

A Thousand Words Tell More Than Just Numbers: Financial Crises and Historical Headlines *

Kim Ristolainen[†], Tomi Roukka,[‡] and Henri Nyberg[§]

February 11, 2022

Abstract

We show that financial crises are preceded by changes in specific types of narrative information contained in newspaper article titles. Our novel international dataset and the resulting empirical evidence are gathered by integrating information from a large panel of economic news articles in global newspapers between the years 1870 and 2016 with conventional macroeconomic and financial indicators. We find that the predictive information of newspaper article titles that signals coming crisis episodes is substantial over and above the macroeconomic and financial indicators. The new indicators capture common features that have often been discussed as potential causes of specific crises but not incorporated yet into empirical models.

JEL codes: G00, G01, N01, C25, C82.

Keywords: financial crisis, text data, leading indicators, topic model .

* We thank Alp Simsek, Janne Tukiainen, Paavo Okko, Matti Virèn, Karlo Kauko and seminar participants at the University of Turku and the Bank of Finland for their useful comments and discussions. We gratefully acknowledge the research assistance provided by Elsi Lindell and Miko Pasanen, as well as the financial support provided by the Emil Aaltonen Foundation, (Ristolainen) OP Group Research Foundation, (Ristolainen) Yrjö Jahansson Foundation, and (Nyberg) the Academy of Finland (grant 321968).

[†] Department of Economics, Turku School of Economics, University of Turku, Finland. Address: Turku School of Economics, Department of Economics, FI-20014 University of Turku, Finland. E-mail address: kim.ristolainen@utu.fi. Corresponding author.

[‡] Department of Economics, Turku School of Economics, University of Turku, Finland

[§] Department of Mathematics and Statistics, University of Turku, Finland

1 Introduction

Financial crises have been of central interest since the global financial crisis of 2008 and the subsequent European sovereign debt crisis. In the recent economic literature, large historical panel datasets such as the one assembled by Jordà et al. (2017) covering more than 150 years of macroeconomic and financial data for multiple countries have generated newfound interest in examining leading indicators. In a series of papers utilizing this new rich data, it has been shown that credit expansions (Schularick and Taylor, 2012) and private sector debt (Jordà et al., 2016b) have predictive power for banking crises, but a crisis preceded by lower capital ratios (Jordà et al., 2020) and large public debt is more costly in general.¹ Although many leading indicators have been found, the prediction performance of these models is still modest: A randomly selected year can be classified as a crisis (or “normal”) year correctly with around 70% accuracy. Despite this (better than a random coin toss) predictive power, the models can still be seen as not accurate enough to persuade policy makers to rely solely on the signals of these models to take systematic preventive action in a timely manner.

Recent research has uncovered many new insights into multiple economic applications and issues by incorporating newspaper data in measuring relevant economic variables. Specifically, text data has been used to measure stock market sentiment (Tetlock, 2007), economic policy uncertainty (Baker et al., 2016), firm-level policy uncertainty (Hassan et al., 2019), and the relationship between news article topics and U.S. economy (Bybee et al., 2020). Text data from economic newspaper articles may contain much broader information about economic issues than just the plain numerical values of GDP growth, interest rates, unemployment, inflation, or even asset prices. Importantly, macroeconomic variables are also usually reported with a lag and typically revised multiple times in the future. In addition, asset prices may be influenced by financial market anomalies, belief formation biases, and heuristics that violate the rational expectations assumption and complicate their use as leading indicators.

In this paper, we use Blei et al. (2003)’s Latent Dirichlet Allocation topic model to assemble a large annual panel dataset containing specific time- and country-invariant economy-related topics in news article titles published in a collection of global newspapers during the period 1870–2016. The topic model is formed without any specific purpose or target, i.e. including only the topics that we might think to be related to financial crises. That is, we aim to include and distinguish all important economy related themes from this large collection of news article titles and also to quantify the popularity of these topics relative to all other economy related topics. We provide the annual frequencies and common words for each of the 110 topics of the final model for all 17 developed countries in the historical macro-dataset assembled by Jordà et al. (2016a), so that the two information sources can be used together in an economic analysis. We validate the dataset (the topics and their frequencies) by showing that the majority of the topics capture a coherent economic factor that makes sense to a human (and not just to a machine)

¹Recently, Greenwood et al. (2020) use historical banking crisis database newly assembled by Baron et al. (2020)—that starts from the year 1870 and covers 46 countries—to show that the interaction of rapid credit and asset price growth in the previous three years makes a crisis much more likely in the following years.

and that the frequencies of the topics capture historical events across topics, countries, and time. Although newspaper articles have increasingly been used in economic analysis in recent years, we are the first to derive topic frequency series from global newspapers for a large panel of countries. In addition, the topics that we estimate are not time or region specific, but they can be found from the whole time span of 150 years and across all 17 countries.

As a second step after first dissecting the text information accurately, we contribute to the literature that aims to understand and predict financial crises by adding a new, important, and previously absent element: text data. We show that throughout history, shifts in the attention to specific latent topics or narratives of newspaper articles precede financial crises. These leading indicators derived from text data outperform macroeconomic and financial data in crisis prediction and most importantly, significantly improve the predictive ability of these models even when the usual indicators are included in the model. Together, the best subset of the text topic series adds around 13 percentage points to the prediction accuracy of the baseline model (containing only macroeconomic and financial variables), bringing the overall classification ability to more than 87%. Even when only five topic frequencies are added to the baseline model with macro and financial variables, the prediction accuracy improves around 7 percentage points. In the 75 crisis episodes found in our dataset, the model using information from the topic-based indicators gives a higher crisis probability than the baseline model for 57 crisis episodes². The superior performance of the model using text data is not specific to crisis types or specific time periods, and it outperforms the baseline model across time before and after the World War II as well as in out-of-sample prediction experiment. These results are confirmed to hold in a large set of robustness checks.

Our text-based leading indicators contain text topics that, for instance, cover issues of *economic policy agreements and cooperation, foreign investments, commodities, international trade restrictions, money markets, diplomatic relations, high-level and government economists, firm-level sales and purchases, crises and dangers, promises, plans and ideas, relief and help, matches and fights, seasonal reporting and bookkeeping, noble metals, jobs, investigations and charges, reports on business conditions, factory problems, and national supply shortages*. Separate case studies of specific crisis episodes show that the topics represent general factors like external shocks, changes in economic policies, economic atmosphere and detailed production/output information, connected to events, and at least partial causes, for the specific historical crisis episodes. For example, the superior performance of the topic-based model for the Japanese financial crisis in 1990 is mostly due to the disappearance of the topic covering issues on *international trade restrictions on imports*. This indicator most likely captures the relevant aspects of the Japan–US trade war in the first half of the 1980s. The trade war ended after the Plaza accord was signed in 1985, which launched the strong appreciation of the yen with respect to the dollar, which further lead to financial speculation and a large asset price bubble. Similarly, the Portuguese financial crisis in 1890 was caused by the large accumulation of public debt and

²In the 18 episodes, where the baseline model gives a higher probability, the difference in probabilities was, on average, only 2.3 percentage points, whereas it was 12.9 percentage points in the cases where the text-based model gave higher probabilities.

a shock to the demand for Portugal’s main export. The increase in the *promises* and *high-level and government economists* topics before the crisis contributed the most to the increased risk of a crisis according to our model. These topics likely capture the change in economic policies led by the government *Partido Regenerado* and the financial minister *De Carvalho*, which ultimately led to the accumulation of public debt.

