Business cycle narratives∗

Vegard H. Larsen† Leif Anders Thorsrud‡

This version
February 12, 2018

Abstract

This article quantifies the epidemiology of media narratives relevant to business cycles in the US, Japan, and Europe (euro area). We do so by first constructing daily business cycle indexes computed on the basis of the news topics the media writes about. At a broad level, the most influential news narratives are shown to be associated with general macroeconomic developments, finance, and (geo-)politics. However, a large set of narratives contributes to our index estimates across time, especially in times of expansion. In times of trouble, narratives associated with economic fluctuations become more sparse. Likewise, we show that narratives do go viral, but mostly so when growth is low. While narratives interact in complicated ways, we document that some are clearly associated with economic fundamentals. Other narratives, on the other hand, show no such relationship, and are likely better explained by classical work capturing the market’s animal spirits.

JEL-codes: C55, E32, E71, N10

Keywords: Business cycles, Narratives, Dynamic Factor Model (DFM), Latent Dirichlet Allocation (LDA)

∗This Working Paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. We thank Knut A. Aastveit, Andre K. Anundsen, Drago Bergholt, Hilde C. Bjørnland, Gunnar Bårdsen, Francesco Furlanetto, and colleagues at BI and Norges Bank for valuable comments. We are grateful to the Dow Jones Newswires Archive for sharing their data with us for this research project. This work is part of the research activities at the Centre for Applied Macro and Petroleum Economics (CAMP) at the BI Norwegian Business School.

†Centre for Applied Macro and Petroleum Economics, BI Norwegian Business School, and Norges Bank. Email: vegard-hoghaug.larsen@bnorges-bank.no

‡Corresponding author. Centre for Applied Macro and Petroleum Economics, BI Norwegian Business School, and Norges Bank. Email: leif-anders.thorsrud@norges-bank.no. Norges Bank, P.O. Box 1179, Sentrum 0107, Oslo, Norway.
1 Introduction

In his presidential address before the American Economic Association’s 2017 meeting, Professor Robert J. Shiller writes:

“The human brain has always been highly tuned toward narratives, whether factual or not, to justify ongoing actions,... Narratives “go viral” and spread far, even worldwide, with economic impact...Though these narratives are deeply human phenomena that are difficult to study in a scientific manner, quantitative analysis may help us gain a better understanding of these epidemics in the future.” (Shiller (2017))

This article quantifies the epidemiology of narratives relevant to economic fluctuations, and business cycles in particular, by asking: To what extent are narratives informative for describing business cycle variation, do they go viral, how do they interact with each other, and are they associated with economic fundamentals or better understood as capturing the market’s animal spirits?

To answer these questions, we restrict our attention to narratives told and spread through the mass media, and construct quantitative measures of narratives based on the news topics the media writes about. Shiller (2017) defines the term narrative to mean a simple story or easily expressed explanation of events that many people want to bring up on news. In Section 2 we discuss why the topic modeling approach provides a good quantitative approximation for narratives, while we in Section 3 describe how we technically construct the news topics and transform them into data useful for a time series analysis. We then proceed in four successive steps.

First, in Section 4, we present a daily coincident index model, built to capture aggregate business cycle dynamics, for three major economies; the US, Japan, and Europe (euro area). Unlike conventional models of this type (Stock and Watson (1988), Mariano and Murasawa (2003), Aruoba et al. (2009), and Marcellino et al. (2016)), however, the model allows for time-varying parameters through a threshold mechanism, and, most importantly, uses the daily narratives as input variables. In turn, this innovation allows us to decompose the changes in the latent daily business cycle indexes into time-varying news topic contributions reflecting the continuously evolving narrative about economic conditions, as described by the media. The resulting indexes and decompositions are reported in Sections 4.1 and 4.2.

Building on these results, in Section 4.3, we explore the extent to which narratives relevant for business cycles go viral and affect economic fluctuations and co-movement across borders. In the process we derive novel virality indexes, which provide quantitative and qualitative information about which narratives go viral, when, and for how long.
In Section 4.4 we investigate how narratives independently spread between economic regions. We do so by using the individual news topic time series, their estimated importance for describing business cycle fluctuations, and so called “Graphical Granger causality” modeling (Lozano et al. (2009), Shojaiie and Michailidis (2010)). This framework allows us to handle the high dimensionality of the problem, but also draw on graph theory to construct measures of node importance and centrality. More than providing a sophisticated analysis of the causal mechanisms underlying information diffusion, our analysis provides the first attempt of quantifying news spillovers relevant for economic fluctuations for the world’s largest economies.

Finally, in Section 4.5, we show that the complex network of news spillovers can be partitioned into (more or less) exogenous components, and thereby used to cast light on whether narratives are associated with economic fundamentals (Beaudry and Portier (2006), Barsky and Sims (2012), Blanchard et al. (2013)), or noise and sentiment (Shiller (2000), Angeletos and La’O (2013)).

Our analysis is explorative rather than grounded in one (single) formal model. We loosely take a rational inattention view (Sims (2003)), where news broadcasted through the media is important because it can reach a broad population of economic agents and alleviate informational frictions, but also potentially have an independent role in explaining economic fluctuations (Dougal et al. (2012), Peress (2014), Larsen and Thorsrud (2017), Shiller (2017)). We operationalize this view by working with a simple underlying hypothesis: To the extent that the media provides a relevant description of the economy, the more intensive a given topic is represented in the media at a given point in time, the more likely it is that this topic represents something of importance for the economy’s current and future needs and developments. For example, we hypothesize that when the media writes extensively about, e.g., regulatory developments, this reflects that something is happening in this area that potentially has economy-wide effects.

Key to our approach is that we use text as data (Gentzkow et al. (2017)), and our focus on news topics. From the Dow Jones Newswires Archive (DJ) we have access to over 40GB of news stories dating back to the early 1990s, covering all areas of economics, a range of countries and regions, and the Dow Jones flagship publication The Wall Street Journal.1 While the Dow Jones news service is far from the monopolistic supplier of economic news, it is among the three biggest suppliers in this global market. Thus, while we can not rightfully argue that we capture all economic news relevant for economic agents in all three countries, we believe the dataset is fairly representative.

1The term “Big Data” is used for textual data of this type because they are, before processing, highly unstructured and contain large amounts of words and articles (Nymand-Andersen (2016)).
literature, while the tone of the news is identified using simple dictionary based techniques (Tetlock (2007)). In general, topic models are statistical algorithms that categorize the corpus, i.e., the whole collection of words and articles, into topics that best reflect the corpus’s word dependencies. In this paper, an unsupervised topic model belonging to the Latent Dirichlet Allocation (LDA) class (Blei et al. (2003)) is used to estimate 80 topics for each country. Each individual topic can be viewed as a word cloud, where the font size used for each word represents how likely it is to belong to this specific topic. We subsequently transform these word clouds into tone adjusted frequency measures, reflecting by how much, and by which tone, each topics is written about on each day in the sample. A vast information set consisting of words and articles can thereby be summarized in a much smaller set of topics facilitating usage in a macroeconomic context. Although topic models hardly have been applied in economic (see, e.g., Hansen et al. (2018) for an exception), their use as a natural language processing tool in other disciplines has been widespread. The LDA’s popularity stems from its success in classifying text and articles into topics in much the same manner as humans would do (Chang et al. (2009)).

We reach five main conclusions. First, in all three countries/regions, the resulting coincident indexes are shown to track the phases of the business cycles with high precision, but performs especially well in the US. For policymakers and forecasters who need to assess the state of the economy in real time to devise appropriate policy responses, the news-based coincident indexes offer a valuable alternative. High-frequency economic statistics covering the broader economy are scarce. Daily news coverage is available in large quantities.

Second, we provide new evidence on the narratives relevant to economic fluctuations. At a broad level, particularly influential news topics include news about macroeconomic developments, the financial market, and (geo-)politics in all three countries. Across time however, there is considerable variation in how narratives contribute to, or describe, economic fluctuations. For example, late in 2007 and through 2008, news about regulatory developments is among the most influential news topics in the US, while earthquake-related narratives became particularly relevant in Japan in 2011. A common pattern across all countries, however, is that in periods associated with recessions, the number of narratives contributing to our index estimates become more sparse than during expansions. Thus, in relation to narratives, expansions are broad based, while recessions are not.

Third, we find that narratives do go viral, as argued by Shiller (2017), but mostly so in times of trouble. In total we identify 13 epidemic episodes between the mid 1995 and 2016, with an average duration of 4-5 months. The narratives contributing the most to these episodes tend to be associated with US-based labor market conditions and (partly)
monetary policy. Interestingly, however, we find little evidence suggesting that epidemics lead to more synchronized international business cycles.

Fourth, the graph describing the network of cross-country news spillovers is dense, but complex. Still, narratives identified with the US dominate, and have predictive power for news in Japan and Europe to a much larger extent than vice versa. The most central nodes in the graphical Granger causality graph are very much the same as those that contribute the most in explaining the fluctuations in the daily coincident indexes, i.e., news about macro economic developments and (geo-)politics, while the least central narratives are found to include news about technology, finance and commodities.

Finally, when partitioning the news topics into more or less exogenous variables using the centrality score computed from the graphical Granger causality graph, we find clear evidence that the most “exogenous” (least connected) narratives are associated with economic fundamentals (total factor productivity (TFP)). Unexpected fluctuations in these narratives lead to persistent, and significant, increases in TFP. In contrast, narratives with a high centrality score show no such relationship. Thus, some narratives confirm to the news-driven business cycle view. Other narratives, on the other hand, are likely better explained by classical work capturing the market’s animal spirits.

