3057	0.5344	0.6531	0.7449
Narrative intensity → VIX	0.6569	0.5125	0.2766	0.7908	0.8141	0.8364
VIX → Intensity	<b>0.0000</b>	<b>0.0001</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>	<b>0.0000</b>
Narrative virality → VIX	0.2081	0.0640	0.0627	0.2323	0.1864	0.1907
VIX → Virality	<b>0.0043</b>	<b>0.0139</b>	<b>0.0160</b>	<b>0.0385</b>	0.0607	0.1244
Narrative intensity → Return	0.1584	0.4307	0.4587	0.5283	0.4812	0.3313
Return → Intensity	0.2890	0.0652	<b>0.0230</b>	<b>0.0066</b>	<b>0.0002</b>	<b>0.0040</b>
Narrative virality → Return	<b>0.0270</b>	0.1528	0.3216	0.2573	0.3368	0.3077
Return → Virality	0.2057	0.2314	0.3696	0.5444	0.6211	0.7750
Narrative intensity → Volume	0.5070	0.1231	<b>0.0236</b>	<b>0.0395</b>	<b>0.0435</b>	0.0663
Volume → Intensity	<b>0.0224</b>	<b>0.0149</b>	<b>0.0242</b>	<b>0.0052</b>	<b>0.0041</b>	<b>0.0138</b>
Narrative virality → Volume	<b>0.0354</b>	0.1595	0.1626	0.2178	0.2929	0.3679
Volume → Virality	0.3654	0.5568	0.7303	0.8499	0.7561	0.8832

This table presents the Wald test *p*-values of Granger-causality between coronavirus narratives and market indices. The Granger-causality tests were applied on level data. Bold font indicates significance level smaller than 5%

intensity and the textual sentiments, especially the negative tones, are Granger-caused by market returns, implied volatility and trading volume. Less or no predictability were found in the reverse direction. The virality measure can be predicted by VIX, but it has no predictability to the market condition measures. That means the coronavirus narratives are significantly influenced by the stock market moves, and reflect the past market performance.

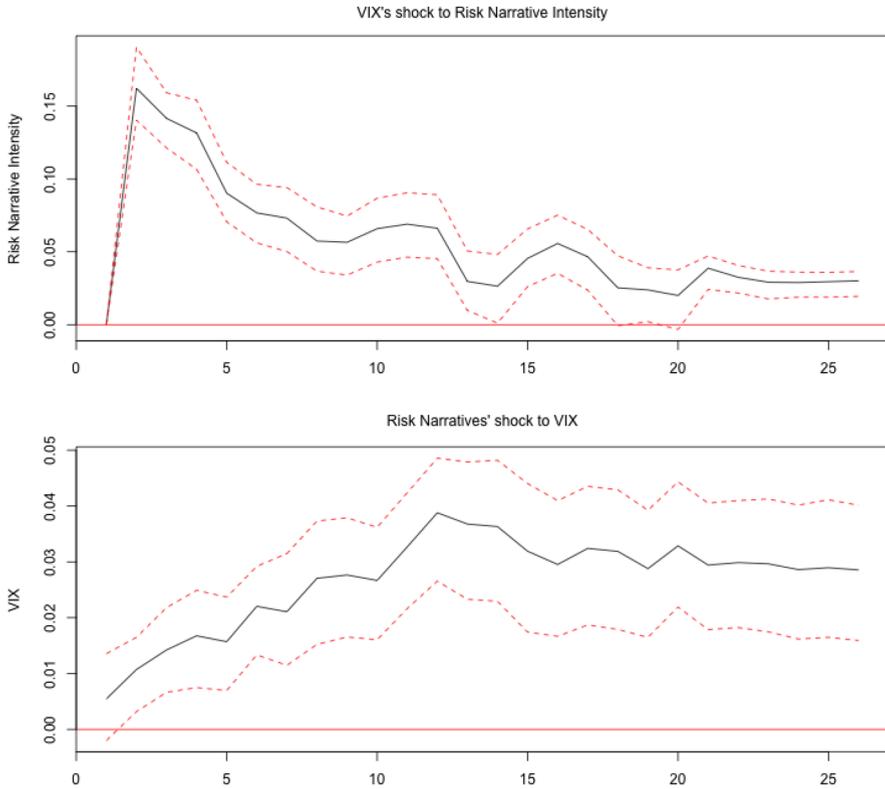
### 5.3 Perennial risk narratives

With the procedure proposed in the method section, we successfully identified a latent topic that can be interpreted as a perennial risk narrative. This topic can be found in all of our tested LDA algorithms with different values of *K* and random seed. We visualize the intensity measure of the topic from the LDA algorithm with



**Fig. 8** Perennial risk narratives. Notes: the upper panel plots the monthly intensity measure of the “risk narratives” topic and the monthly VIX. The topic was detected by applying LDA on historical general market narratives from January 1st, 1980 to December 31st, 2019. The grey-shaded area indicate the “out-of-sample” period, January 1st, 2020–December 31st, 2020. The middle panel plots the weekly intensity measure of the “risk narratives” topic and the weekly VIX from September 2019 to December 2020. The grey-shaded area indicates the collapse period of S&P 500 Index from February 21st, 2020 to March 23rd, 2020. The lower panel plots the daily S&P 500 index on the same time range of the middle panel. The grey-shaded area indicates the collapse period of S&P 500 Index from February 21st, 2020 to March 23rd, 2020