The results of this paper imply that there is much more generalizable public information than just the usual macroeconomic and financial indicators that is relevant for signaling a financial crisis in the future. From the perspective of policy makers, news article information can be used to form more accurate early warning models for financial crises. Text data can be used to measure more specific and detailed general factors that the usual economic statistics may not capture. The importance of these factors to financial crises may be of general knowledge, but incorporating them into prediction models has been difficult in previous research. The results also point out that there is still a lot that we do not understand about financial crises as our theory-implied indicators cannot explain the large difference of 13 percentage points in the prediction accuracy.

The article is organized as follows. In section 2, we show how we form our new text-based measures with topic modeling that incorporates historical news article data. In section 3, we present the results for the financial crisis prediction analysis and various robustness analyses of the main results. In section 4, we focus on the interpretation of different topic-based leading indicators and analyze the relationship between historical crisis episodes and the new leading indicators with several case studies. Finally, in section 5, we conclude.

2 Quantifying economic news articles

2.1 Historical newspaper data

We collect a large corpus of historical newspaper texts from ProQuest Historical newspapers. Our aim is to build a collection of global newspapers that would have reached readers on a global scale for the last 150 years. The newspapers covered are *The New York Times*, *The Washington Post*, *The Globe and Mail*, *The Times of India*, *The Guardian*, and *The Wall Street Journal*. We collect the titles of the 200 most relevant³ economic articles for each newspaper-year-country combination that mention at least one economic keyword and country name⁴. The countries included are Australia, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom (UK), Italy, Japan, The Netherlands, Norway, Portugal, Sweden, and the United States (US).

³Relevance is defined by the ProQuest search engine.

⁴The list of 104 economic keywords was assembled with a search and error procedure in which different economy-related keywords were used to find economy-related news articles from Proquest historical newspapers. After the news articles were collected, a random manual audit was performed by reading the articles of specific decades and countries. During this procedure, texts and keywords that had selected them were removed if the article was not related to the economy. Ten percent of the different decade-country combinations were randomly selected for the audit. This process resulted in the removal of keywords based on the root word "work", because they mostly resulted in sports- and art-related articles.

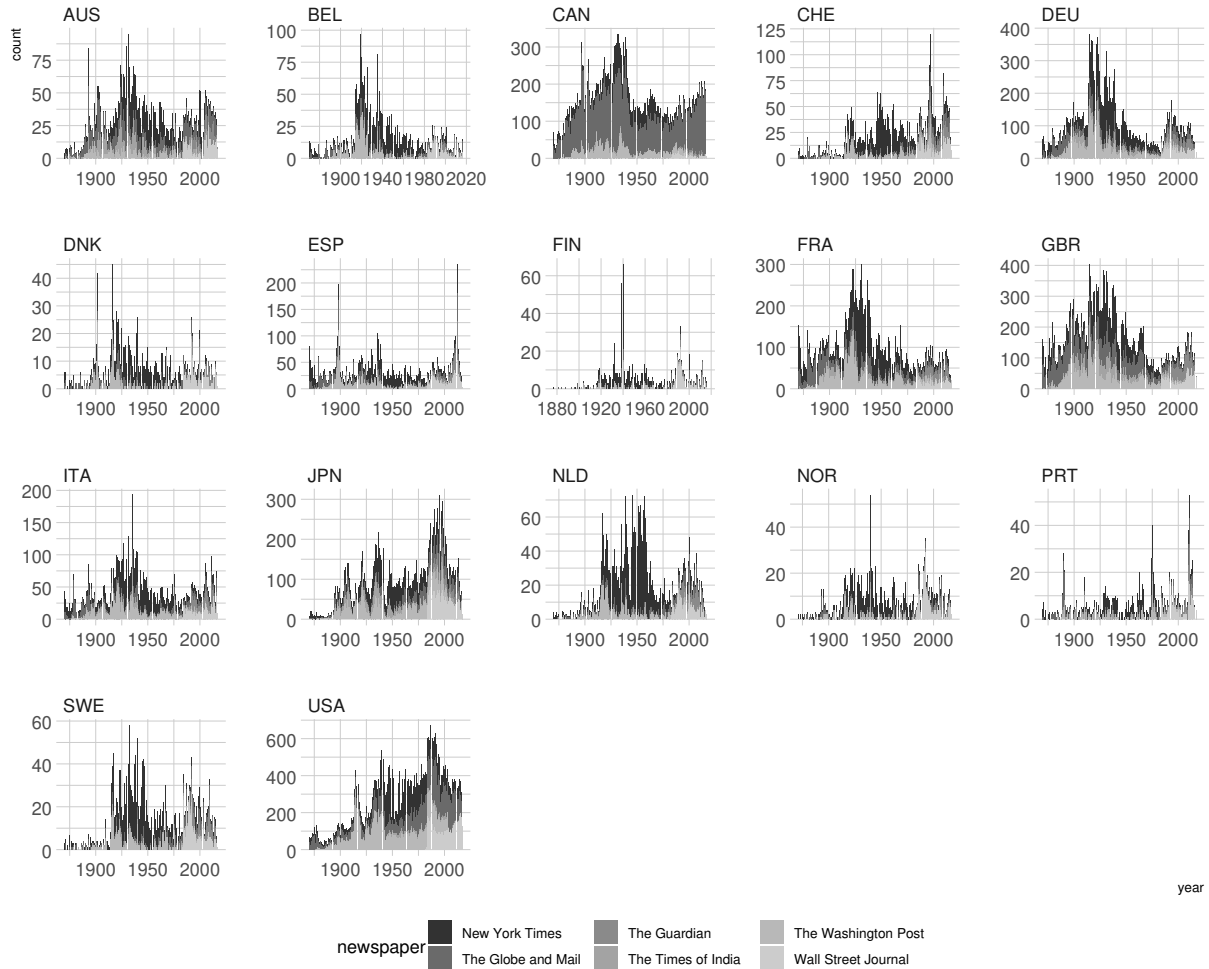


Figure 1: Annual article counts in time and across countries.

After an extensive text audition, we decided to focus on the article titles⁵. Each article has to mention the name of the country and at least one of the economic keywords in the article title. This requirement makes it more likely that the chosen article discusses about economic topics related to that specific country, when compared to the procedure where it would be enough that the keywords are mentioned in the full text. We use the titles in the following analysis instead of the full texts, because the lengths of the texts vary more among the latter than in the former. In addition, article titles explicitly summarize the content of the news articles, whereas the full texts also include a lot more useless content. The use of titles also keeps the data collection process feasible. With this text collection procedure, the final corpus includes 171.711 economy-related article titles.

Figure 1 displays the distribution of the newspaper article titles across countries, years, and newspapers. It is no surprise that the United States has the largest number of articles mentioning the economy—more than 40,000. Behind the US, Japan, Germany, the United Kingdom, Canada, and France form a group with around 15,000–25,000 articles. Italy, Spain,

⁵The manual audit is described more thoroughly in the Appendix.

Switzerland, The Netherlands, Belgium, and Sweden are mentioned in 2,300–7,000 articles. The countries with fewer than 2,000 total articles are Portugal, Denmark, Sweden, and Finland. We do not see the large dispersion in country articles as a problem as the number of articles reflects a country’s economic importance and relevance in the world. For example, it can be seen that the number of articles mentioning Japan has increased gradually since 1950s reflecting the change in the country’s role in the global markets. Another good example is the UK, for which the number of articles decreased gradually from the 1900s onward as it is losing its place as a global superpower to the US. Finally, as an example, the number of articles mentioning Finland reflects well how a small country is discussed on a global scale. The peaks for Finland occur during the Second World War and in the 1990s when this small Nordic country experienced one of the worst banking crises in history and also recovered partly due to the rise of mobile phones and Nokia. The data collection procedure is described in more detail in the Appendix.