This article contributes to a broader literature that seeks to understand the role of narratives in economics (Shiller (2017)). To this end we establish a number of new “stylized facts” about the relationship between business cycles and narratives, epidemics, and cross-country spillovers for the three major economies the US, Japan, and Europe. As such, we also relate more loosely to a large literature investigating international business cycle synchronization (Kose et al. (2003), Stock and Watson (2005), Mumtaz et al. (2011) Kose et al. (2012)). In contrast to earlier studies in this literature, however, we are the first to focus on narratives. Likewise, by investigating the relationship between narratives and economic fundamentals we speak to a huge and long-lasting literature where changes in economic agents’ expectations, due to either news (new information) or animal spirits (noise/sentiment), are the main underlying driver of business cycle fluctuations (Pigou (1927), Keynes (1936), Beaudry and Portier (2006), Barsky and Sims (2012), Blanchard et al. (2013), Shiller (2000), Angeletos and La’O (2013)).

This paper is also directly related to a large literature, starting with Burns and Mitchell (1946), that seeks to measure business cycles and construct coincident indexes. Regarding the latter, Stock and Watson (1988), Mariano and Murasawa (2003), Aruoba et al. (2009), and Marcellino et al. (2016) provide prominent contributions, and Balke et al. (2017) and

---

2A closely related field, particularly in finance, investigates the (causal) role of the media itself (Dougal et al. (2012), Peress (2014)). Interestingly, in the first “Handbook of Media Economics” (Simon P. Anderson and Strömbäck (2015)) there is a separate chapter about “The Role of Media in Finance” (Tetlock (2015)), but no equivalent chapter about “The Role of Media in Macroeconomics”.

Shapiro et al. (2017) are examples of newer work using text as data. Neither of these studies do, however, provide a narrative account of business cycle fluctuations.

The approach taken here speaks to a growing number of studies in economics using text as data (Bholat et al. (2015), Gentzkow et al. (2017)). On this point, commonly used methods in economics involve some kind of subjectively chosen keyword search and auditing (Baker et al. (2016)), or narrative methods for shock identification Ramey (2016)). For uncovering the narratives relevant for economic fluctuations, the topic modeling approach offers a conceptual advantage over other often applied textual data techniques because it provides interpretable output in a highly automated fashion.3

Lastly, on the methodological side, we draw on recent advances presented in Larsen and Thorsrud (2018) for constructing time series measures of text, and the model proposed in Thorsrud (2016b,a) for constructing the coincident indexes. Both of these studies explore the relationship between news and economic fluctuations in Norway. Here we extend this line of research to three of the biggest economies in the world, and take a narrative perspective. Naturally, we provide a number of news results, and propose new tools for measuring the extent to which narratives go viral, cross-country spillovers, and whether or not narratives are associated with economic fundamentals or animal spirits.

2 On narratives

Humans are inherently storytellers, and the academic literature on narratives is vast. Most work, however, is not found in economic journals, but rather in fields related to linguistics, psychology, anthropology, and history. Here, as alluded to already, we follow Shiller (2017) and define the term narrative to mean a simple story or easily expressed explanation of events that many people want to bring up on news. We then construct measurable approximates to this definition based on the news topics the media writes about, and subsequently link those to economic fluctuations. Accordingly, we will be using the terms narrative and news (topic) interchangeably. More formally, the narrative of a story will consist of one or more news topics. To elaborate on why this approximation is reasonable, what it allows us to measure, and why it might fall short, we take inspiration from the well known cognitive psychologist Jerome Bruner, and in particular Bruner (1991).

First, our interest is not so much in how narratives as text are constructed, but rather how they operate as instruments of mind in the construction, or reflection, of reality.

3For studies that seek to uncover the economic relationships between more concretely defined events or concepts, like, e.g., political uncertainty or monetary policy shocks, a keyword/event search approach might be better suited. For capturing narratives relevant for aggregate business cycles, a keyword/event based approach is not suited unless the researcher knows apriori what to search for.
Obviously, our focus in centered on a narrowly defined aspect of reality, i.e., economic fluctuations, and our sources for constructing measurable narrative approximates are limited to textual news broadcasted through the media. Still, as noted by Shiller (2000); "Significant market events generally only occur if there is similar thinking among large groups of people, and the news media are essential vehicles for the spread of ideas".

We look upon narratives as time dependent, and accounts of events occurring over time. At the same time, “...the particulars of narratives are tokens of broader types” (Bruner (1991)). The modeling approaches adapted in this study reflect these views. As described in greater detail in Section 3, a news story is a weighted sum of different word distributions, i.e., topics. The particular topic composition of a given story, at a given point in time, might very well be unique, but the topic distributions that the narrative constitute are potentially shared by many other narratives. Likewise, to capture the time dependent nature of narratives, we allow the mapping between narratives and economic fluctuations to be time-varying (see Section 4).

However, we do not require the stories in news to be true. Rather, the narrative “truth” is “judged by its verisimilitude rather than its verifiability” (Bruner (1991)). In our setting this means that objective reporting (if that exists) and speculative news stories about market developments, or even news stories about events not happening (if such reporting exists), are all treated equal.

Finally, we take the view that there is only a loose link between the intentional states of a narrative, and the subsequent actions it might induce. Relatedly, the meaning of a story is not simply the sum of its partial expressions, and the interpretation of it will likely depend on the readers background knowledge and context. Admittedly, while neither of these effects are well captured by our approach, it is difficult to envision how quantitative analysis of aggregate economic fluctuation and narratives can fully encapsulate such effects.

3 Data

The main raw data used in this analysis consist of a long sample of daily news extracted from the Dow Jones Newswires Archive (DJ). In total we utilize an extraction of over 40GB of raw textual data in XML format from this historical database, which covers a large range of their news services, including content from The Wall Street Journal. All text is business-focused, written in English, and covers the US, the Asian, as well as the European market.

The data span the period 1990 to 2016, and includes almost 11 million news articles. Each article listed in the database comes with a number of meta data such as publication
time and region. To classify news as either US, Japan, or Europe specific, we rely on the
tags provided by DJ, and partition the dataset accordingly. After removing duplicates
and articles that only include updates of earlier published news, the resulting regional
data sets include 4754040, 682424, and 1969222 articles for the US, Japan, and Europe,
respectively. For all three areas the partitioned data sets end in 2016. For the US we
have news observations starting in 1990, while for Japan and Europe the start dates are
1994 and 1995, respectively.

Arguably, what we categorize as country-specific news relies on the DJ definitions, and
does not end up as three completely non-overlapping datasets (see Table 16 in Appendix
C). As news likely does not stop at the border, we do not find this especially problematic.
Another potential limitation is that we have to rely on the DJ region classification tag,
and do not use economic news published in region-specific media. As The Wall Street
Journal is the largest newspaper in the United States in terms of circulation, but likely
not in Japan and Europe in general, our raw data might be more representative for the
US, than for the two other areas.4

To make the textual data applicable for time series analysis, we proceed in three steps
illustrated in Figure 1. Technically, these are the same data processing steps as proposed
in Larsen and Thorsrud (2018). We provide a summary of the computations below. In
the interest of preserving space, technical details are relegated to Appendix C.1 to C.3.

3.1 Cleaning

The share size of the three datasets makes statistical computations challenging. However,
as is customary in the Natural Language Processing (NPL) literature, some steps are
taken to clean and reduce the raw dataset before estimation (Gentzkow et al. (2017)).
First, a stop-word list is employed. This is a list of common words not expected to have
any information relating to the subject of an article. Examples of such words are the, is,
are, and this. In total, the stop-word list together with the list of common surnames and
given names removed roughly 1800 unique tokens from the corpus. Next, an algorithm
known as stemming is run. The objective of this algorithm is to reduce all words to their
respective word stems. A word stem is the part of a word that is common to all of its
inflections. An example is the word effective whose stem is effect. Finally, a measure
called tf-idf, which stands for term frequency - inverse document frequency, is calculated.
This measures how important all the words in the complete corpus are in explaining single
articles. The more often a word occurs in an article, the higher the tf-idf score of that

4Obviously, for us, language barriers are a non-trivial friction in terms utilizing truly country-specific
media. Likewise, obtaining textual data of the size and coverage as here is costly. We are grateful to the
Dow Jones Newswires Archive for sharing their data with us for this research project.
word. On the other hand, if the word is common to all articles, meaning the word has a high frequency in the whole corpus, the lower that word’s tf-idf score will be. Around 150,000 of the stems with the highest tf-idf score are kept, and used as the final corpus.

### 3.2 Topic extraction

The “cleaned”, but still unstructured, datasets are decomposed into news topics using a Latent Dirichlet Allocation (LDA) model (Blei et al. (2003)). The LDA model is one of the most popular clustering algorithms in the NPL literature because of its simplicity, and because it has proven to classify text in much the same manner as humans would do (Chang et al. (2009)).