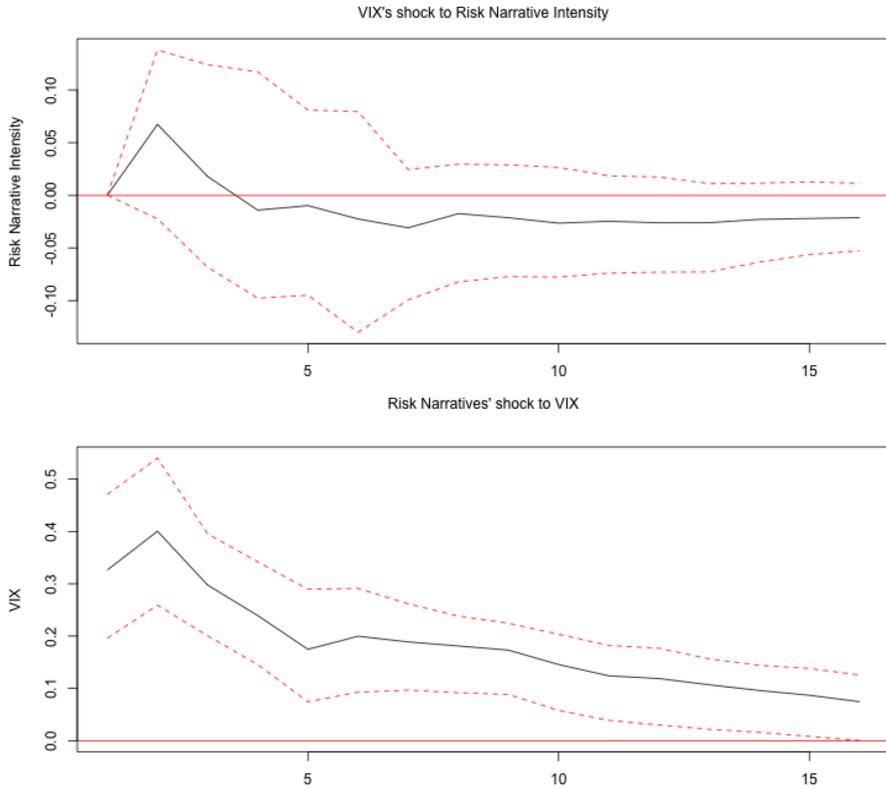
$K$  equals to 20 in Fig. 8. The intensity measure (aggregated by sum) of this topic matches with VIX surprisingly well for the period in which VIX is available. Both series spike simultaneously around almost all historical major events, including the Black Monday in 1987, the burst of the dot bubble in 2000 and the 2008 financial crisis. Given that the events have heterogeneous triggers and context, our text-based measure must capture some homogeneous components of the market narratives during bad times. By reading a selection of articles with the highest probability of this topic during multiple events, we confirm that “risk narrative” is an accurate description. A selection of snippets of the risk narrative around the 1987, 2008 and 2020 market downturns can be found in Appendix 2.



**Fig. 9** IRFs of risk narratives and VIX (daily). Notes: the upper panel plots the response of risk narrative intensity measure to one standard deviation shock of VIX. The lower panel plots the response of VIX to one standard deviation shock of the risk narrative intensity measure. Both series were standardized to be zero mean and unit standard deviation in the system. The IRFs are generated from a VAR model, with daily risk narrative intensity and daily VIX

The train/test split procedure proves that the co-movement between the risk narrative and VIX is not from data-mining. The algorithm training and the topic selection were conducted using texts before 2020, and the trained algorithm was applied on the unseen data in 2020 to calculate the topic intensity index. Focusing on the window between September 2019 and December 2020, we find that both series suddenly increase one day after the COVID-19 outbreak in Europe. The perfect timing suggests that the outbreak in Europe triggers a sudden panic, which initiated a surge of risk perception and a series of severe market declines.

When  $K$  equals 20, the monthly narrative intensity series has 0.62 correlation with VIX, the highest among those of all the 20 latent topics. This pattern holds when we use different  $K$ s in the LDA. Both ADF test and Phillips–Perron test show that both series are stationary either daily or monthly. An OLS regression shows a significant positive relationship between the two stationary series. Thus, we put the two level series in a VAR system to examine the relationships. We



**Fig. 10** IRFs of risk narratives and VIX (monthly). Notes: The upper panel plots the response of risk narrative intensity measure to one standard deviation shock of VIX. The lower panel plots the response of VIX to one standard deviation shock of the risk narrative intensity measure. Both series were standardized to be zero mean and unit standard deviation in the system. The IRFs are generated from a VAR model, with monthly risk narrative intensity and monthly VIX

show the impulse response functions (IRFs) using daily data in Fig. 9. A one standard deviation shock of VIX causes an immediate increase of risk narrative intensity for about 15% standard deviation, but the effect decreases quickly. By contrast, a positive one standard deviation shock of risk narrative intensity cause a long-term increase of VIX, with the greatest effect, 4% of the standard deviation of VIX, on day 12. The Granger-causality tests with Newey–West standard errors on the two daily series show a dual-way Granger-causal relation.<sup>25</sup> To find their relationship in the long term, after aggregating the data monthly, we run the VAR again and plot the IRFs in Fig. 10. Monthly, VIX has no predictability on the risk narrative intensity measure, but the risk narrative intensity index can predict VIX. The impulse response functions show that a positive shock of monthly risk narrative intensity causes a long-term increase on VIX. One shock of the narrative causes a 30% and 40% standard deviation increase on VIX in 1 and 2 months,

<sup>25</sup> We calculate the Newey–West standard errors with the same number of lags as the VAR model.

**Table 5** VAR with the virality measure of the risk narratives

	Regressand: S&P 500 return	Regressand: volume	Regressand: virality	
	Virality	Virality	S&P 500 return	Volume
Lag 1	0.001 (0.004)	0.054 (0.056)	- 0.036 (0.053)	- 0.004 (0.002)
Lag 2	- 0.010 (0.005)	0.004 (0.071)	<b>0.108</b> (0.053)	-0.001 (0.002)
Lag 3	- 0.008 (0.005)	0.061 (0.072)	<b>0.152</b> (0.053)	- 0.003 (0.002)
Lag 4	<b>0.014</b> (0.005)	- 0.087 (0.071)	<b>0.193</b> (0.053)	0.001 (0.002)
Lag 5	- 0.006 (0.004)	0.017 (0.056)	<b>0.068</b> (0.033)	<b>0.005 (0.002)</b>
$\chi^2(5)$ [Joint]	27.342	5.762	30.553	9.128
p-value	0.000	0.330	0.000	0.104

This table presents the estimated coefficients and Wald test results of the five lags of variables in the VAR with the virality measure of the risk narratives. The regression is based on 7,805 observations from January 1980 to December 2020. Standard errors are reported in parentheses

Italic font indicates significance at the 10% level; bold font indicates significance at the 5% level; and italic and bold font indicates significance at the 1% level