2.2 Topic modeling

We use a topic model to learn latent topics in the historical newspaper corpus. The most commonly used topic model is LDA (Blei et al., 2003), which is an unsupervised learning model (when referring to the classification of supervised and unsupervised statistical and machine learning models). These models uncover structures in the data without having any knowledge of the data labels. In this context, we do not provide the model the information that some news are titled as, for example, sports or economic news. The LDA model assumes that each news article title is generated by a generative process where each title is possibly a combination of multiple topics (title can be represented by distribution of topics) and each topic is a distribution of words from a fixed vocabulary. The word latent refers to the latent nature of the topics in the text corpus, and the word Dirichlet refers to distributions from which the topic and word distributions are assumed to be generated.

Let us assume that there is a corpus of T news titles with N words in each title from a vocabulary of size V and a total number of K topics. The generative process from which we assume the titles to be generated can be presented in the following way. For title d there is a distribution $\theta_d \sim \text{Dirichlet}(\alpha_1, \alpha_2, \dots, \alpha_K)$ of topics, and for each topic k , there is a word distribution $\beta_k \sim \text{Dirichlet}(\eta_1, \eta_2, \dots, \eta_V)$. The actual title d is generated so that for each word position j in title d a topic $z_{d,j}$ is generated from the topic distribution so that $z_{d,j} \sim \text{Multinomial}(\theta_d)$. Next, a word $w_{d,j}$ is generated for each word position j in title d from the word distribution of the assigned topic $z_{d,j}$ so that $w_{d,j} \sim \text{Multinomial}(\beta_{z_{d,j}})$. Smaller values of the parameter α of the Dirichlet distribution make the titles consist more of single topics, whereas larger values force titles to consist more of multiple topics. The process can be more formally written as the following joint probability function:

$$P(\theta, \beta, Z, W) = \prod_{k=1}^K P(\beta_k | \eta) \prod_{d=1}^M P(\theta_d | \alpha) \prod_{i=1}^N P(z_{d,n} | \theta_d) P(w_{d,n} | \beta, z_{d,n}). \quad (1)$$

In this generative process, we know only $w_{d,j}$ and the number of topics K , as it is predetermined by the researcher. The distributions, the topic assignments, and the other parameters are unknown. We can estimate the following posterior distribution in Eq. (2) with Gibbs sampling:

$$P(\theta, \beta, Z|W) = \frac{P(\theta, \beta, Z, W)}{P(W)}. \quad (2)$$

This procedure searches for the parameter values that make the process to generate newspaper article titles that are as close to the ones that we observe in the actual text corpus.

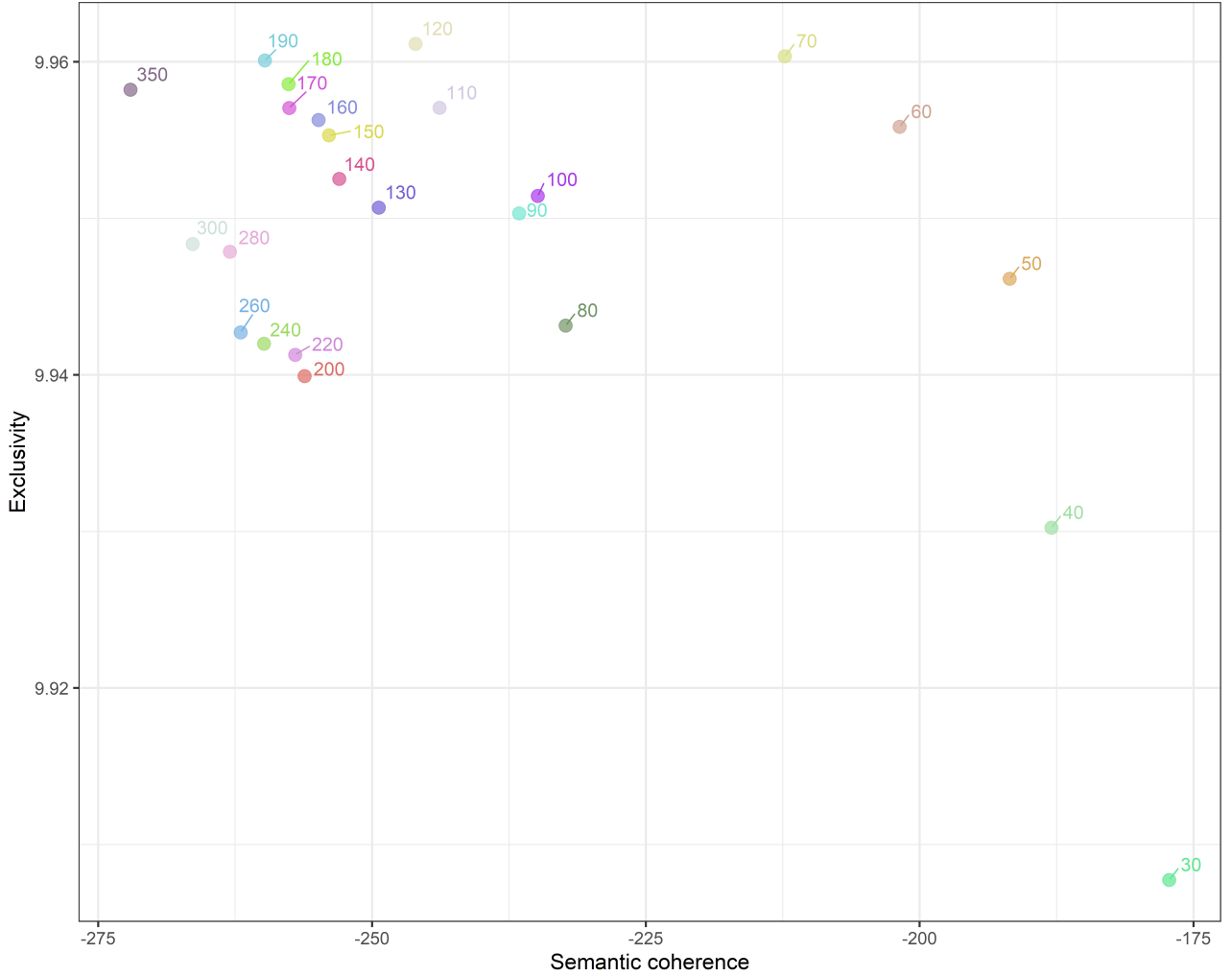


Figure 2: The topic model’s semantic coherence and exclusivity for different values of K (pre-determined number of topics). Semantic coherence measures how often the most frequent words of a given topic co-occur in the news titles; the higher the number of topics, the lower the probability of co-occurring topic top words. Exclusivity measures how unique the most frequent words of a given topic are; the higher the number of topics, the higher the probability of unique topic top words.