The LDA is an unsupervised topic model that clusters words into topics, which are distributions over words, while at the same time classifying articles as mixtures of topics. A unsupervised learning algorithm is an algorithm that can discover an underlying structure in the data without being given any labeled samples to learn from. The term “latent” is used because the words, which are the observed data, are intended to communicate a latent structure, namely the subject matter (topics) of the article. The term “Dirichlet” is used because the topic mixture is drawn from a conjugate Dirichlet prior.\(^5\)

\(^5\)As such, the LDA shares many features with latent (Gaussian) factor models used in conventional econo-
Different algorithms exist for solving the LDA model. We follow Griffiths and Steyvers (2004), and estimate the model using Gibbs simulations. Technical details and a short description of estimation and prior specifications are described in Appendix C.1. Here we note that we extract $K = 80$ topics from each of the three cleaned datasets. We subjectively chose $K = 80$ for two reasons. First, this was the choice showing the best statistical results in Larsen and Thorsrud (2018) and Thorsrud (2016b,a). Second, we have experimented with estimating both fewer and more topics. It is our experience that with $K$ substantially higher than 80, each topic starts to become highly event specific, i.e., there are signs of over-fitting. Conversely, extracting substantially fewer than 80 topics results in too general topics. Thus, in sum, our choice of $K = 80$ is based on a compromise between fitting the corpus well, getting interpretable topics, as well as earlier experience.

The LDA produces two outputs; one distribution of topics for each article in the corpus, and one distribution of words for each of the topics. Our primary interest is in the latter distributions, which are illustrated using word clouds in Figure 2. Now the LDA estimation procedure does not give the topics any name or label. To do so, labels are subjectively given to each topic based on the most important words associated with each topic. For example, as seen from Figure 2, the most important words associated with the US topic number $T_0$ are monetary, inflation, and bernanke. Thus, we label this topic Monetary Policy. While it is, in most cases, conceptually simple to classify the topics, the exact labeling plays no material role in the experiment, it just serves as a convenient way of referring to the different topics (instead of using, e.g., long lists of words). A full list of the different topics, their most important words, and our subjective labeling is given in Tables 9 to 11 in Appendix A.  

### 3.3 Topic time series

Given knowledge of the topics (and their distributions), the topic decompositions are translated into tone adjusted time series. To do this, we proceed in three steps described in detail in Appendix C.2 and C.3. In short, for each of the three cleaned datasets we first collapse all the articles for a particular day into one document, and then compute, using the estimated word distribution for each topic, the topic frequencies for this newly formed

---

6 The further improve the reader’s understanding of what the different topics are (and are not), we investigate, in Appendix B, how the topics relate to external texts freely available to the public.
Figure 2. Word clouds and topic categorization. For each word cloud the size of a word reflects the probability of this word occurring in the topic. Each word cloud only contains a subset of all the words in the topic distribution. Topic labels are subjectively given.

document. This yields a set of $K$ daily time series. Then, for each day and topic, we find the article that is best explained by each topic, and from that identify the tone of the topic, i.e., whether or not the news is positive or negative. This is done using an external
word list and simple word counts, similar to in Tetlock (2007). The word list used here classifies positive/negative words as defined by the Harvard IV-4 Psychological Dictionary. For each day, the count procedure delivers a statistic containing the normalized difference between positive and negative words associated with a particular article. These statistics are then used to sign-adjust the topic frequencies computed in step one. Finally, we remove high frequency noise from each topic time series by using a 60-day (backward looking) moving average filter, and, as is common in factor model studies (Stock and Watson (2012)), standardize the resulting series. Figure 8, in Appendix A, illustrates the resulting series for the 18 word clouds presented in Figure 2.

Notice from the description above that also the tone adjustment procedure explicitly uses the output from the topic model. Still, the method used for identifying the tone of the news using dictionary based techniques is simple, and could potentially be improved upon with more sophisticated algorithms (Pang et al. (2002)). While leaving such endeavors for future research, Thorsrud (2016b) shows that working with topic frequencies without tone adjustment results in a loss of important information.

4 Business cycle narratives

To link the daily news topics time series to aggregate economic fluctuations, we estimate a coincident index of business cycles utilizing the joint informational content in quarterly output growth and the daily news narratives using a Dynamic Factor Model (DFM). This approach builds on conventional models proposed in, e.g., Stock and Watson (1988), Mariano and Murasawa (2003), Aruoba et al. (2009), and Marcellino et al. (2016), and has two important characteristics. First, since our best measure of aggregate economic fluctuations, changes in Gross Domestic Product (GDP), is observed at the quarterly frequency, the aggregation from higher to lower frequency variables is handled using a cumulator variable approach (Harvey (1990), Banbura et al. (2013)). Second, to summarize the informational content in the large panel of variables in a parsimonious manner, a factor modeling approach is implemented.

The novelty of the DFM used here is that we include daily news variables instead of hard economic statistics as observable variables (in addition to GDP), but also that the model allows for time-varying parameters with a latent threshold mechanism. This model property is motivated by our narrative definition (see Section 2), enforces dynamic sparsity, and has also proven to be important for both forecasting and more structural interpretation in other high-dimensional settings (Zhou et al. (2014), Scott and Varian (2013), Thorsrud (2016b,a)).

We obtain GDP statistics, measured in constant prices, for the US, Japan, and Europe.
from the Federal Reserve Bank of St. Louis FRED database. The raw data is transformed into quarterly growth rates, and normalized. Then, a separate model is specified and estimated for each country. Following Thorsrud (2016b), and letting bold-font letters denote vectors and bold-font capital letters matrices, the DFM containing quarterly GDP growth and the daily news topic variables, can be written in a compact form as:

\[ y_t = Z_t a_t + e_t \]  
\[ a_t = F_t a_{t-1} + R_t \Sigma_t \omega_t \]  
\[ e_t = P e_{t-1} + u_t \]

with

\[ y_t = \begin{pmatrix} y_t^{k_q} \\ y_t^d \end{pmatrix} \quad \text{and} \quad a_t = \begin{pmatrix} a_t^{k_q} \\ a_t^d \end{pmatrix} \]

where \( t \) is the daily time index, \( k_q \) and \( d \) denote the quarterly and daily observation intervals, respectively, and the model has been written with simple autoregressive time series processes of order one for notational simplicity.\(^7\)

Equation (1a) is the observation equation of the system. \( y_t^{k_q} \) and \( y_t^d \) are \( N_q \times 1 \) and \( N_d \times 1 \) vectors of quarterly and daily variables, respectively, with \( N = N_q + N_d \). In this applications, \( N_q = 1 \) and \( N_d = K = 80 \). \( Z_t \) is a \( N \times N_a \) matrix with dynamic factor loadings linking the variables in \( y_t \) to the latent dynamic factors in \( a_t \), and are described in greater detail below. The vector \( e_t \) contains the idiosyncratic errors. It is assumed that these evolve as independent AR(\( p \)) processes given by (1c), where \( u_t \sim i.i.d. N(0, U) \).

Equation (1b) is the transition equation of the system. The common factors follow a VAR(h) process. \( \omega_t \sim i.i.d. N(0, I) \) and \( \Sigma_t \) is a diagonal matrix with \( \Sigma_t \Sigma_t' = \Omega_t \), allowing for stochastic volatility. The individual elements in \( \Sigma_t \) are assumed to follow random walk processes.\(^8\)

The last element in \( a_t \), the scalar \( a_t^d \), is interpreted as the latent common daily business cycle index. The other elements in \( a_t \), and in \( F_t \) and \( R_t \), contain cumulator variables used to handle the mixed-frequency property of the model. In the interest of brevity we describe the time aggregation procedure in Appendix D.7.

Dynamic sparsity is enforced on the system through the time-varying elements in \( Z_t \), which are modeled following the Latent Threshold Model (LTM) idea by Nakajima and West (2013). For one particular element in the \( z_t^d \) vector, \( z_{i,t} \), the LTM structure can be written as:

\[ z_{i,t} = z_{i,t}^* s_{i,t} \quad s_{i,t} = I(|z_{i,t}^*| \geq d_i) \]

\(^7\)The model can easily be generalized to include variables of other frequencies as well (see Thorsrud (2016b) for details).

\(^8\)While not explicitly discussed in this study, earlier studies show that allowing for stochastic volatility tend to improve the model performance in this type of DMFs (see, e.g., Thorsrud (2016a)).
where
\[ z_{i,t}^* = z_{i,t-1} + w_{i,t} \] (3)
with \( w_{i,t} \sim i.i.d. N(0, \sigma_{1,w}^2) \), and \( w_t \sim i.i.d. N(0, W) \) where \( W \) is a diagonal matrix. In (2) \( \varsigma_{i,t} \) is a zero one variable, whose value depends on the indicator function \( I(|z_{i,t}^*| \geq d_i) \).

If \(|z_{i,t}^*|\) is above the threshold value \( d_i \), then \( \varsigma_{i,t} = 1 \), otherwise \( \varsigma_{i,t} = 0 \).

The motivation for the LTM mechanism can easily be understood by an example. If the interpretation of narratives evolve and justify ongoing actions differently across time, or, if some narratives are more important in some periods than in others, a constant parameter model will fail. The researcher might simply conclude that a given narrative has no relationship with \( d_t \), i.e., that \( z_{i,t}^d \) equals zero for all time periods, because, on average, periods with a positive \( z_{i,t}^d \) cancels with periods with a negative \( z_{i,t}^d \). The LTM mechanism potentially captures such cases in a consistent and transparent way.

A more detailed description of the time-varying DFM model, and estimation, is given in Appendix D. Here we note that the DFM is estimated by decomposing the problem of drawing from the joint posterior into a set of much simpler ones using MCMC simulations, with prior specifications discussed in Appendix D.6.