**Table 6** VAR with the intensity measure of the risk narratives

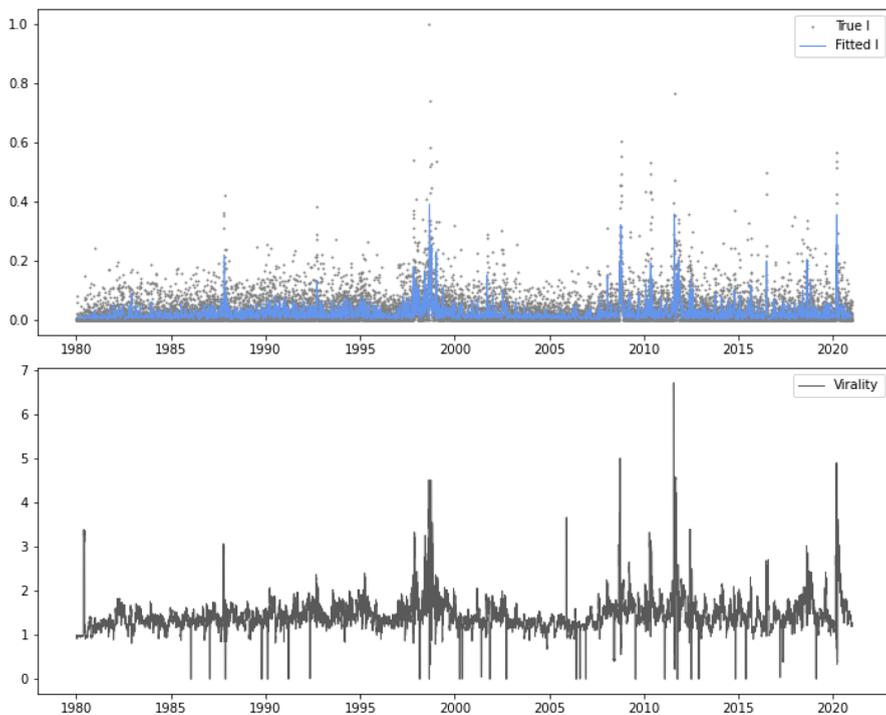
	Regressand: S&P 500 return	Regressand: volume	Regressand: intensity	
	Intensity	Intensity	S&P 500 return	Volume
Lag 1	- <b>0.007</b> (0.003)	0.001 (0.037)	<b>0.198</b> (0.081)	<b>0.015</b> (0.004)
Lag 2	0.002 (0.003)	0.040 (0.038)	- 0.143 (0.081)	0.002 (0.004)
Lag 3	- 0.002 (0.003)	- 0.024 (0.038)	- 0.156 (0.081)	0.004 (0.004)
Lag 4	- 0.002 (0.003)	0.057 (0.037)	- 0.081 (0.081)	- 0.004 (0.004)
Lag 5	0.001 (0.003)	0.020 (0.036)	- <b>0.172</b> (0.050)	- 0.002 (0.003)
$\chi^2(5)$ [Joint]	8.995	5.737	25.993	31.147
p-value	0.109	0.333	0.000	0.000

This table presents the estimated coefficients and Wald test results of five lags of independent variables in the VAR with the intensity index of the risk narratives. The regression is based on 7,805 observations from January 1980 to December 2020. Standard errors are reported in parentheses

Italic font indicates significance at the 10% level; bold font indicates significance at the 5% level; and italic and bold font indicates significance at the 1% level

respectively. The effect then gradually decrease but lasts about 15 months. The Granger-causality tests with Newey–West standard errors show that risk narrative Granger-causes VIX but not the other way around. The results show that the risk narrative measure might be impacted by implied volatility in the short term, but it has long-term predictability on implied volatility, resembling a “feedback loop” [a term borrowed from Shiller (2015)].

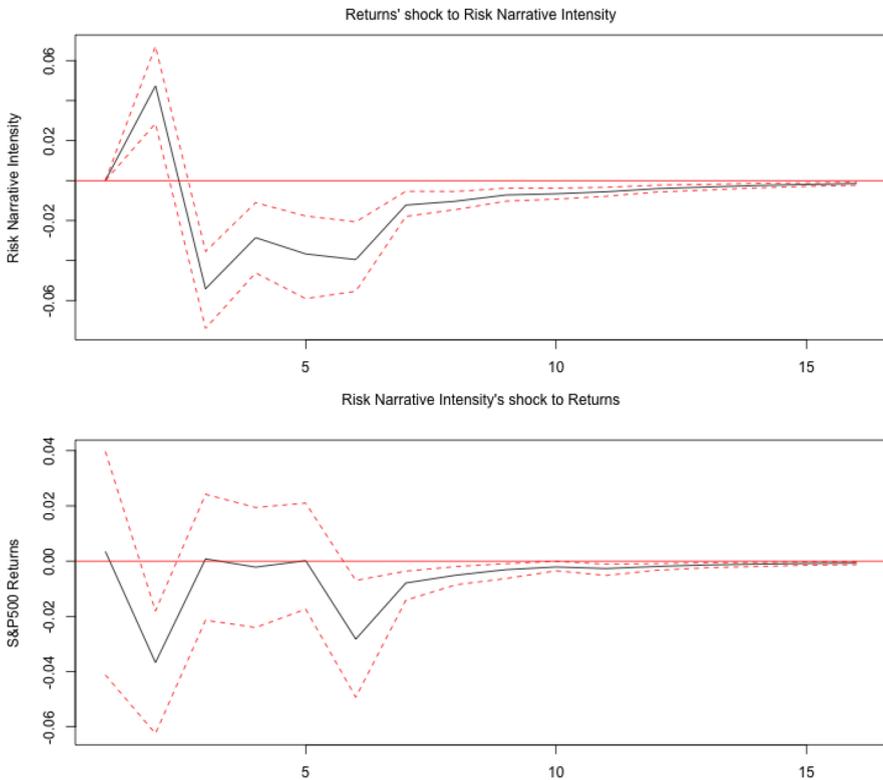
We further apply our modified epidemiological model on the daily risk narrative intensity measure to estimate its virality. The narrative intensity and virality measures are plotted in Fig. 11. Both indices spikes around historical periods



**Fig. 11** Fitted virality with the adjusted SIR model in the risk narratives. Notes: this figure plots the evolution of infected population (upper panel) and effective reproduction number (lower panel) of the risk narratives epidemic, estimated using the adjusted SIR model. On the upper panel, the dots represent the true observations of the infected population  $I$  and the blue line represents the estimated curve of  $I$ . On the lower panel, the dark curve represents the estimated values of  $R_t$

with high market volatility, including the 1987 market crash, the dot-com bubble, and the global financial crisis. Given that the risk narratives become viral only during bad times, it is likely to be a perennial economic narrative, a concept introduced in Shiller (2019). The risk narrative epidemic in 2020 is triggered by the world-wide outbreak of COVID-19, indicating that the narrative became viral again through its mutation to including the association between negative economic outlook and the development of the COVID-19 pandemic.