As the number of topics K is a predetermined hyperparameter of the model, the optimal value of K has to be estimated via a grid search. We use a model’s exclusivity and semantic coherence (Mimno et al., 2011) together with human judgment as criteria to select the optimal

value for K . Semantic coherence is high when the most frequent words in a topic co-occur frequently. Exclusivity is high when a topic’s most frequent words are not frequent in other topics. There is a trade-off between these two measures and the number of topics. When there are many topics, a model’s exclusivity is often high, and its semantic coherence is often low, and vice versa when there are only a few topics. This means that often human judgement is the final criteria to select the optimal amount of topics. Semantic coherence and exclusivity for different values for K are plotted in Figure 2. We chose the topic model with 110 topics as it holds a sufficiently large number of topics to describe the large news coverage of the text corpus, and the majority of the topics were identifiable at an adequate level. Based on our findings of our new title corpus and by using human judgement, 110 as the number of topics is sufficiently large to distinguish certain important themes from each other, as well as, small enough not to have too narrow, event specific, topics. Finally, we form the annual topic frequency $TF_{k,t,c}$ for a specific topic k by averaging the proportion of each topic k across the L articles that are published in year t and mention country c . More formally,

$$TF_{k,t,c} = \frac{1}{L} \sum_{l=1}^L \theta_{k,t,c}. \quad (3)$$

A few authors have implemented LDA in economic research in different applications. Larsen et al. (2021) estimate a topic model by incorporating all news from the Dow Jones newswire archives from 1990 to 2016. The authors show that a subset of the topics predict inflation and inflation expectations. Bybee et al. (2020) construct a topic frequency series of the *Wall Street Journal* business articles from 1984 onward and show that these topics explain fluctuations in economic variables well. In Bybee et al. (2021), the authors use topic modelling in an asset pricing framework and extract risk factors from news narratives. Our methods and purpose differ in three important aspects from theirs. First, we construct the series from 1870 onward from articles in six global newspapers. Second, we form the series separately for each country by identifying economy-related articles mentioning specific countries. Third, we clean the text data so that we remove all time- and country-variant words from the corpus before we estimate the topic model. The purpose of using this procedure is to get general and timeless topics that can be utilized in an econometric analysis. This way, we do not get topics that discuss “World War II” or “Ronald Reagan”, but general versions of these specific topics, like war and presidents. These procedures are needed to analyze the relationship between financial crises and text-based information, because crises are rare events that occur relatively seldom—every 28 years on average (Jordà et al., 2010).

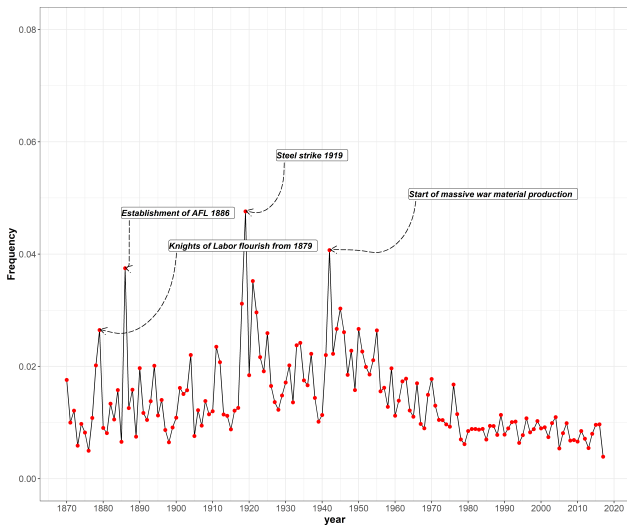
frequent in a specific topic but less frequent in others. This way, we do not put too much emphasis on words that are very common across many topics. An example of an easily interpretable topic is topic 4 (*economy-growth-spur-signs-grows-slowng-pace*), which clearly discusses the direction and magnitude of economic growth. Topics 5 (*central-bank-governor-institution-raises-eyes-credit*) and 6 (*banks-savings-accounts-cash-lending-institutions-fail*) are likely to loosely cover issues on central banks and banking. Furthermore, topics 12 (*investors-briefs-look-wary-feel-yields-interest*), 37 (*labor-party-unions-leader-strikes-reds-vote*), and 55 (*gold-metal-standard-silver-reserves-reserve-shipment*) could be about views on financial markets, left-wing politics, labor unions and strikes, and noble metals.

Not all topics are as interpretable as others from only looking at the seven most exclusive words, which is why we also look at the most representative titles of each topic to identify the content of the topic. By viewing the titles that have the greatest share of a given topic in the title-topic distribution we can attain better insight of the actual content of that particular topic, and hence it also helps labeling the topics. With this procedure, we are able to label more than 82% of the topics clearly, and only one topic remains completely unnamed. We have to keep in mind that the corpus used in the model estimation is cleaned from the impact of time- and country-variant words and names, which makes the interpretation task a little harder. This implies that a topic that discusses war, for example, cannot have words like “Vietnam” or “Iraq” included in it. These kinds of context words would make the identification of topics easier, but they would also make topics time, country and event specific. This would make the frequency series possibly increasing only during a single period or in a single country, which we do not want as the topic frequency series should be as common as possible in the sense that their frequencies could be monitored all the time, and they could possibly be relevant in any country or year.

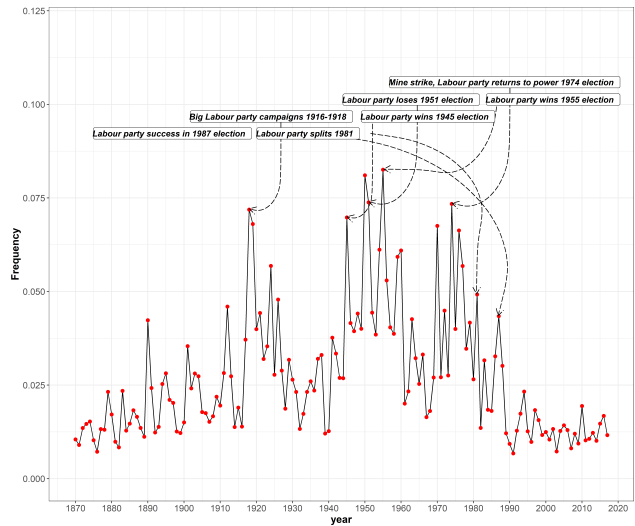
2.3.2 Association with historical events

Second, we want to make sure that the time series of the topic frequencies captures major historical events related to issues that the topic is supposed to approximate. Next, we look at examples of the topic frequencies TF derived from the topic distributions θ_d . Figures 4a–4c plot the frequency of topic 37 labeled *labor, unions and strikes* for the US, the UK and France from 1870 onward. In the United States, the frequency of this topic spikes in 1879, when Knights of Labor started to flourish, in 1886 when the American Federation of Labor (AFD) was founded, during the Great Steel Strike of 1919, and in 1942 when massive labor-intense war material production started. In the United Kingdom, the series captures the great losses and wins of the Labor Party and the mine strike in 1974. In France, the series is elevated during major strikes in 1974, 1984, 1997, 1998, and 2016. It peaks in 1942–1944 during the so-called *service du travail obligatoire*, when thousands of French workers were sent to Nazi Germany for hard labor.