For all specifications we allow for one lag in the equation for the idiosyncratic errors \( (p = 1) \), and up to ten lags for the latent common business cycle index \( (h = 10) \). The (full) estimation sample ends 31 December 2016 for all three countries. Due to data availability, estimation starts in 12 January 1990, 29 June 1994, and 1 July 1995 for the US, Japanese, and European model, respectively. Finally, we globally identify the sign and size of the latent factor by restricting the factor loading for the first element among the \( N_d \) variables to equal 1 for all time periods. We choose the normalizing variables by looking at the simple correlation between linearly interpolated output growth and the daily news topics. Accordingly, for the US, Japan, and Europe we use the Macroeconomics, Outlook, and Macroeconomics, news topics, respectively. Bai and Ng (2013) and Bai and Wang (2014) show that these restrictions uniquely identifies the factor and the loadings, but leaves the transition equation dynamics completely unrestricted.

4.1 The daily news-based coincident indexes

Figure 3 reports the estimated news-based coincident indexes for the US (NCI-US), Japan (NCI-Japan), and Europe (NCI-Euro). The gray shaded areas illustrate recession periods as defined by NBER (US), ECRI (Japan), and CEPR (euro area), while the black stars report observed quarterly GDP growth. In each graph we also report alternative existing state-of-the-art coincident index estimates. For the US, Japan, and Europe this is the

NBER is the National Bureau of Economic Research, ERCI is the Economic Cycle Research Institute, while CEPR is the Centre for Economic Policy Research. Of these, only the chronologies provided by the
Figure 3. $\Delta GDP^a$ is standardized output growth. It is recorded at the end of each quarter. The colored solid line is the standardized (median) estimate of the daily business cycle index, while the dotted colored lines are the 68 percent probability bands. The gray shaded areas illustrate recession periods as defined by NBER (US), ERCI (Japan), and CEPR (euro area).

daily $ADS$ index (Aruoba et al. (2009)), the monthly $CLI$ index (Eurostat), and the NBER and CEPR are regarded as representing official business cycle dates. To the best of our knowledge, no official business cycle dating committees exists for Japan.
monthly ECOIN index (Altissimo et al. (2010)), respectively.

By simple visual inspection we observe that the estimated news-based indexes track the state of the economies very well, and that results for the US seem to be especially good. The financial crisis is common for all indexes, while the recession in the early 1990s is US specific. Likewise, the two long downturns in the late 1990s and early 2000s are specific for Japan, while the troubled times following the Great Recession are partly shared by both Japan and Europe. In relation to this, it is interesting to observe the substantial increase in uncertainty associated with NCI-Euro in the periods following the financial crisis.\footnote{10}

To formally evaluate the models we use classification tests. Like in Travis and Jordà (2011), and in the tradition of Burns and Mitchell (1946), we categorize aggregate economic activity into phases of expansions and contractions and evaluate the indexes’ ability to classify such phases using Receiver Operating Characteristic (ROC) curves and area under the curve (AUROC) statistics. As measures of the unknown “truth”, we use the business cycle chronologies illustrated in Figure 3, i.e., the business cycle phases defined by the NBER, ERCI, and CEPR. Since these chronologies are available at a daily frequency only for the US economy, daily classifications are obtained by assuming that the economies remain in the same phase on each day within the monthly classification periods for Japan and the euro area.

Focusing on the AUROC statistics, Table 1 summarizes the business cycle classification scores, while Figure 10 in Appendix A reports the associated ROC curves. As a perfect classifier receives an AUROC of 1, we observe from the table that the NCI-US index is tracking the official NBER business cycle chronology very well. Also the NCI-Euro index is doing a reasonably good job at classifying the phases of economic fluctuations. The worst performing index, in terms of AUROC, is NCI-Japan, which receives a score of 0.76. Still, this is far better than random guessing, which would give an AUROC of 0.5.

To put the performance of the news-based indexes into perspective, we also evaluate the classification performance of the alternative state-of-the-art coincident indexes illustrated in Figure 3. Of these, only the ADS index is available on a daily frequency. For the monthly CLI and ECOIN indexes we construct daily analogs by assuming that every day within a month equals the observed monthly value. Again, Table 1 summarizes the results. In all three countries the existing indexes perform slightly better than the news-based indexes. However, the differences are not large, and at most 12 percent, for the euro area. In addition, the news-based indexes are available at a daily frequency, which

\begin{itemize}
  \item The time-varying changes in the variance of the NCI errors are illustrated in Figure 9 in Appendix A. Unexpectedly, all models pick up a substantially higher variance during the financial crisis episode than in other parts of the sample. Convergence statistics indicating that the MCMC algorithm has reached the ergodic distribution are discussed in Appendix F.
\end{itemize}
Table 1. Receiver Operating Characteristics and area under the curve (AUROC) statistics. By definition the AUROC can not exceed 1, perfect classification, or be lower than 0.5. We compute the AUROC score non-parametrically using the algorithm described in Travis and Jordà (2011).

<table>
<thead>
<tr>
<th></th>
<th>NCI-US</th>
<th>ADS</th>
<th>NCI-Japan</th>
<th>CLI</th>
<th>NCI-Euro</th>
<th>ECOIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUROC</td>
<td>0.946</td>
<td>0.996</td>
<td>0.760</td>
<td>0.790</td>
<td>0.853</td>
<td>0.969</td>
</tr>
</tbody>
</table>

the alternative indexes typically are not.

In sum, these results illustrate how informative the news-based approach is in terms of capturing economic fluctuations. For countries where high-frequency hard economic variables are not easily available, the news-based approach offers a valuable alternative. Moreover, in contrast to existing coincident indexes, the news-based approach gives the researcher, or index user, potential knowledge about the narratives important for understanding economic fluctuations. An issue we now turn to.

4.2 Business cycle decompositions

In this section we investigate “the epidemiology of narratives relevant to economic fluctuations” (Shiller (2017)). We do so by utilizing an attractive feature of the DFM modeling framework, namely that the state evolution of the model (the daily business cycle index(es)) can be decomposed into news surprises driven by the developments in the observable variables (the news topics). Technically, this is done using Kalman Filter iterations and decomposing the state evolution at each updating step into news contributions using the Kalman Gain (see Appendix E), and the recursive nature of the filter. Following Koopman and Harvey (2003), let:

\[ a_{t|t} = a_{t|t-1} + K_t v_t \]  

be the standard Kalman filter equation for updating the latent state estimate \( a_t \) given knowledge of the Kalman Gain matrix \( K_t \), with:

\[ a_{t|t-1} = F_t a_{t-1|t-1} \]

\[ v_t = y_t - Z_t F_t a_{t-1|t-1} \]  

Now, plugging (5) into (4) one obtains:

\[ a_{t|t} = F_t a_{t-1|t-1} + K_t (y_t - Z_t F_t a_{t-1|t-1}) \]

\[ = (I - K_t Z_t) F_t a_{t-1|t-1} + K_t y_t \]

11 Although the DFM model, with the LTM mechanism, is built to filter out uninformative data, it is very likely that a more elaborate data (pre)selection procedure could improve the results further. High frequency (hard) economic indicators can also be included into the model alongside the news topic variables. We leave such attempts for future research.
Table 2. Top 10 news topic (surprises). The ranking is based on sorting the output from equation (7) in descending order.

<table>
<thead>
<tr>
<th>NCI-US</th>
<th>NCI-Japan</th>
<th>NCI-Euro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor market</td>
<td>Outlook</td>
<td>Macroeconomics</td>
</tr>
<tr>
<td>Stocks</td>
<td>Motor</td>
<td>Middle East</td>
</tr>
<tr>
<td>Monetary policy</td>
<td>Financial companies</td>
<td>Trading data</td>
</tr>
<tr>
<td>Clients</td>
<td>Fed</td>
<td>Fiscal policy</td>
</tr>
<tr>
<td>Congress</td>
<td>Russia</td>
<td>Bonds</td>
</tr>
<tr>
<td>Regulations</td>
<td>Stock listings</td>
<td>Credit rating</td>
</tr>
<tr>
<td>Strategy</td>
<td>Market commentary</td>
<td>Nordic countries</td>
</tr>
<tr>
<td>Petroleum</td>
<td>Natural disasters</td>
<td>Australia</td>
</tr>
<tr>
<td>Education</td>
<td>Communication</td>
<td>Public safety</td>
</tr>
<tr>
<td>Market performance</td>
<td>Car technology</td>
<td>Investing</td>
</tr>
</tbody>
</table>

which can be inverted to obtain the moving average representation of the unobserved states as a function of the observed variables. Or, in other words, how the model interprets surprising news fluctuations when updating the state estimates.

Defining $w_{i,t} = K_{i,t}v_{i,t}$ as the weighted forecast error contribution from topic $i$ at time $t$, and:

$$w_i = \frac{1}{T} \sum_{t=1}^{T} (w_{i,t})^2$$

as the mean squared error, Table 2 reports the 10 most influential news topics on average across the sample. In general, news surprises about macro economic developments (e.g., Labor market, Macroeconomics and Outlook), the financial market (e.g., Stocks and Trading data), and (geo-)politics (e.g., Monetary policy, Fiscal policy, Congress, Middle East, and Russia) are important in all three countries. Still, constructing a story based on words drawn from the topic distributions summarized in the three columns in Table 2 would clearly result in three different narratives. For example, a grand narrative about Japan would be much more likely to contain topics related to the motor and car industry, and natural disasters, than a story for the US or euro area. Likewise, for a US-specific story, topics related to Petroleum and Regulations are likely much more prominent than in any of the other two countries.