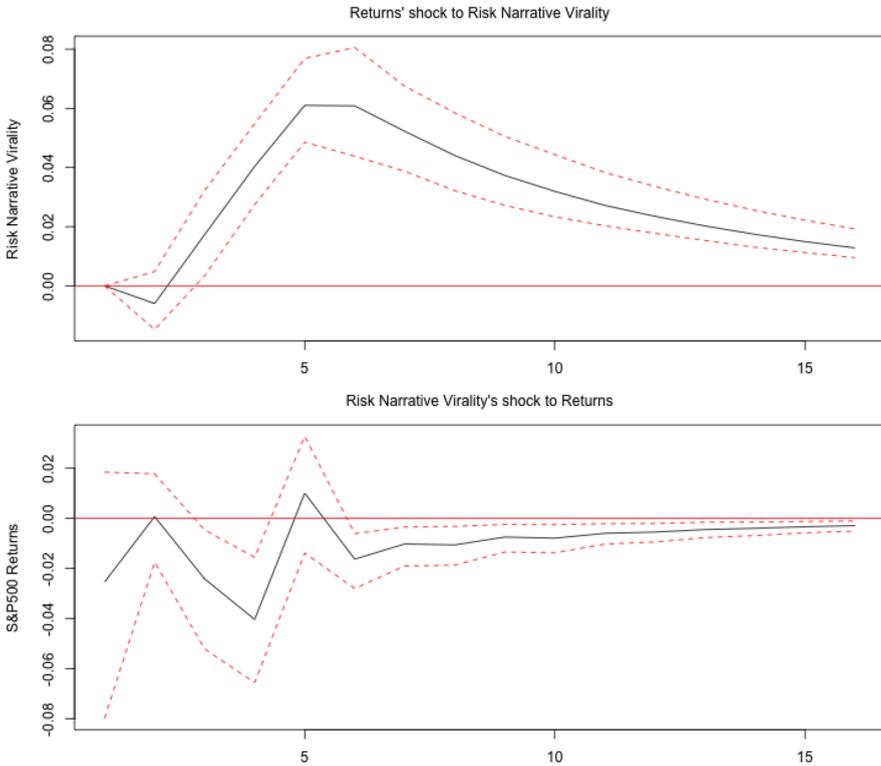
With the VAR model, following Tetlock (2007), we test the intertemporal links between the risk narratives and the stock market. We adopt the model only on the intensity and the virality measures here. The narrative is identified as a more specific narrative, there is no need to consider its sub topic (topic distribution and topic consensus). The narrative intensity measure also conveys sentiment information because this narrative itself has stable negative sentiment. The results are reported in Tables 5 and 6. Interestingly, we find a two-way Granger-causal relationship between stock market returns and the virality measure of risk narratives, and a one-way Granger-causal relationship from market returns and volume to narrative



**Fig. 12** Impulse response of the VAR with the risk narrative intensity. Notes: this figure plots impulse responses of one standard deviation shock of S&P500 return and risk narrative intensity from a VAR model. This VAR model is formed with three endogenous variables in the order of risk narrative intensity, detrended log of trading volume and S&P500 returns. The exogenous variables include five lags of VIX, dummy variables of weekdays and January. The horizontal axis represents the lags

intensity. The impulse response functions are plotted in Figs. 12 and 13. Based on Fig. 12, a return’s shock causes an immediate increase of risk narrative intensity on the next day, but causes a permanent decrease since the third day. On the other hand, a risk narrative intensity’s shock causes a quick decrease of returns on the second day and a decrease of returns for about two trading weeks. Based on Fig. 13, a return’s shock causes a permanent increase of risk narrative virality. A shock of risk narrative virality, on the other hand, causes a permanent decrease of market returns since the third day and last for three trading weeks. This results support the power of the text-based risk narratives and the potential of our virality measure, that is, the risk narratives likely contain fundamental information that has not been priced.

In summary, the risk narrative contains information about both market volatility and market values. This narrative went viral during early stage of the COVID-19 pandemic. Its intensity is impacted by implied volatility in the short term, but has a much longer term (15 months) impact on implied volatility, implying a “feedback loop” effect. The



**Fig. 13** Impulse response of the VAR with the risk narrative virality. Notes: this figure illustrates the impulse responses to one standard deviation shock of S&P500 return and risk narrative virality from a VAR model. This VAR model is formed with three endogenous variables in the order of risk narrative virality, detrended log of trading volume and S&P500 returns. The exogenous variables include five lags of VIX, dummy variables of weekdays and January. The horizontal axis represents the lags

intertemporal relationship between the risk narrative and stock market returns implies that, this narrative is Granger-caused by market returns, but can predict market return in the long term. The risk narrative can be used as a long-term early warning signal for stock market conditions.

## 6 Conclusion

The COVID-19 pandemic provides a natural experiment for narrative studies in finance. In this paper, we characterize the properties of the COVID-19 narratives and their changes around a severe equity market decline. By examining a set of interpretable features, we show that the COVID-19 narratives' evolution is consistent with the findings and proposals in the narrative literature. Around the period of the severe market decline, the pandemic-relevant narratives went viral and dominant, but also fragment into competing topics, became more pessimistic

and uncertain, and were more linked to the “recession” and “crisis” context. The features of the event narratives can be used to monitor the general beliefs in the market. A VAR-based analysis shows that the coronavirus narratives circulating in financial markets are impacted by the markets. Future studies can run cross-sectional analysis to examine their potential as early warning signals of drastic market volatility.

Though a creative narrative selection workflow, we successfully identify a perennial risk narrative, which went viral during the early stage of COVID pandemic, and all other periods with significant market downturns. The risk narratives went viral since the COVID-19 outbreak in Italy, when the world realized the potential damage of the pandemic. At the same time, the U.S. stock market tumbled heavily. From our Granger-causality-based tests, we find that the risk narratives can predict VIX and stock market returns in the long term, implying that there is new fundamental information contained in the narratives. Our findings thus encourage the use of narratives to evaluate long-term market conditions and to early warn event-driven severe market declines.