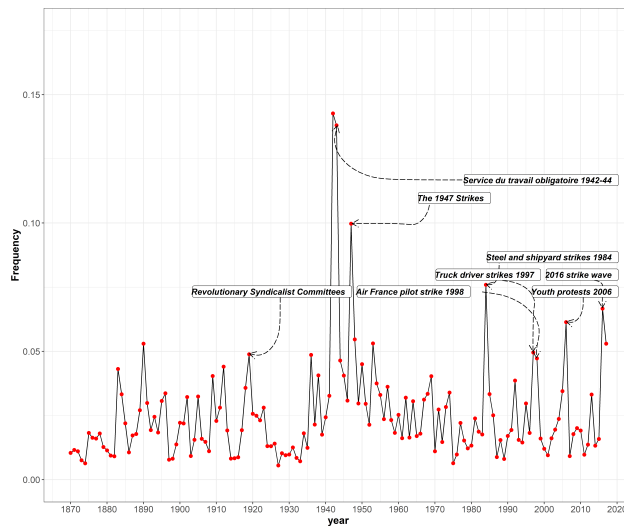
Figures 5a–5c plot the frequencies of topic 28 for the United States, Finland, and Japan. The most common exclusive words for this topic are *banking-system-giant-scandal-beats-sector-*



(a) United States

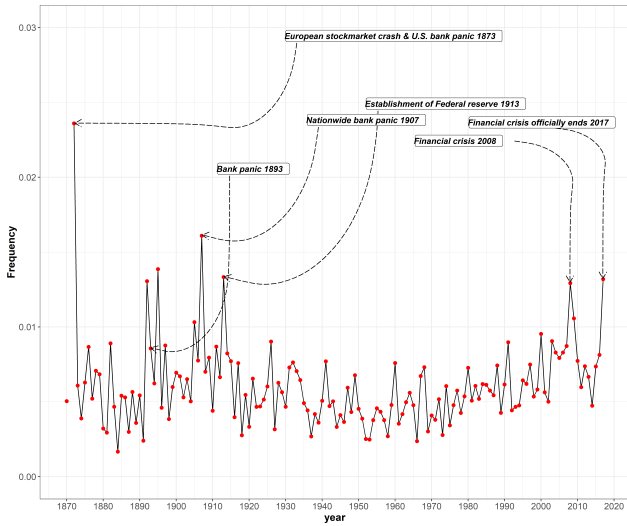


(b) United Kingdom

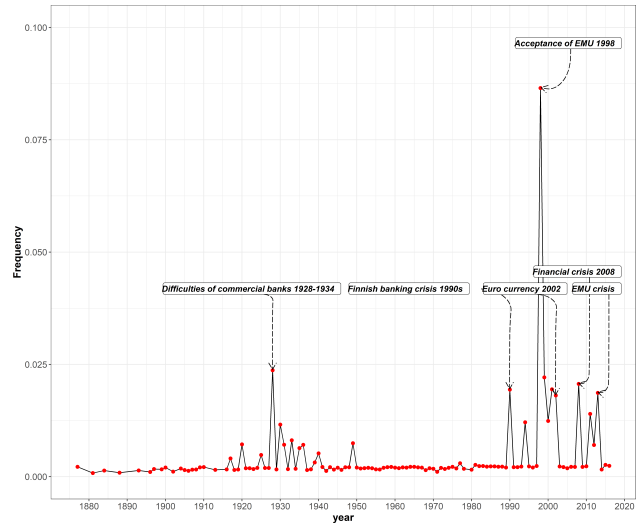


(c) France

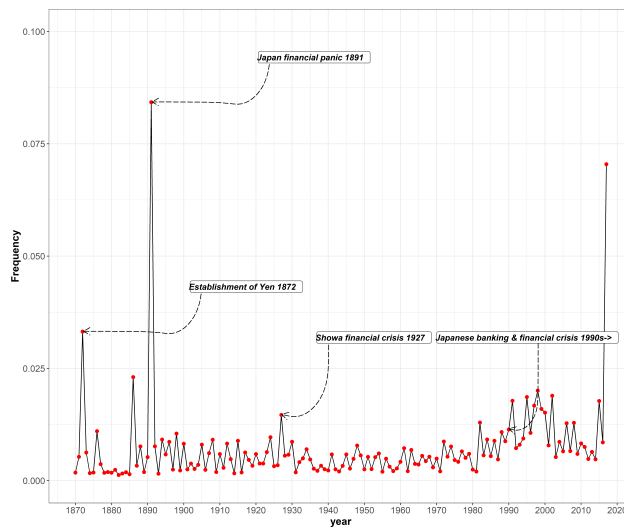
Figure 4: The frequency of topic 37 *Labour, unions and strikes* and historical events in selected countries.



(a) United States



(b) Finland



(c) Japan

Figure 5: The frequency of topic 28 *Banking crises* and historical events in selected countries.

crises, which point to banking system difficulties. Thus, the topic was labeled as a *banking crises* topic. For the US, the series spikes during the panics of 1873, 1893, and 1907 and the global financial crisis (GFC) of 2008. In Finland, the series co-moves with similar events, such as the difficulties of commercial banks in 1928–1934, the Finnish banking crisis in the 1990s, and the GFC of 2008. The series is also elevated in 1998 and 2002, which were the years when Finland was accepted into the European Monetary Union (EMU) and when the country started using the euro as currency. Similarly, in Japan, the series rises when the yen was established in 1872, during Japan’s financial panic in 1891, the Showa financial crisis of 1927, and most of the 1990s when Japan experienced a severe financial crisis and a recession. This visual evidence indicates that the topic frequency series captures relevant and significant historical events that are related to the topic.

2.3.3 Topic distance and prevalence

To see how these different topics are related to each other and which ones are more prevalent than others, we plot the so-called ”inter-topic distance map”, where the different topics are placed according to their word probability distribution and their size according to the overall prevalence of that topic. To measure the inter-topic distances, we use the Jensen-Shannon divergence metric with multi dimensional scaling (MDS). For the scaling, we use the classical Principal Components algorithm. (Sievert & Shirley, 2014)

Figure 6 shows the inter-topic distance map for the dataset. The figure shows that topic 46 (*financial market movements*) is the most prevalent topic; that is, it has, on average, the highest share across all news article titles. Topics 20 (*telegraphs and cables*), 103 (*stock exchange*), 1 (*positive market movements*), 25 (*economic policy agreements and cooperation*), 4 (*economic growth*), 60 (*pacts and treaties*), 5 (*central bank*), and 37 (*labor, unions, and strikes*) also have visibly higher prevalence than many other topics.

In addition, it can also be seen from the figure that the topics form several clear clusters. That is, certain topics are more closely related to each other than to the rest of the topics. Many of the prevalent topics are present in these clusters. In the lower left corner of Figure 6, there is a cluster that includes topic 103 (*stock exchange*), topic 46 (*financial markets movements*), topic 65 (*positive quarterly corporate reports*) and topic 18 (*specific location events of interest to markets*). The second cluster in the lower right corner is formed by topic 20 (*cables and telegraphs*) and topic 105 (*communities and institutions*), which include more general society-related news. Topic 49 (*government level conflicts and war*) and topic 75 (*public appointments*) are also relatively near this cluster, which can also be seen as more general news affecting the whole economy. In the right middle, there is a cluster of topics that includes topic 37 (*labor, unions and strikes*), topic 48 (*government bills, resolutions and statements*), topic 14 (*commercial treaties*) and topic 33 (*high officials*). This cluster seems to be focused on news concerning the public sector. Topic 82 (*diplomatic relations*) is also relatively close to the cluster, which seems reasonable. In the right upper corner, there is a clear distinct cluster including topic 60 (*pacts and treaties*) and topic 78 (*free trade difficulties and barriers*). This

cluster possibly consists of news about different kinds of important agreements and restrictions on trade. In the upper middle right, there is a small cluster of three topics: topic 17 (*competition in international trade*), topic 83 (*change in international trade barriers*) and topic 30 (*international trade restrictions - imports*). This cluster quite clearly covers topics that include news about international trade. In addition, there is a cluster that includes topics 4 (*Economic growth*), 3 (*Industrial production*), 56 (*Exchange rate movements*), 24 (*Unemployment rate*), and 67 (*Periodic production numbers*). This cluster is formed with topics covering the news about macroeconomic conditions. Furthermore, one notion is that topics 1 (*Positive market movements*) and 39 (*Company acquisitions*) are clearly not related to any other topics; that is, they are far away from other topics and also from each other.