Table 3, for the US, and Tables 13 and 14 in Appendix A, for Japan and the euro area, list the most influential narratives across six different sub-samples, as well the first sentences of particularly representative news articles during these time periods. While some of the same news topics tend to top the lists in every period, we observe a relatively large variation in the ranking of the other narratives. For example, during the period 1999-2002, topics associated with Internet and Persuasion are in the top of the list for the US,
Table 3. Top five news topics across sub-samples for the NCI-US index. The Example narratives are found by querying the corpus for news articles where the five news topics listed in column two combined receive a high weight. Only the first sentences of each story are included in the table. The date of publication is printed in parenthesis.

<table>
<thead>
<tr>
<th>Top 5 news topics</th>
<th>Story example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1995 - 1999</strong></td>
<td></td>
</tr>
<tr>
<td>Labor market</td>
<td>(1996-04-24) Western Germany’s consumer price index (CPI) is estimated to have risen a preliminary 0.2% in April from February and 1.5% from a year ago, a survey conducted by AP-Dow Jones shows... Economists concurred that the expected increase in the price index is largely due to an increase in energy prices...</td>
</tr>
<tr>
<td>Europe</td>
<td></td>
</tr>
<tr>
<td>Market perfor.</td>
<td></td>
</tr>
<tr>
<td>East Asia</td>
<td></td>
</tr>
<tr>
<td>Petroleum</td>
<td></td>
</tr>
<tr>
<td><strong>1999 - 2002</strong></td>
<td></td>
</tr>
<tr>
<td>Labor market</td>
<td>(2000-05-22) So you’ve started a successful company before your 30th birthday. Big deal. Navin Chaddha has co-founded five. What’s more, the 29-year-old electrical engineer has assisted and even invested hundreds of thousands of his own dollars in at least eight other start-ups...</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td></td>
</tr>
<tr>
<td>Internet</td>
<td></td>
</tr>
<tr>
<td>Persuasion</td>
<td></td>
</tr>
<tr>
<td><strong>2002 - 2006</strong></td>
<td></td>
</tr>
<tr>
<td>Stocks</td>
<td>(2003-09-05) Look past the ongoing sabotage and strife in Iraq and you will see that the Bush administration is eager to pull off the most ambitious economic reform in a Middle Eastern country since the dissolution of the Ottoman Empire... The administration wants to promote free trade for the entire gamut of Arab countries,...</td>
</tr>
<tr>
<td>Labor market</td>
<td></td>
</tr>
<tr>
<td>Events</td>
<td></td>
</tr>
<tr>
<td>Terrorism</td>
<td></td>
</tr>
<tr>
<td>Strategy</td>
<td></td>
</tr>
<tr>
<td><strong>2006 - 2009</strong></td>
<td></td>
</tr>
<tr>
<td>Labor market</td>
<td>(2008-09-17) Congressional auditors are questioning whether the Interior Department is collecting all the royalties energy companies owe for petroleum developed on federal property... Last year, the MMS collected more than $11.4 billion in oil, natural-gas and other mineral royalties... Congress this week is debating proposals to allow more offshore oil drilling...</td>
</tr>
<tr>
<td>Regulations</td>
<td></td>
</tr>
<tr>
<td>Congress</td>
<td></td>
</tr>
<tr>
<td>Natural gas</td>
<td></td>
</tr>
<tr>
<td>Strategy</td>
<td></td>
</tr>
<tr>
<td><strong>2009 - 2013</strong></td>
<td></td>
</tr>
<tr>
<td>Labor market</td>
<td>(2010-03-03) When it comes to talking about what is holding back the economy, politicians in Washington should look in the mirror. Inaction and infighting on the government level have resulted in a loss in confidence among consumers and business owners that their elected officials are doing the right thing when it comes to healing the economy or bringing down unemployment...</td>
</tr>
<tr>
<td>Clients</td>
<td></td>
</tr>
<tr>
<td>Elections</td>
<td></td>
</tr>
<tr>
<td>Sports</td>
<td></td>
</tr>
<tr>
<td>Congress</td>
<td></td>
</tr>
<tr>
<td><strong>2013 - 2016</strong></td>
<td></td>
</tr>
<tr>
<td>Labor market</td>
<td>(2013-05-16) Even though inflation measures have fallen sharply in recent months, Federal Reserve officials aren’t ringing alarm bells about it as they have done in the past. Fed officials have said they take comfort that the public’s expectation of future inflation, as registered in surveys of households and bond markets, has remained stable...</td>
</tr>
<tr>
<td>Monetary policy</td>
<td></td>
</tr>
<tr>
<td>Documentation</td>
<td></td>
</tr>
<tr>
<td>Clients</td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td></td>
</tr>
</tbody>
</table>

whereas the topic *Terrorism* enters the list during the 2002-2006 period. Likewise, the *Terrorism* narrative enters the top five list during the 2013-2016 period in the euro area together with the *Monetary policy* topic. Interestingly, and something we will come back
to, the narrative focus on monetary policy is also shared by the US and Japan during this time period. The news article excerpts reported in the tables illustrate how the discovered topic structure in the corpus, together with the DFM decomposition, provides meaningful mappings. It is, for example, easy to argue that the excerpts for the US are about at least *Europe*, *Petroleum*, and *Market performance* (1995-1999), *Regulations*, *Congress*, and *Natural gas* (2006-2009), and *Labor market*, and *Congress* (2009-2013).

Figure 4 provides an illustration of how news surprises in the US affect the NCI-US estimates over time, at a daily frequency. Two distinct results stand out. First, the timing of when specific topics become important, either positively or negatively, resonates well with the conventional narrative held about economic developments the last two decades. At the risk of cherry picking, we give some examples: Prior to, and going into the 2001 recession, surprising news related to the *Internet*, *Design*, *Education*, *M&A*, and *Volatility* topics pulled the coincident index upwards, while narratives related to *Labor markets*, *Bankruptcies*, and *Automobiles* pulled the coincident index downwards. Thus, interpreted through the lenses of the model proposed here, the burst of the dot-com bubble is well identified, but the news topic developments directly related to the grand dot-com narrative was not as bad as the model expected. Conversely, news topic developments related more towards the general economic conditions came in worse than predicted. The story related to the financial crisis in 2007/2008 is of a somewhat different type. Now surprising negative movements in topics as *Strategy*, *Bonds*, and *Regulations*, stand out. Lastly, turning to the slow recovery period following the financial crisis, we observe that unexpected news about *Congress*, *Economic crisis*, *Funding*, *Environment*, and *Commodities* contributed negatively to growth, while topics related to *Labor market*, *Sports*, *Commentary*, *Natural gas*, and *Elections* helped pull the index upwards.

Second, the degree of sparsity enforced on the factor loading space changes considerably across time. For example, during the 1990s few factor loadings have a high probability of being zero. In the period following the financial crisis, however, the degree of sparsity is much larger, with only a few time-varying factor loadings being larger (in absolute value) than their respective threshold. It is also interesting to see how the degree of sparsity seems to increase around recession periods. That is, when times are bad, our results indicate that the set of narratives relevant for economic fluctuations is smaller. Interestingly, this finding is very much in line with theory models explaining how news coverage becomes more homogeneous around major events, and thereby increasing the correlation among economic agents’ actions (*Nimark and Pitschner* (2016)). Thus, in relation to narratives, booms are broad-based while busts are not.
Figure 4. US news topics and their contribution to coincident index estimates across time. The reported decompositions are based on running the Kalman Filter using the posterior median estimates of the hyper-parameters and the time-varying factor loadings (at each time $t$). In the interest of readability, the topic names are reported on two y-axes with two-step increments. For example, the Labor market topic is associated with the first row (from above) in the figure, while the Stocks topic is associated with the second row (from above). White areas illustrate the time-varying sparsity patterns. Recession periods, defined by NBER, are illustrated using gray shading.
Figures similar to 4 are reported for Japan and the euro area in Figures 11 and 12 in Appendix A. Instead of going into the details, we highlight that we see clear sparsity patterns around recession periods, like in the US. In Europe, for example, Credit rating, Bonds, Investing, Outlook, and Funding are almost the only news topics contributing to explaining the negative developments in the euro-area business cycle index during, and following, the financial crisis. Similarly, in Japan narratives related to Electronics, Retail, Income, and Growth contributed especially negatively during 2009, while the period between 2010 and 2011 is partly dominated by negative news topic surprises attributed to Politics and US politics.

Finally, although most topics are easily interpretable and provide information about what is important for the current state of the economy, some topics either have labels that are less informative, or reflect surprising categories. From the US-based decompositions, in Figure 4, examples are the Sports, Entertainment, and Food topics. That said, such exotic or less informative named topics, are the exception rather than the rule. It is also the case that a news article is a mixture of topics. To the extent that different topics, meaningful or not from an economic point of view, stand close to each other in the decomposition of the corpus they might covary and therefore both add value in terms of reflecting the current state of the economy.

We conclude that the decompositions of the business cycles into narrative contributions tell a story about economic fluctuations reasonably in line with historical experience. This should not be too surprising, given that the narratives we know are the ones we have been served, partly through the media. What is perhaps more surprising is that it is quantified so well. The finding about narrative sparsity around recessions is novel, and some of the influential news topics clearly represent (economic) concepts or events that would have been very difficult, if not impossible, to capture using conventional economic data.