## Appendix 1: the adjusted SIR model

According to Kermack and McKendrick (1927), in a region with a constant population ( $N$ ) and interaction rate among the individuals, after the emergence of the first case of disease, the population can be divided into three groups, namely susceptible population ( $S$ ), the part of population who have not had the disease and are vulnerable to it; infective population ( $I$ ), the portion of population who have caught the disease and are infectious; and recovered population ( $R$ ), who have recovered from the disease and have acquired permanent immunity. Through contagion and recovery, the numbers of people in the three groups will change overtime with the dynamics described by three key difference equations:

$$dS/dt = -\beta SI \quad (7)$$

$$dI/dt = \beta SI - \gamma I \quad (8)$$

$$dR/dt = \gamma I. \quad (9)$$

Here, after normalizing with  $N$ ,  $S$ ,  $I$  and  $R$  represent the percentage of the population who are susceptible, infective and recovered (and have immunity), separately. The contagion rate  $\beta$  defines the expected portion of new infection from the total number of interactions between the infected and susceptible population,  $SI$ . The recovery rate  $\gamma$  determines the speed of recovering. The basic reproduction number ( $R_0$ ) is defined as the number of new cases that are, on average, expected to be infected by a single infected individual before a widespread immunity and before any attempt or measure has been made at immunization or spreading. It is simply estimated as

$$R_0 = \beta/\gamma. \quad (10)$$

The contagion rate  $\beta$  must exceed the recovery rate  $\gamma$  (a basic reproduction number larger than 1) for an epidemic to initiate and continue. The curve will be flatter or narrower if the  $R_0$  are smaller.

Another version of reproduction number, also consider the susceptible population ratio, is more useful during an epidemic. It is called effective reproduction number and usually denoted as  $R_t$ . As shown below, it is simply the product of the basic reproduction number and the susceptible population ratio at time  $t$ . The relationship shows that the effective reproduction of the virus is scaled down by the susceptible population ratio:

$$R_t = R_0 S_t \quad (11)$$

The basic model has been widely used in epidemiology literature and was suggested by Shiller (2019) to be able to potentially model narrative spreading. Cinelli et al. (2020) use the basic SIR to model the COVID-19 social media infodemic. We will model the infodemics with the basic SIR model but address the re-occurrence pattern with some modifications. The key variable of our interest is the effective reproduction number  $R_t$ .

We adjust the constant variable model by allowing the parameters to vary.

First, detect the turning points—where the infected population ratio starts increasing again from a decreasing trend.

Second, estimate  $\beta$  and  $\gamma$  for the period before the first turning point.

Third, estimate  $\beta$  and  $\gamma$  for the post-mutation periods with the adjusted  $S$ ,  $I$ ,  $R$  initial values.

```

beta ,gamma = zeros(T,1)
S,I,R = zeros(T,1)
# Step 1
turning_points , turning_scales = find_turning_points(I_true[:])
# Step 2
[beta0 ,gamma0] = SIR(I_true[:turning_points[0]])
beta ,gamma,S,I,R =
update_parameters(beta0 ,gamma0,I_true[0])
# Step 3
for t in range(len(turning_points)):
    try:
        post_mutation_I =
            I_true[turning_points[t]:turning_points[t+1]]
    except:
        post_mutation_I = I_true[turning_points[t]:]
    adj_S , adj_I , adj_R =
    ADJ_initials(S[-1],post_mutation_I[0],R[-1],turning_scales[t])
    [beta1 ,gamma1] = SIR(post_mutation_I , adj_S , adj_I , adj_R)
    beta ,gamma,S,I,R =
    update_parameters(beta1 ,gamma1, adj_S , adj_I , adj_R)
R0 = beta./gamma
Rt = R0.*(1-I_true-R)

```

In the pseudo-program,  $T$  means the number of total observations; SIR is the function to estimate beta and gamma using the observed infected population ratio series and initial values for all three population ratios; find\_turning\_points is the function defined to find the turning points and their scales of turning in a time series; ADJ\_initials is a function defined to adjust the three population ratio at turning points based on the scale of turning, the actual  $I$  and the last values of the previously estimated  $sS$  and  $R$ ; update\_parameters is a function for updating the beta, gamma and the  $S, I, R$  series for each period. We need to decide how to find the turning points and decide the scale of adjustment of  $R$  and  $S$  for each turning point. We smooth the  $I$  series and find the turning points based on whether the smoothed  $I$  change from a decreasing trend to an increasing trend. We assume the scale of  $S$  adjustment should depend on the angle of the turn around the turning point. We start from a simple linear relationship as follows:

$$\Delta R(t) = \hat{R}(t) \left( \frac{1}{1 + e^{-\text{angle}}} \right) + \epsilon_t \tag{12}$$

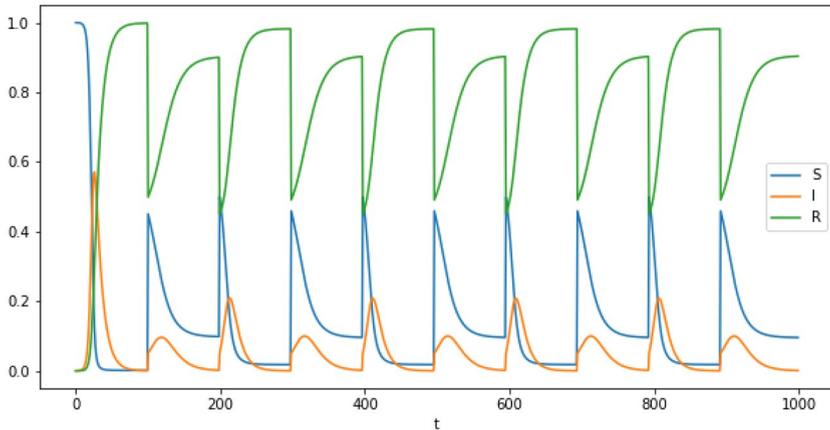
$$\text{angle} = (|g_{\text{post}}| + |g_{\text{pre}}|) / 2. \tag{13}$$

The adjustment of  $S$  around turning points depends on the current observed  $I(t)$ , the estimated  $R(t)$ <sup>26</sup> and the scale of the raw change of the growth rate of  $I(t)$  between time  $t + 1$  and  $t$ , that is the difference between the growth rate before the turning point,  $g_{\text{pre}}$ , and the growth rate after the turning point,  $g_{\text{post}}$ . Specifically,  $\hat{S}(t + 1) = (1 - I(t)) \left( \frac{\hat{S}(t + \Delta R)}{\hat{S}(t) + \hat{R}(t)} \right)$ . The angle is estimated as the average of absolute slopes before and after each point in the sample period. We use a sigmoid function to evaluate the scale of turns, so that the maximum of adjustment would be equal to the current  $R$ . It implies that when there are extreme narrative mutations making the narrative very attractive again, large portion of the recovered population become susceptible again. This mechanism is mimicking the re-occurring patterns in narratives.