Finally, the ambiguous topic 110 has a very low relative share across the titles (document-topic distribution) and seems not to be related to any other of the topics, which we take as a good sign, confirming our previous thoughts that this particular topic has no clear interpretation. In conclusion, the finding that the topics included in the same clusters are closely related to each other can be seen as supporting evidence for the view that our interpretation of the contexts of the topics is appropriate.

3 Predicting financial crises

3.1 Topic attention and common crisis indicators

The financial crisis literature has proposed a number of potential crisis indicators, many of which have been empirically confirmed to have predictive power regarding crises. We analyze how these commonly used and accepted financial crisis indicators perform when compared to our newspaper topic frequency variables. As shown in the previous section, the topic frequencies relate to different economic variables and capture significant economic events. That is why it is possible that these variables have additional, more detailed information on the same issues that the usual leading indicators are supposed to capture, like asset price and credit booms, and thus improve prediction accuracy. Furthermore, the topics are likely to have information about a much wider set of issues for which we have not had numerical measures (data) previously, which can also improve prediction accuracy. Shiller (2017) argues that many historical crisis episodes have been caused at least partially by narratives. News articles capture and create general narratives that either get attention or not. This attention might last for a long time, fade away or resurface at a point in time. Our 110 individual topic frequency series can be seen to measure the attention to specific general narratives that are related to the economy, as we do not want to limit the topic model to extract only the topics that might be related to financial crises. Therefore, we are able to get more detailed information about the relative importance and the coverage of all of the economy related topics, and thus, also themes that we ex-ante do not necessarily associate with financial crises.

To analyze the predictive power of topic frequencies for financial crises, we estimate a panel logistic regression (Eq. (4)) with a binary financial crisis variable as the dependent variable

and lagged three-year moving averages of potential crisis indicators as explanatory variables⁷. The potential new topic based crisis indicators are selected to the prediction model with the commonly used forward stepwise selection procedure utilizing the Akaike Information Criteria (AIC). More formally,

$$\log \left(\frac{Pr[Y_{i,t} = 1 | X_{i,t-1}]}{Pr[Y_{i,t} = 0 | X_{i,t-1}]} \right) = \alpha_i + \beta X_{i,t-1}, \quad (4)$$

where $Y_{i,t}$ represents the binary crisis indicator variable indicating whether there is a financial crisis ($Y_{i,t} = 1$) in country i at year t using predictive information contained in $X_{i,t-1}$. We use data for this crisis indicator variable from Baron et al. (2020), who improved previous historical crisis chronologies by also including crisis episodes without panics by investigating bank equity declines. The potential indicators in the model include change in credit to GDP, return on equity, real GDP growth, and inflation. We also include the change in house prices, bank leverage, bank non-core funding ratio, and bank liquidity as separate robustness tests because including these variables decreases the sample size significantly. In addition to these macroeconomic and financial variables from Jordà et al. (2016a)’s and Baron et al. (2020)’s historical databases, the topic frequencies that were estimated in the previous section are included as predictors in the model (via $X_{i,t-1}$) as lagged three-year moving averages. We include country fixed effects in the model to control for the unconditional probability of a crisis that might differ across countries and compute standard errors of the estimated coefficients according to Driscoll and Kraay (1998). The main data used in the estimation includes observations from 17 developed countries for the period 1870–2016 and a total number of 75 separate crisis observations.

For the purpose of comparing the prediction ability of the usual crisis indicators and the topic frequencies, we estimate separately a baseline model composed of the common macroeconomic and financial crisis indicators and then different models that include topic frequency variables with or without the usual macroeconomic and financial indicators. The models that include only topic-based variables can be used to assess how these variables perform on their own. The specifications that also include the common crisis indicators can be used to assess whether newspaper topics add to the predictive ability of the model, or in other words, include additional useful information about financial crisis probability in addition to the usual indicators.

Throughout the analysis and following the common practice in closely related crisis prediction studies, we assess the prediction performance of each model with the area under the receiver operating characteristic (ROC) curve (AUC). As a brief introduction of the AUC, we predict a crisis when the fitted crisis probability (see (4)), later also out-of-sample probability forecast in Section 3.2.2) exceeds a probability threshold $c \in (0, 1)$. The ROC curve plots each combination of a true positive rate (sensitivity, crisis years labeled correctly) and a false positive rate (i.e., 1-specificity, “normal” years incorrectly labeled as crises) when varying the

⁷Jordà et al. (2016b) and Jordà et al. (2020) use five-year moving averages of the main explanatory variables. Greenwood et al. (2020) use three-year changes, and Schularick and Taylor (2012) use the five lags of the explanatory variables.

threshold c . Finally, the AUC, the area under the ROC curve, gets values between 0 and 1, where 0.5 means that the prediction accuracy is equal to a random coin toss, and a value of 1 equals a perfect classifier. For a detailed introduction to the ROC curve, see, e.g., Berge and Jordà (2011).