4.3 Going viral?

Shiller (2017) argues, but does not quantify, that “narratives “go viral” and spread far, even worldwide, with economic impact”. Accordingly, a reasonable testable hypothesis is that there exists a significant relationship between how important similar news topics are in explaining business cycle developments across countries and economic fluctuations, at least periodically. We investigate this hypothesis by first constructing statistics measuring how similar the news topics are across countries. Then, we weight these similarity measures with how important the news topics are in explaining business cycle developments and derive what we label virality indexes. These indexes give a quantitative measure of the degree to which (similar) narratives relevant for growth go viral. Finally, we exploit the high frequency nature of our data, and investigate if there is any significant relationship
between the virality indexes and economic fluctuations across countries.

To measure topic similarity across countries, we use the Jensen-Shannon divergence (JSD). This is a method for measuring the similarity between two probability distributions. The JSD is based on the Kullback-Leibler divergence, but it is symmetric, always a finite value, and bounded between 0 and 1. Formally, for two discrete probability distributions $P$ and $Q$:

$$JSD(P||Q) = \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M)$$ (8)

where $M = \frac{1}{2}(P + Q)$, and $D(P||M)$ is the Kullback-Leibler divergence:

$$D(P||M) = \sum_i P_i \log_2 \frac{P_i}{M_i}$$ (9)

Here, with reference to Section 3.2, and Table 1, $P$ and $Q$ are two word distributions ($\Phi_k$) associated with two different topics. Treating the US economy as the common “numeraire”, we compute the $JSD(P||Q)$ for all combinations of topics in the US and either Japan or Europe. This results in two $K \times K$ matrices, one for each country pair, with JSD scores. Table 12, in Appendix A, reports the topic combinations with the lowest JSD score (most similar), and shows that the mappings make sense intuitively. For example, the US topics we have labeled Fiscal policy, Funding, and, Telecommunication, have gotten the same labels in both Japan and Europe, while the US topic Monetary policy has gotten the label Fed/BoJ and Fed in the Europe and Japan, respectively. In some cases, however, there are larger, less intuitive, discrepancies. An example is the US-based topic labeled Canada by us, which according to the JSD score is most similar to the European and Japanese topics Outlook and Fiscal policy.

The virality index $VIR_{s,US}^t$ between country $s$ and the US is constructed as follows:

$$VIR_{s,US}^t = \sum_{i=1}^{80} \sum_{j=1}^{80} \left[ \frac{\tilde{w}_{i,j}^s w_{i,t}^S}{(c + JSD_{i,j}^s)} \right]$$ (10)

Here, $\tilde{w}_{i,j}^s = w_{i,t}/\sum_i^K w_{i,t}$, with $w_{i,t}$ defined in Section 4.2, i.e., the normalized weight given to topic $i$ in explaining the movements in the business cycle index in country $s$ at time $t$, while the $JSD_{i,j}^s$ term defines how similar topic $i$ in country $s$ is to topic $j$ in the US. $c$ is a small constant ensuring that we do not divide the expression by 0, which is the lower limit of the VIR indexes.

Figure 5 reports the two virality indexes. On average, the indexes fluctuate mildly. However, at times the indexes spike, and some narratives go viral and become an epidemic. This pattern is especially pronounced following the financial crisis in 2008, when the frequency, duration, and magnitude of the spikes all increase significantly relative to the periods before. More formally, using a peak-finding algorithm to compute the number
of peaks, and their duration, we identify only two peaks prior to 2008, see Figure 13 in Appendix A. This is in the late 1997 for the \( \text{VIR}^{\text{Japan,US}} \) index, and in early 2000 for the \( \text{VIR}^{\text{Euro,US}} \) index. The length of these episodes are roughly 3 and 6 months. In contrast, in the periods following 2008, we identify in total 11 epidemics with durations up to 8 months.\(^{12}\) The average duration of the epidemics are estimated to be around 5 and 4.5 months for the \( \text{VIR}^{\text{Japan,US}} \) and \( \text{VIR}^{\text{Euro,US}} \) indexes, respectively, where events happening late in the sample tend to pull these averages up.

Borrowing from Shiller (2017) and the spread of disease literature and the benchmark SIR model of Kermack and McKendrick (1927), our results indicate that the contagion rate \( (co) \) to recovery rate \( (re) \) ratio has increased over time. That is, (narrative) epidemics in the post 2008 period are more severe than in previous periods. Many different explanations can rationalize this finding. It is for example easy to argue that the introduction of internet and social media likely have increased both \( co \) and \( re \) (Zhao et al. (2013)). However, in terms of Figure 5, it seems strange that this should have happened exactly in the mids of the financial crisis in 2008, suggesting instead that the epidemics observed during and after 2008 might be of a very different type than those encountered during the 1990s and early 2000s.

In Figure 13, in Appendix A, we also report the topic mappings contributing the most to the VIR estimates during the epidemic periods discussed above. Three broad findings stand out. First, epidemics are mostly associated with the US Labor market topic. In

\(^{12}\)We have also tried defining periods of virality using a generalized version of the sup augmented Dickey-Fuller test (Phillips et al. (2015)). However, this test has low power in terms of correctly classifying spikes/bubbles when the duration of each is small relative to the total sample size. As this is the case here, the number of periods defined as explosive are far fewer than suggested by Figure 5.
Table 4. Epidemics and economic fluctuations. For each month in the sample we compute the mean and standard deviation of the three news-based coincident indexes, as well as their correlation with the NCI-US index, using the daily observations. Contagion periods (Cont.) are defined using the timing and durations implied by the results in Figure 13, in Appendix A. Periods of no contagion are defined as normal times (Norm). Significant differences in the moments (Diff) are tested using the Welch’s t-test. The superscripts ***, **, and * denote the 1%, 5%, and 10% significance level, respectively.

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>Japan</th>
<th>Europe</th>
</tr>
</thead>
<tbody>
<tr>
<td>E(X)</td>
<td>-0.30</td>
<td>0.06</td>
<td>-0.36**</td>
</tr>
<tr>
<td>STD(X)</td>
<td>0.07</td>
<td>0.07</td>
<td>-0.00</td>
</tr>
<tr>
<td>COV(X, US)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

almost all episodes this topic features as a central component in the explaining the spikes in the VIR indexes. Second, there are three exceptions to this first point, namely the spike in the $VIR_{Euro,US}^E$ index in 2000, and the spikes in the $VIR_{Japan,US}^E$ index in 2014 and 2015. The former is undoubtedly related to the burst of the dot-com bubble, while the two latter are associated with the US Monetary policy topic. Third, the diversity of topics needed to explain a sizable share of the epidemic episodes varies considerable across time. During the spike in the $VIR_{Euro,US}^E$ index in September 2009, only one topic mapping is needed to explain up to 40 percent of the index. In contrast, during the September 2013 epidemic in the same index, 13 topic mappings are needed. Thus, some epidemic episodes have a “sharp” narrative interpretation, while others are more complex. Based on the topic contributions, and the timing, we can for example conjecture that the 2009 episodes are related to the Great Recession, while the 2011 episodes are related to the massive earthquake that hit Japan this year, sparking well known global concerns about both finance, trade, and energy related topics. We do not find, however, any relationship between the estimated duration of the epidemics, and the number of topic mappings needed to explain a sizable share of the VIR indexes during such episodes.

The estimated timing of the VIR epidemics suggest that they are associate with bad events, and thus potential negative economic developments. The results reported in Table 4 confirms this impression. Higher values of the VIR indexes are associated with lower growth rates than in “normal” times in all three countries, and significantly so in the US and Japan. On the other hand, we do not find any significant differences in the covariances between the country pairs during periods of epidemics relative to normal times. If anything, it becomes lower between the US and Europe. To the extent that increases in the $VIR_{Euro,US}^E$ index are considered as some type of common shock(s) to the international business cycle (Kose et al. (2003), Stock and Watson (2005)), this means that
their (short-term) propagation differ across countries, potentially leading to divergence, as opposed to convergence, of international business cycles (Mumtaz et al. (2011) Kose et al. (2012)).

To summarize, the preceding analysis has shown that narratives do “go viral” and spread worldwide, as argued by Shiller (2017), but mostly so in times of trouble. The narratives contributing the most to the epidemic episodes tend to be associated with US-based macroeconomic developments and (partly) monetary policy.

4.4 Behind the news

No causal inference is sought, or can be inferred, from the preceding analysis. Here, we take one step towards a more structural understanding of information diffusion. In particular, we ask how narratives independently spread between economic regions, and whether news topics that are important for describing, e.g., the US business cycle, have predictive power for narratives in Japan and Europe, or vice versa.

We answer these questions by building on the well known Granger causality concept (Granger (1969)). A variable is said to Granger cause another, if the first series contains additional information for predicting the future values of the second series, beyond the information in the past values of this second series. While originally formulated in a low dimensional setting, recent work trying to infer causal relationships among components of biological systems has extended this reasoning to high dimensional problems through the usage of “Graphical Granger causality” modeling (Lozano et al. (2009), Shojaie and Michailidis (2010)). These methods offer efficiency gains over more standard (pairwise) Granger causality tests because of the usage of regression methods with variable selection and regularization (Arnold et al. (2007)), i.e., Lasso and its variants, and are tailored for high dimensional problems, as here.