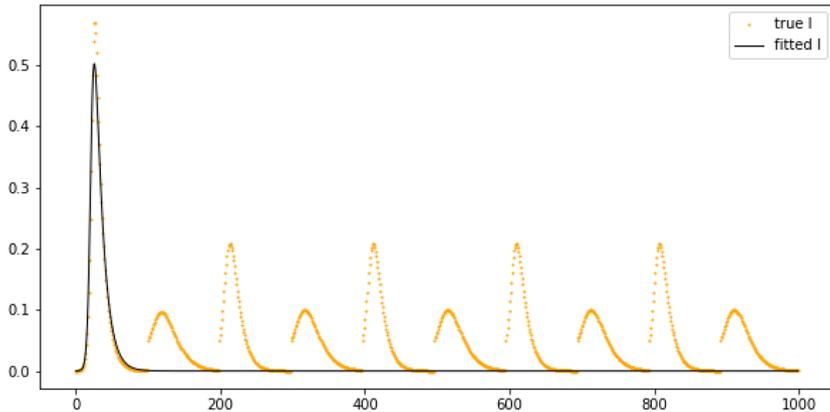
Time-fixed  $\beta$  and  $\gamma$  parameters cannot accurately reflect the narrative spreading dynamics when re-occurrences are present. In a simple simulation, we show the advantages of our adjusted SIR model fitting.

In the data generating process, we set the initial infected population ratio  $I(0)$  as 0.0001 and initial recovered population ratio  $R$  as zero. The  $S(0)$  is then calculated as  $1 - I(0) - R(0)$ . We set 10 regimes within a period of 1,000 time units. For regimes with an odd order, we set  $\beta$  as 1.2, and other regimes (with even orders) have  $\beta$  equals 0.75. All regimes have the same  $\gamma$ , which equals 0.1. The fixed  $\gamma$  is set to reflect our assumption that people forget or lose interest to a narrative constellation with similar speed. Whenever there is a regime change, we set the infected population to be 0.05 (a sudden increase of the infected) and the recovered population to be half of the last observation (half of the recovered population become susceptible again). The simulated infodemic is plotted in Fig. 14. Simply applying

<sup>26</sup> Please note that  $R(t)$  here means the recovered population ratio, instead of the effective reproduction number  $R_e$ .

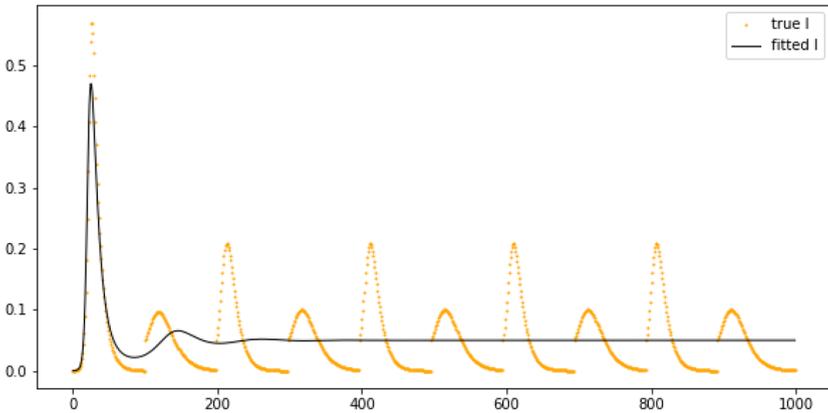


**Fig. 14** Simulated epidemic with 9 re-occurrences. This figure shows the evolution of the simulated epidemics. In the data-generating process, we set the initial infected population ratio  $I(0)$  as 0.0001 and initial recovered population ratio  $R$  as zero. The  $S(0)$  is then calculated as  $1 - I(0) - R(0)$ . We set 10 regimes within a period of 1,000 time units. For regimes with an odd order, we set  $\beta$  as 1.2, and other regimes (with even orders) have  $\beta$  equals 0.75. All regimes have the same  $\gamma$ , which equals 0.1. The fixed  $\gamma$  is set to reflect our assumption that people forget or lose interest to a narrative constellation with similar speed. The blue line represents the susceptible population ratio  $S$ ; the orange line represents the infected population ratio  $I$ ; and the green line represents the recovered/removed population ratio  $R$



**Fig. 15** Fitted  $I$  with the basic SIR model. Notes: this figure plots the evolution of the simulated epidemics, estimated using the basic SIR. The orange points represent the true observations of the infected population  $I$  and the dark line represents the estimated curve of  $I$ .

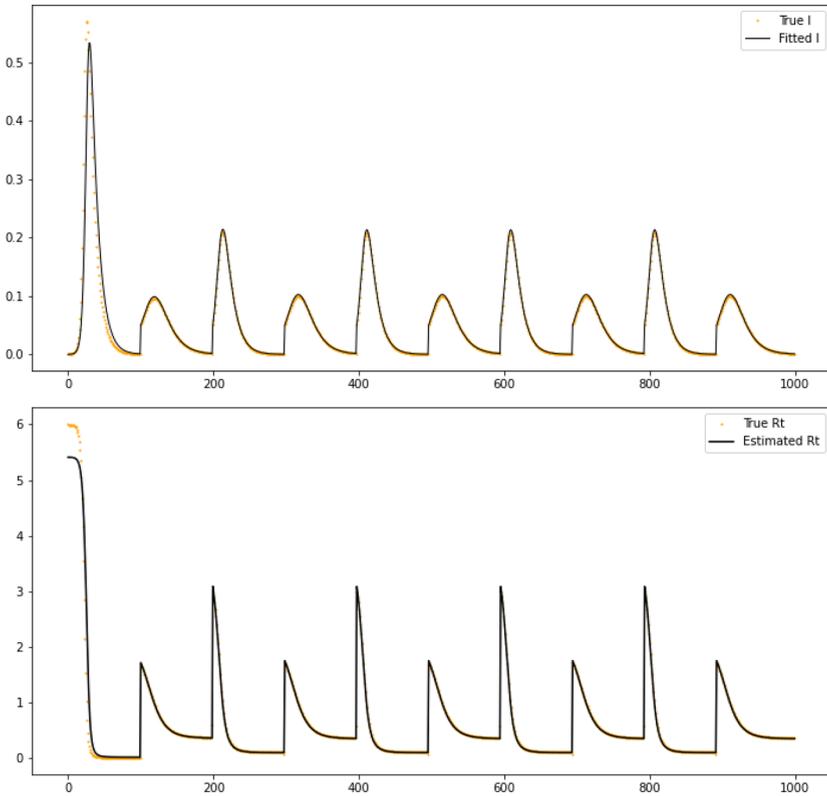
the basic SIR to fit the data, the estimated evolution of  $I$  series is plotted in Fig. 15. When we apply a SIR allowing part of the recovered population transfer to  $S$  again, the estimated evolution of  $I$  series is plotted in Fig. 16. Finally, using the method proposed, the estimated evolution is plotted in Fig. 17. The comparison shows that a simple SIR model can only work for early stage and will lose the ability to predict