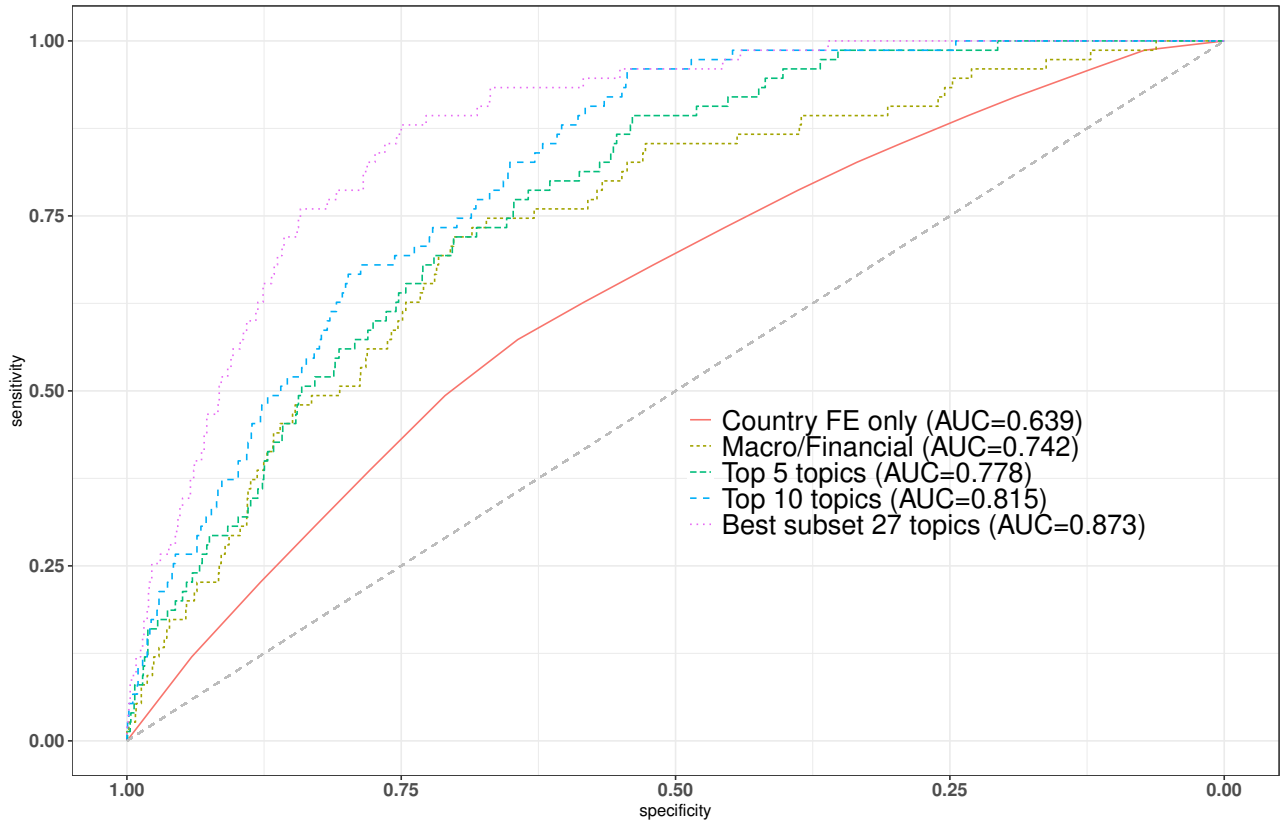


Figure 7: ROC curves for financial crisis prediction models. Model with macroeconomic and financial indicators vs. models with topic frequencies only (i.e. without macroeconomic and financial indicators).

The ROC curves for the baseline model (with the macroeconomic and financial variables only), a model with only country fixed effects, and for three model specifications with only topic frequencies (text data) included as predictive variables are plotted in Figure 7. As the number of potential new topic-based crisis indicators to choose from is very large, we use a forward stepwise selection process to select the optimal prediction model that we refer to as the best subset model. This procedure starts with a null model with only country fixed effects included and then variables are added one by one to the model by first choosing the variable that decreases the AIC value of the model the most. Then the variable that decreases the model the second most is added. This procedure is continued until adding new variables no longer decreases the model's AIC value. In this context, with only different topic frequency variables in the selection pool, the model selection procedure selects 27 different topic frequency variables out of 109 topics⁸.

⁸We excluded topic 110 from the model selection as we were not able to clearly label it.

The AUC value of around 0.74 for the baseline model is very similar to what has been obtained in the literature where the entire timespan of almost 150 years has been included in the analysis. This shows that the usual indicators found in previous studies, such as the change of credit to GDP ratios and equity returns, are already able to achieve modest predictive power. However, the model with the best subset of 27 topic frequency indicators is able to achieve an AUC value of around 0.87, which is more than 13 percentage points higher than with the baseline model. However, there are far more variables in this model when compared to the baseline model with only four indicators. We also estimated the models where we utilize only the 5 and 10 best topics that were added by the forward stepwise selection procedure. Surprisingly, even the model with five topic variables is able to achieve an AUC value of 0.78, which is still statistically significantly higher than the corresponding value of the baseline model.

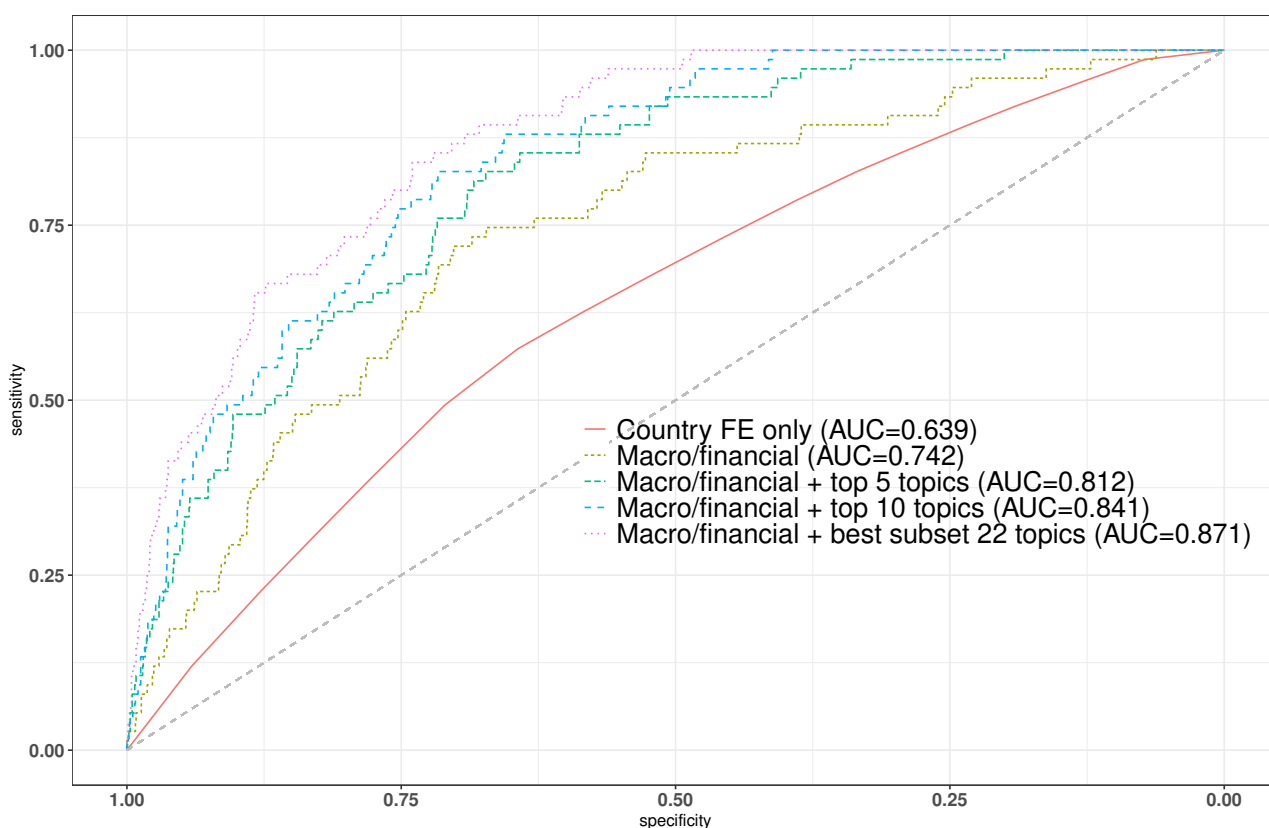


Figure 8: ROC curves for financial crisis prediction models including macroeconomic and financial indicators with and without topic frequencies.

Next, we estimated models where the null model (“Macroeconomic and financial”) included all macroeconomic and financial indicators, and then the stepwise selection procedure was given the opportunity to add topic variables to the model in addition to the usual indicators. The model that included the best subset of 22 topic variables and these common crisis indicators also performed very well and achieved an AUC value of 0.87. That is basically the same as with the model that included only the 27 topic series without any of the common crisis indicators. When only 5 or 10 of the first selected topic variables are included in the model with the common indicators, the AUC values are still remarkably higher relative to the baseline model.

The corresponding ROC curves are plotted in Figure 8 and reported in more detail in Table 1.

Table 1: Financial crisis prediction results of different logit regression specifications.

	AUC	Diff.to.baseline	p.value	Diff.upper	p.value
Country FE only	0.639				
Macro/financial	0.742				
Macro/financial + top 5 topics	0.812	0.070	0.000		
Macro/financial + top 10 topics	0.841	0.099	0.000	0.028	0.002
Macro/financial + best subset 22 topics	0.871	0.129	0.000	0.030	0.002
Top 5 topics	0.778	0.036	0.053		
Top 10 topics	0.815	0.074	0.002	0.038	0.010
Best subset 27 topics	0.873	0.131	0.000	0.058	0.000

Notes: In this table, in addition to the AUC statistics in the second column, the statistical difference in the AUC of the benchmark model (Macroeconomic and financial predictors only) or the model reported in the row above is tested, with the p-values using the DeLong et al. (1988) test.