Let $y_j^T = (y_{j,1}^T, \ldots, y_{j,T}^T)'$ be a $T \times 1$ response variable $j$, and $X = [X_1^T, \ldots, X_J^T]$ be the predictor matrix for $j = 1, \ldots, J$ groups of covariates (including $y$). Each matrix $X_j^T = [L^1 x_j^T, \ldots, L^h x_j^T]$, where $x_j^T = (x_{j,1}^T, \ldots, x_{j,T}^T)'$, $L$ is the lag operator and $h$ is the maximum number of lags. Then, we answer the question posted above, i.e., how narratives independently spread between economic regions, by estimating the group Lasso of Yuan and Lin (2006):

$$
\hat{\beta}(\lambda) = \arg\min_{\beta} \|y_j^T - X \beta\|_2^2 + \lambda \sum_{j=1}^J \|\beta_{G_j}\|_2
$$

(11)

for each $j = 1, \ldots, J$. $\beta_{G_j} = \{\beta_k; k \in G_j\}$ and $G_j$ denotes the set of group indexes. In our case each group is of equal length, and correspond to all the lagged variables belonging to one group. $\lambda$ is the Lasso regularization parameter that shrinks or sets some of the groups (coefficients) to 0. Thus, the group Lasso is faithful to the original (pairwise) Granger
Figure 6. A network graph of the graphical Granger causality results. Each node is a narrative (news topic time series). In the interest of visual clarity, their name is not reported. The (gray) edges connecting the nodes are directed, and illustrate the direction of predictability across narratives. The highlighted nodes are those that are estimated to be the most (least) central narratives in the graph, see Table 6.

causality concept, where \( x^{i \neq j} \) is said to Granger cause \( y^j \) only if the entire lagged series \( X^{i \neq j} \) provides additional information for the prediction of \( y^j \).

We have \( J = K \times 3 = 80 \times 3 = 240 \) individual news topic time series, or groups. Before estimation, to reduce noise, the individual news topic time series are aggregated to monthly values, and the predictor matrix is standardized to make estimation scale invariant. We consider up to a half-year of lags, with \( h = 3 \). As \( T \ll (J \times h) \), a standard regression framework is infeasible, while the Lasso applies because of the regularization term. For each \( j \), we set \( \lambda^i_{\text{max}} \) such that it gives the largest non-null model. The group Lasso solution path is then computed by evaluating on 100 equally spaced \( \lambda \)’s between 0 and \( \lambda^i_{\text{max}} \). The optimal \( \lambda^i_{\text{opt}} \) is chosen based on the BIC, as in Lozano et al. (2009).

We focus on cross-country spillovers, and say that a topic \( i \) in country \( s_1 \) Granger causes topic \( j \) in country \( s_2 \) when \( \hat{\beta}^{i}_{G_{i \neq j}}(\lambda^j_{\text{opt}}) \neq 0 \). More generally, Shojaie and Michailidis (2010) show how the output from procedure described above admits a graphical interpretation. In particular, we can construct the adjacency matrix of a directed acyclic graph (DAG) by stacking the estimated \( J \times 1 \) coefficient vectors \( \hat{\beta}^j(\lambda^j_{\text{opt}}) \), for \( j = 1, \ldots, J \), into a \( J \times J \) matrix \( A \), whose \((i, j)\)-th entry indicates whether there is an edge (and its weight) between nodes \( i \) and \( j \). Below, to simplify the interpretation, we do not count relationships where there is a two-way predictive relationship, and set elements in \( A \) where both the \((i, j)\)-th and \((j, i)\)-th are non-zero to 0.
Table 5. Graphical Granger causality. For each country, the Tot. columns report the number of outgoing edges in percent of total potential connections. The remaining columns decomposes this fraction into cross-country contributions.

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th>Japan</th>
<th>Europe</th>
<th>Tot.</th>
<th>US</th>
<th>Japan</th>
<th>Europe</th>
<th>Tot.</th>
<th>US</th>
<th>Japan</th>
<th>Europe</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>9.92</td>
<td>49.06</td>
<td>50.94</td>
<td>4.91</td>
<td>48.73</td>
<td>51.27</td>
<td>6.66</td>
<td>36.38</td>
<td>63.62</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The network graph in Figure 6 illustrates the complexity of the problem, and shows that the interconnectedness of narratives across the US, Japan, and Europe is large. Still, given the large number of potential connections, the density of the graph is rather small, and estimated to be approximately 5 percent.\(^\text{13}\) The statistics reported in Table 5 break the graph density into country specific contributions. Out of 12800 potential connections, US-specific narratives dominate, and Granger cause roughly 10 percent of the foreign news topics. The direction of predictability is divided equally towards topics in Japan and Europe. The importance of Japan is only half that of the US, while narratives classified as being euro area-specific Granger cause roughly 7 percent of the foreign topics. However, in contrast to the results for the US- and Japan-specific news topics, the direction of the European-specific predictability is clearly tilted towards Japan.

As these results are new, they are hard to compare to existing knowledge. Still, the US-based dominance is well in line with common perception, and adds to the evidence about the US’s role in the global economy more broadly (Kose et al. (2017)).\(^\text{14}\)

To gain knowledge of the narratives’ importance in the graphical Granger causality network, and relate this importance to the narratives’ importance for economic fluctuations, we compute a measure of the graph node’s centrality using the much applied “betweenness” measure (Freeman (1977)). This centrality metric measures how often each graph node appears on a shortest path between two nodes in the graph, and is computed as:

\[
c(u) = \sum_{i,j \neq u} \frac{n_{ij}(u)}{N_{ij}}
\]

where \(n_{ij}(u)\) is the number of shortest paths from \(i\) to \(j\) that pass through node \(u\), and \(N_{ij}\) is the total number of shortest paths from \(i\) to \(j\). In addition, a cost, equaling \(1/\tilde{w}_s\),

---

\(^{13}\)The density of the graph is computed as the number of non-zero elements in the adjacency matrix \(A\) relative to the number of total elements.

\(^{14}\)In unreported results we confirm that these findings hold when partitioning the sample into three equally sizes sub-samples, and re-estimating the graphical Granger causality graph for each. Relatedly, Table 8, in Appendix A, shows that among the daily business cycle indexes themselves, neither the NCI-Japan nor the NCI-Japan index Granger cause the NCI-US index, while the NCI-US index Granger causes at least the NCI-Euro index.
Table 6. Topic centrality. The centrality ranking is computed using the weighted “betweenness” measure of Freeman (1977). The In degree and Out degree counts reflect how many series that predict the listed topics, and how many topics the listed topics themselves predict, respectively.

<table>
<thead>
<tr>
<th>Name</th>
<th>Most central</th>
<th></th>
<th>Least central</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In degree</td>
<td>Out degree</td>
<td>In degree</td>
<td>Out degree</td>
</tr>
<tr>
<td>Euro T5-Macroeconomics</td>
<td>8</td>
<td>20</td>
<td>Japan T55-Intervention</td>
<td>19</td>
</tr>
<tr>
<td>Euro T48-Middle East</td>
<td>8</td>
<td>11</td>
<td>Euro T66-Justice</td>
<td>9</td>
</tr>
<tr>
<td>US T30-Regulations</td>
<td>9</td>
<td>12</td>
<td>Japan T19-Months</td>
<td>2</td>
</tr>
<tr>
<td>Japan T6-Fed</td>
<td>13</td>
<td>11</td>
<td>US T25-Clients</td>
<td>0</td>
</tr>
<tr>
<td>US T55-Labor market</td>
<td>10</td>
<td>14</td>
<td>US T28-Software</td>
<td>0</td>
</tr>
<tr>
<td>Euro T14-Fiscal policy</td>
<td>17</td>
<td>10</td>
<td>US T38-Stocks</td>
<td>0</td>
</tr>
<tr>
<td>US T16-Market performance</td>
<td>6</td>
<td>18</td>
<td>US T57-Australia</td>
<td>2</td>
</tr>
<tr>
<td>Japan T62-Car technology</td>
<td>14</td>
<td>10</td>
<td>Euro T52-Credit rating</td>
<td>0</td>
</tr>
<tr>
<td>Japan T3-Aviation</td>
<td>21</td>
<td>11</td>
<td>US T65-Bankruptcies</td>
<td>0</td>
</tr>
<tr>
<td>Japan T58-Natural disasters</td>
<td>28</td>
<td>3</td>
<td>US T74-Commodities</td>
<td>0</td>
</tr>
</tbody>
</table>

is assigned to each edge in the graph, where $\tilde{w}^*_i$ was defined in Section 4.3 as the average normalized weight given to topic $i$ in explaining the movements in the business cycle index in country $s$. Thus, when computing the shortest path between two nodes in the graph, we rather traverse across edges which are important for explaining aggregate economic fluctuations.

Table 6 reports the 10 most and least important narratives according to (12). The news topics that are found to be important for explaining the economic fluctuations in the US, Japan, and Europe (confer Table 2), are also among the most important narratives in the graphical granger causality graph. The Macroeconomics topic in Europe, for example, is at the top of the list, and has an in and out degree in the network of 8 and 20, respectively. Conversely, at the bottom of the list we find the US Commodities topic. This news topic times series is not predicted by any of the other narratives, and therefore has a very low $c(u)$ ranking. Still, even though many of the least central news topics have low in degree, many of them have a relatively high out degree. The colored nodes in Figure 6 illustrate this, where the narratives with a low $c(u)$ score tend to be found far out in the network graph, while more connected news topics tend to be found closer to the center of the graph.

Figures 14 and 15, in Appendix A, provide examples of how two of the most and least central narratives in the network graph are connected to other topics. The figure is constructed as a subgraph of Figure 6. As the figures illustrate, the narratives Macroeconomics and Commodities tend to predict narratives of a similar type in other countries. For the Middle East and Bankruptcies topics, however, the predictive relationships are
more diverse.