**Fig. 16** Fitted  $I$  with the SIR model allowing re-infection. This figure plots the evolution of the simulated epidemics, estimated using the basic SIR allowing part of the recovered population become susceptible again. The orange points represent the true observations of the infected population  $I$  and the dark line represents the estimated curve of  $I$

the trajectory of infected population ratio as multiple re-occurrences become present. A modified SIR model allowing part of the recovered population to become susceptible again can partly reflect the re-occurrence patterns in a narrative epidemic; however, with fixed parameters, the estimated trajectory is likely to flatten out in the long term. Our adjusted SIR fitting method, on the contrary, has a performance in reflecting the true patterns. Fig. 17 shows that the adjusted SIR can reflect the effective reproduction number very well. The performance is robust to different regime settings in the simulation.

It is theoretically possible to also allow the adjusted SIR to detect turning points from a slow growing stage to a fast growing stage, however, it is practically difficult to differentiate the turning points with the points around highly curved trajectory due to high  $R_t$ . For this reason, in this version of paper, we assume there is no such mutations suddenly increase contagion. To apply this method in narrative series, we need to first smooth the data for turning points detection. Without smoothing the series, we might detect some “turning points” caused by high frequent fluctuations. In our COVID narratives, we find that the ratio of news discussing COVID is lower during weekends.



**Fig. 17** Fitted  $R_t$  with the adjusted SIR model in simulation. Notes: this figure plots the evolution of the infected population (upper panel) and effective reproduction number (lower panel) of the simulated infodemics, estimated using the adjusted SIR model. On the upper panel, the orange points represent the true observations of the infected population  $I$  and the dark line represents the estimated curve of  $I$ . On the lower panel, the orange points represent the true values of the effective reproduction number  $R_t$  and the dark line represents the estimated values of  $R_t$

## Appendix 2: snippets of risk narratives

### Article 1

Date: February 26th 2020; Topic probability: 60.6%

“When the stock market shudders a few days in a row, it is tempting to do something, and fast. Read this first.”

“When the stock market falls, as it has [missing word] all week, there is a natural desire to search for explanations and consult crystal balls. While **the spread of the coronavirus has been a catalyst, nobody knows exactly why the market moved the way it did, including whether underlying economic troubles are contributing to the severity of the gyrations.** And nobody can tell you how the

virus will affect the United States, including whether the outbreak's short-term economic impact will reduce long-term profits.”

“**Bull markets generally don't last this long. Maybe this is the end**, and if it is, so be it. Stocks will fall 10% and flatline for a while. Or maybe they drop 20%. Or 40. But wherever they bottom out and start rising again is likely to be higher than the average price you paid for your stocks over the past 10 or 15 years.”

## Article 2

Date: 21st March 2020; Topic probability: 70.1%

“There is a lot of the characteristics that we've witnessed so far already that are very much **end-of-crisis things**,” Paulsen said. “Things that you typically get at the end of the panics, as opposed to its start.”

“Ask the question, what are the odds of where we might be a year from now,” suggested Paulsen. “To me, I'm **fairly confident it will be higher than it is right now**, even if it goes lower for a while first.””

“It's really the speed of this decline that is interesting,” said Sean Naughton, senior vice president of U.S. equities at RBC Wealth Management, who sees some resemblances to **the market crash in 1987** and others.”

“Market observers are looking to medical professionals to help unwind some of the biggest questions about this market downturn, which has been centered around **the worldwide spread of the virus**.”

“There is a possibility that there is a lot of information we don't know about the spread of this virus,” Naughton said. “I think it's going to be critically important for us to get comfortable that there is some control mechanisms on the pandemic.””

“One thing that is much different this time is that the banks are in much better position,” Naughton said. The annual financial stress tests that banks have had to do since the **financial crises in 2008 and 2009** has them in a much better position than they were heading into that crisis. “That could help soften the blow if some of these worse case scenarios come out,” Naughton added.”

“Long-term winners in the stock market generally have to go against the grain at some point. Some are still looking at this volatile market as an opportunity. “**The best opportunities** in my entire investment career is the middle of panics like this,” Paulsen said.”

## Article 3

Date: 20th October 1987; Topic probability: 72.6%

“But all investors the world over have been warned that this bull market, like all bull markets in the past, will someday come to an end. And the chances are that the triple jolt has been enough to cause foreign investors to hesitate before providing more loans to sustain America's spending spree of the '80s.”

“The triple jolt does one other thing. It calls attention to the relationship between **the stock market crash of 1929** and the Great Depression that followed it. That single and now immensely relevant fact is that the American economy was doing reasonably well when the stock market crash came and did not sink into disastrous depression until late in 1932.”

#### Article 4

Date: 25th October 1987; Topic probability: 72.0%

“**THE economy will never be the same.** No matter where the stock market heads in coming months, the whole process of doing business, making a living, choosing investments—even buying a house—has been radically transformed. Last Monday’s market collapse was one of those pivotal events that shape everything that comes afterward.”

“All last week, the collapse of the market, and its breathtaking ups and downs, were the national obsession. There had been warnings about economic danger signs before, but Wall Street’s free-fall delivered the message more clearly than all the talk: **Good times can end.** For two generations of Americans who know about economic disasters only from history books, the collapse of ’87 marks the end of an era of economic optimism. Even if the economy stays on a basically steady course, as most experts predict, the tumble was “a little like seeing the atom bomb go off,” said Richard R. West, dean of New York University’s Graduate School of Business. “Once you’ve seen it, you know it could happen again. The world has changed.” ”

#### Article 5

Date: 26th October 2008; Topic probability: 89.9%

“If he were alive today, former Vice President Spiro T. Agnew would tell you in no uncertain terms how hard it is to keep the nattering nabobs of negativism down. He would warn you to be wary of the hopeless, hysterical hypochondriacs of history.”

“On the way up, this **irrational exuberance** is all well and good as long as there is a ready inventory of fools willing to pay more than the last fool. When the supply of foolish buyers runs out, the issue becomes the supply of fools who want to sell for less than the last guy, which is where we find ourselves today.”

“Friday’s 4% slide in the Dow Jones industrial average left the market barometer down 5% for the week, 27% since Sept. 1 and 37% for the year.”

“**“Stock market bubbles and crashes are caused by herd mentality,”** Dr. Howard said. “It’s scary to me because we make our own heaven, and we make our own hell.” ”

“The herd mentality is more fearful in a falling market, says George Loewenstein, a Carnegie Mellon University business professor who specializes in the emerging field of behavioral economics. That’s because individuals start fearing the economic consequences of the market sell-off: whether they will have a job; whether they should reduce spending; whether they’ll be able to make loan payments or get a loan to buy a car.”

““It’s going to be **terrible**,” he said. “People are **either uncertain or afraid**. Those two moods ... do not bode well.” ”

““Part of the reason the market is so volatile now is that we’re in unbelievably uncharted territory. **No one seems to know what’s going on**,” he said. “Even business people are having a hard time wrapping their minds around it.””

## Article 6

Date: 19th August 2007; Topic probability: 82.1%

“We all understand that panic selling traditionally is an ally of the stock buyer. Dramatically plunging prices ultimately will lead to a major low point, as everyone gnashes their teeth, trying to gauge “the” time to move. Unfortunately, while we all may understand that bottoms are carved out when volume surges and prices are in a free fall, **it’s always tough to guess when enough is enough**.”

“Then there are the emerging markets: China, India, and others. The associated countries’ economies are growing at a phenomenal pace, and many analysts believe that expansion will continue into the next decade. Perhaps that’s true, but buying before panic is ripe can still be a painful process. Remember, the United States was on the verge of **a fabulous century in the 1930s**, but that didn’t stop the Great Depression.”

“No one doubts that the current breakdown will end any differently than ever before in the history of markets. When the sellers exhaust their supply, the market will turn. Watch for enormous volume, exasperation, despair and then capitulation. Maybe the news will be worse; that’s likely. Maybe the news stories will be dismal; that’s likely. Maybe everyone will throw in the towel in unison; that’s likely. In any event, the market will turn at some point; that’s likely.”

## Declarations

**Conflict of interest** The authors have no competing interests to declare that are relevant to the content of this article.

## References

- Ardia, D., Bluteau, K., & Boudt, K. (2019). Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, 35(4), 1370–1386.

- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593–1636.
- Bertsch, C., Hull, L., & Zhang, X. (2021). Narrative fragmentation and the business cycle. *Economics Letters*, 201, 109783.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Bodnaruk, A., Loughran, T., & McDonald, B. (2015). Using 10-k text to gauge financial constraints. *Journal of Financial and Quantitative Analysis*, 50(4), 623–646.
- Cinelli, M., Quattrocioni, W., Galeazzi, A., Valensise, C.M., Brugnoli, E., Schmidt, A.L., Zola, P., Zollo, F., & Scala, A. (2020). The COVID-19 social media infodemic. 10(1): 1–10, arXiv preprint [arXiv:2003.05004](https://arxiv.org/abs/2003.05004).
- Egami, N., Fong, C.J., Grimmer, J., Roberts, M.E., & Stewart, B.M. (2018). How to make causal inferences using texts. arXiv preprint [arXiv:1802.02163](https://arxiv.org/abs/1802.02163).
- Eliasz, K., & Spiegler, R. (2020). A model of competing narratives. *American Economic Review*, 110(12), 3786–3816.
- Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, 64(5), 2023–2052.
- Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, 68(3), 1267–1300.
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267–297.
- Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, 33, 171–185.
- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London.*, 115(772), 700–721.
- Larsen, V. H., Thorsrud, L. A., & Zhulanova, J. (2021). News-driven inflation expectations and information rigidities. *Journal of Monetary Economics*, 117, 507–520.
- Liu, S., & Han, J. (2020). Media tone and expected stock returns. *International Review of Financial Analysis*, 70, 101522.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1), 35–65.
- Mikolov, T., K. Chen, Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint [arXiv:1301.3781](https://arxiv.org/abs/1301.3781).
- Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. In: *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, p. 262–272.
- Mostafazadeh, N., Chambers, N., He, X., Parikh, D., Batra, D., Vanderwende, L., Kohli, P., & Allen, J. (2016). A corpus and cloze evaluation for deeper understanding of commonsense stories. In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, p. 839–849.
- Nimark, K. P., & Pitschner, S. (2019). News media and delegated information choice. *Journal of Economic Theory*, 181, 160–196.
- Nyman, R., Kapadia, S., & Tuckett, D. (2021). News and narratives in financial systems: exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, 127, 104119.
- Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11. [arXiv:1908.10084](https://arxiv.org/abs/1908.10084).
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An r package for structural topic models. *Journal of Statistical Software*, 91(1), 1–40.
- Romero, D.M., Meeder, B., & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In: *Proceedings of the 20th International Conference on World Wide Web*, p. 695–704.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423.
- Shiller, R. J. (2015). *Irrational exuberance: Revised and expanded* (3rd ed.). Princeton University Press.
- Shiller, R. J. (2017). Narrative economics. *American Economic Review*, 107(4), 967–1004.
- Shiller, R. J. (2019). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

- ter Ellen, S., Larsen, V. H., & Thorsrud, L. A. (2019). *Narrative monetary policy surprises and the media*. Norges Bank.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139–1168.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Yuting Chen<sup>1,2</sup>  · Don Bredin<sup>1</sup> · Valerio Potì<sup>1</sup> · Roman Matkovskyy<sup>2</sup>

<sup>1</sup> College of Business, University College Dublin, Blackrock, Dublin, Ireland

<sup>2</sup> Rennes School of Business, Rennes, France