Lastly, it is worth noticing before turning to the next section that among the 10 least central narratives in Table 6, we find 6 US news topics. Of these, both the Stocks and Clients are also among the 10 most important in terms of describing the US business cycle, confer Table 2 in Section 4.2.

4.5 News or noise?

The literature we speak to is divided in its view on whether narratives contain fundamental economic information, or just noise and sentiment. One branch of the literature can be associated with the news-driven business cycle view. Here, changes in expectations, due to news (new information), is put forward as the primary driver of economic fluctuations, and linked to economic fundamentals, i.e., total factor productivity (Barsky and Sims (2012), and Blanchard et al. (2013)). An alternative view of narratives and their role in explaining economic fluctuations builds on the classical work of Pigou (1927) and Keynes (1936) on capturing the market’s animal spirits where changes in agents’ expectation can be totally self-fulfilling or not rooted in economic fundamentals at all. Such mechanisms have for example been highlighted by Shiller (2000), and recent work by Angeletos and La’O (2013).

Since changes in expectations are not directly observable, and since economic feedback loops easily can confound the cause and effect relationship, it is intrinsically difficult to discriminate between these two opposing views. Empirical investigations have therefore resorted to using various high frequency and hard to predict economic variables, e.g., asset prices or consumer sentiment (Beaudry and Portier (2006), Barsky and Sims (2012)), to approximate news and changes in expectations. In contrast, our approach permits the usage of a primary source of (potential) new information directly, namely the news narratives.

To this end, we build on the results presented in the previous section and partition the high dimensional news topic dataset into what we loosely call “propagators” and “initiators”. The “propagators” are news topics with a high centrality score in the graphical Granger causality network. Such narratives predict many of the other series, but are also themselves predicted by a large share of other news topics. In contrast, the “initiators” are more exogenous. At the extreme they are not predicted by any of the other series, but they do still themselves have predictive power for other narratives (confer Table 6). Thus, any unexpected changes in these less central parts of the network should be less likely to be due to potential feedback loops, and more likely to represent new information.

Building on this simple logic, and focusing on the US, Figure 7a plots the first principal component estimate of the five most “exogenous” US-based news topic time series, i.e.,
those with in degree equaling 0 from Table 6, together with total factor productivity (TFP). The factor estimate explains 55 percent of the total variation across the five variables, and is reported on a quarterly frequency. The TFP measure is adjusted for capacity utilization using the methodology suggested by Basu et al. (2006), and obtained from the Federal Reserve Bank of San Francisco web pages (Gerstein (2018)). As seen from the figure, the TFP estimate shows much more high frequency variation than the news factor. Still, there is a clear tendency for the two series to move together. Their contemporaneous correlation is 0.2.

To investigate the dynamic relationship between the news factor and TFP, and show how unexpected fluctuations in the news factor affects TFP, we formulate a simple bivariate Structural Vector Autoregression (SVAR) with these two variables. In the tradition of Beaudry and Portier (2006) and Barsky and Sims (2012), shocks to the news factor are identified using a recursive ordering where TFP is ordered first in the system and the news factor last. Thus, unexpected innovations in the news factor are orthogonal to contemporaneous TFP disturbances, and can only affect TFP with a lag. According to the new-driven business cycle view, and to the extent that shocks to the news factor contain new information, we expect a delayed but persistent increase in TFP. On the other hand, if the narratives just contain sentiment and noise, TFP should not respond at all to unexpected shocks in the news factor.

Figure 7b reports the cumulative response, i.e., the level, of TFP following a shock to the news factor. During the first year following the initial impulse, TFP is more or less unaffected. Then it increase significantly, and remains at a higher level than prior to the shock. This response pattern is as predicted by the news-driven business cycle view, and
suggests that the news factor carries fundamental information, and not only noise and sentiment. The news shock also explains a large fraction of the variation in TFP. At the 10- and 40-quarter horizons, for example, as much as 22 and 52 percent of the variation in TFP can be attributed to the news shock.

Words clouds for the five narratives used to construct the news factor, the US-based topics Clients, Software, Stocks, Bankruptcies, and Commodities, are illustrated in Figure 16 in Appendix A. Examples of stories representative for these topics are reported in Table 7. As before, the narrative realism of the news topic-based approach stand out. The stories are clearly about technological changes, but also partly associated with developments in financial markets. However, as seen from Figure 17a, in Appendix A, the news factor does not work as a stand-in for surprising movements in asset markets. In particular, when we augment the SVAR model with quarterly returns from the Dow Jones Industrial Average, and order this variable above the news factor (but below TFP) in the recursively identified SVAR, our results remain basically unchanged from the benchmark case in Figure 7b.

The flip side of the argument used above is that unexpected innovations to the narratives with a high centrality score, i.e., the “propagators”, should be less likely to lead to a significant TFP response. Figure 17b confirms this hypothesis. When computing the first principal component of the two US-based news topic variables with the highest
centrality score, confer Table 6, and re-estimating the bivariate SVAR described above with this factor instead of the earlier news factor, we obtain insignificant results.

To the best of our knowledge, quarterly TFP statistics do not exist for Japan and the euro area (and, due to data availability they are hard to construct). Still, using interpolated quarterly TFP estimates based on the yearly statistics provided by the European Commission, we can get an impression of whether or not shocks to the US-based news factor tend to affect productivity levels globally as well. The results from this experiment are reported in Figure 18 in Appendix A. Following a news shock, the level of TFP in the euro area increases significantly, in line with the results for the US, although with a substantial lag of up to two years. For Japan, however, we get insignificant results. In that respect, it is interesting to note that among the 88 outgoing edges from the five US-based initiators used to construct the news factor (confer Table 6), 60 percent go directly to European news topics. Thus, in line with earlier results, there seems to be a stronger relationship between the US and Europe, than with the US and Japan, also when it comes to narratives associated with economic fundamentals.\textsuperscript{15}

While our results clearly suggest that narratives, or at least some of them, carry fundamental information, we can not rightfully argue that these narratives cause TFP. There are well known potential problems with using SVAR models to try to uncover the structural effects of anticipated shocks (news shocks) (Sims and Zha (2006), Forni et al. (2017), Blanchard et al. (2013)). More broadly, establishing a causal relationship between narratives and economic developments, in terms of potential outcomes (Rubin (2005)), is difficult because of the obvious simultaneity between economic events and media coverage of the same events. Without some truly exogenous information, decoupling the effect of the new information component (the economic event) from the effect of the ether (the media generating the narrative or reporting on the event) is challenging.

Still, our results are very much in line with other newer studies trying to understand the underlying relationship between news and economic fluctuations using exogenous events and high-frequency data. For example, Larsen and Thorsrud (2017) use an exogenous strike in the newspaper market to show that up to 40 percent of the predictive effect from news topics to daily asset returns can be attributed to the causal effect of the media itself. Similarly, it is interesting to note that the narratives defined as “initiators” here overlap well in theme and meaning with the news topics associated with productivity

\textsuperscript{15}At the 40-quarter horizons, 47 and 7 percent of the variation in the euro area and Japanese TFP measures, respectively, can be attributed to the news shock. The close to idiosyncratic behavior of Japanese productivity growth is also found in Crucini et al. (2011). They compute a common (yearly) component of productivity growth across G7 countries, and document that as little as 16 percent of the variation in TFP in Japan can be attributed to a common global component. In contrast, for the US this number is 43 percent.
for the Norwegian economy in Larsen and Thorsrud (2018). In that study, using a very different approach, news topics labeled Funding, Stock market, and IT/startup are among the most influential. These narratives share many important words with in particular the Bankruptcies, Stocks, and Software topics found to be important here.

5 Conclusion

To what extent are narratives informative for describing business cycle variation, do they go viral, how do they interact with each other, and are they associated with economic fundamentals or better understood as capturing the market’s animal spirits?

In this article we focus on the three major economies the US, Japan, and the euro area, and show how unstructured textual news data can be used to provide quantitative answers to these questions. We do so by first constructing daily business cycle indexes computed on the basis of the news topics the media writes about. We then derive virality indexes capturing the extent to which narratives relevant for growth go viral and affect economic fluctuations across borders, and finally use so called “Graphical Granger causality” modeling to cast light on cross-country narrative spillovers and whether or not narratives carry news or noise.

The resulting coincident indexes are shown to classify the phases of the cycle with high precision. At a broad level, the most important news narratives are shown to be associated with general macroeconomic developments, finance, and (geo-)politics. However, a vast set of narratives contribute to our index estimates across time, especially in times of expansion. In times of trouble, narratives associated with economic fluctuations become more sparse. Likewise, we show that narratives do go viral, with an average epidemic duration of 4-5 months, but mostly so in times of trouble. Finally, while narratives interact in complicated ways, we document that some news topics are clearly associated with economic fundamentals, as predicted by the news-driven business cycle view. Other narratives, on the other hand, show no such relationship, and are likely better explained by classical work capturing the market’s animal spirits.

More than providing definite answers, we offer a number of new results about the relationship between business cycles and narratives, and an analytical framework for quantifying such interactions. Natural extensions to the approach taken here include: comparing the topic model approach to other automated Natural Language Processing techniques and further investigate how textual data can be translated into useful time series; exploiting the high frequency nature of the news data and natural frictions in information flow (e.g., time zones), to design experiments better suited for understanding the underlying structural relationship between narratives and economic fluctuations.
References