In Table 1, we also present the differences between AUCs of different models relative to the AUC value of the baseline model. In addition, we report the corresponding p-value of DeLong et al. (1988)'s statistical test where the null hypothesis is that the AUC value of a model is not larger than in the baseline model. In all six model specifications that include topic frequency variables, the AUC value is statistically significantly higher than in the baseline model. This implies that newspaper article titles and more specifically, the time-varying attention to specific topics in these articles contain highly useful information about the risk of a financial crisis in the near future. In column (4) of Table 1, we present the difference of the model's AUC value relative to the model that included more topic variables. The p-values in column (5) reveal that adding more topic variables to the model, that is, moving from 5 to 10 topics and from 10 to 22 topics, improves the prediction ability even more statistically significantly.

Another way of comparing the predictive power of different models is to compare the actual crisis probabilities that the models produce for the actual crisis years. The difference of the crisis probability for all 75 crisis episodes between the model that includes the best subset of 22 topic indicators in addition to the macroeconomic and financial variables and the baseline model is plotted in Figure 9. The figure shows that for a large majority (57 out of 75) of crises, the model using topic frequencies gives a higher crisis probability for the crisis year. The difference is very large in many cases. For 27 episodes, the difference is larger than 10 percentage points, and for 15 episodes, it is larger than 20 percentage points. In the extreme cases, like the crises in Japan in 1890, Denmark in 1907, Portugal in 2008, and Belgium in 2011, the model using topic frequencies as indicators gives 35- to 47-percentage-points larger crisis probabilities relative to the baseline model.

As prediction ability in this context means how well the probability estimates can be used to classify crisis years from normal years, it is enough that the model gives a 2% crisis probability for crisis years and 1% for normal years. In this case, the model can be seen to classify crisis years from normal years perfectly, and it will achieve $AUC = 1$. However, in practice, when policy makers use these models to assess whether preventive actions are necessary, it would be

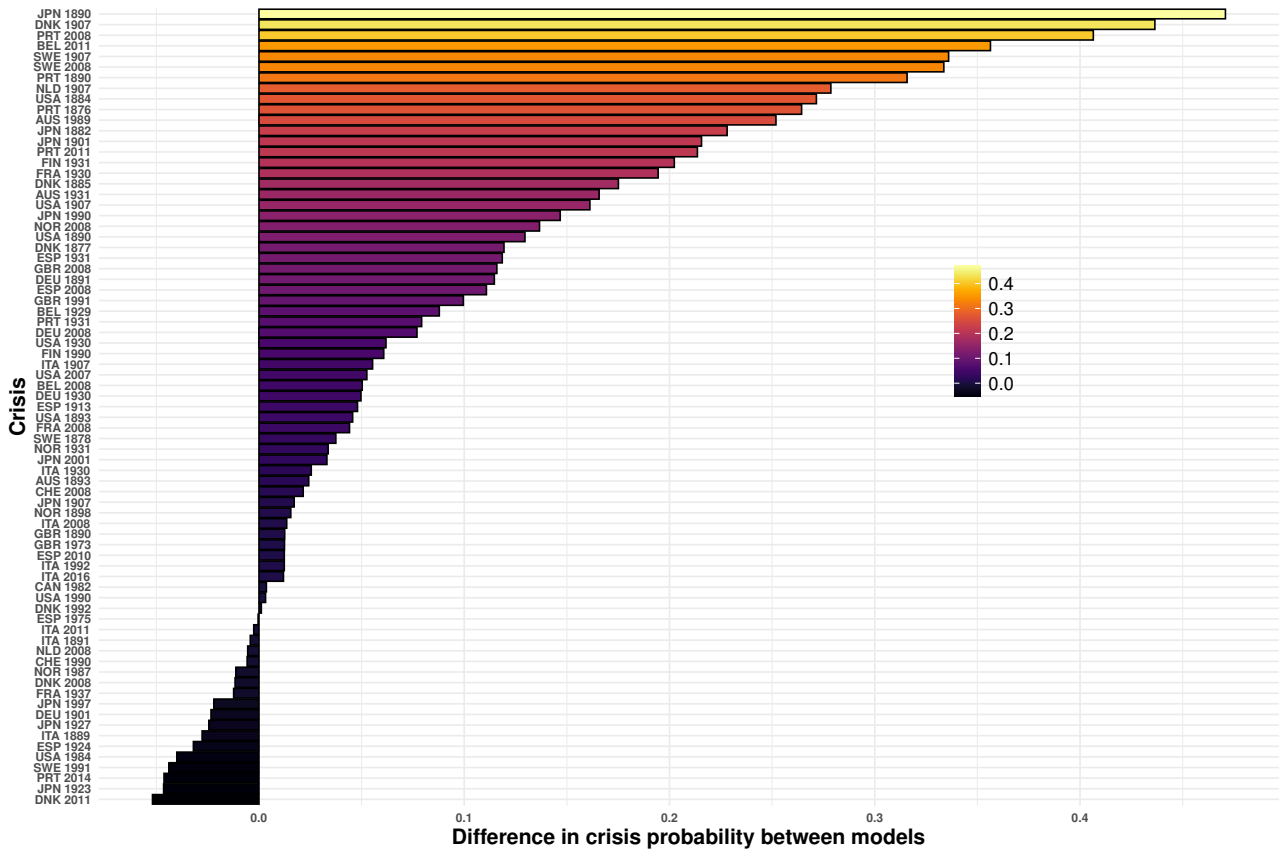


Figure 9: Differences in the crisis probabilities of all 75 crises using the models including macroeconomic and financial indicators with and without the best subset of topic indicators. Positive numbers imply the superior performance of the model with text data information.

desirable that the difference between these crisis probabilities is as large as possible to enhance the interpretability and information signals for a policy maker to trust the models. From this point of view, the topic-based model is better than the baseline model in 57 times out of 75. For the remaining 18 cases for which the crisis probability is larger for the baseline model, the difference is not that large: For 14 cases, it is smaller than 6 percentage points. The text based model outperforms the baseline model in 25 (or 69.4%) cases prior to World War II and in 32 (or 82.1%) cases after it. This implies that the superior prediction performance is not related to a specific time period, but the text based leading indicators outperform the baseline model consistently over time.

In the next section, we perform an extensive number of robustness checks on this main finding on a substantial improvement in prediction accuracy by using the information contained in the text data (Section 3.2). Furthermore, in Section 4, we also make the results more interpretable regarding the specific indicators and their contribution to different and specific crisis episodes.

3.2 Robustness analyses

3.2.1 Overfitting

When a crisis prediction model is estimated from a dataset with a large set of potential predictors and a relatively small number of cases (75 crises episodes), overfitting is a potential issue that can make the results look too optimistic and result into models with poor generalization ability. How likely is it that adding enough uninformative variables to a baseline prediction model increases the AUC value statistically significantly (in the same manner as our topics series) without any genuine information that would be relevant for crisis prediction? At the previous section we have already taken this possibility partly into account by using only the top 5 or top 10 topic series as new or additional predictors, which reduces the likelihood that the improvements in prediction accuracy are a result of poor luck. Although the best topic indicators can be seen to make sense on why they improve the prediction ability of a model and they fit to the 'stories' how historical crisis episodes have occurred, we still see it useful to statistically analyse these issues in more detail.

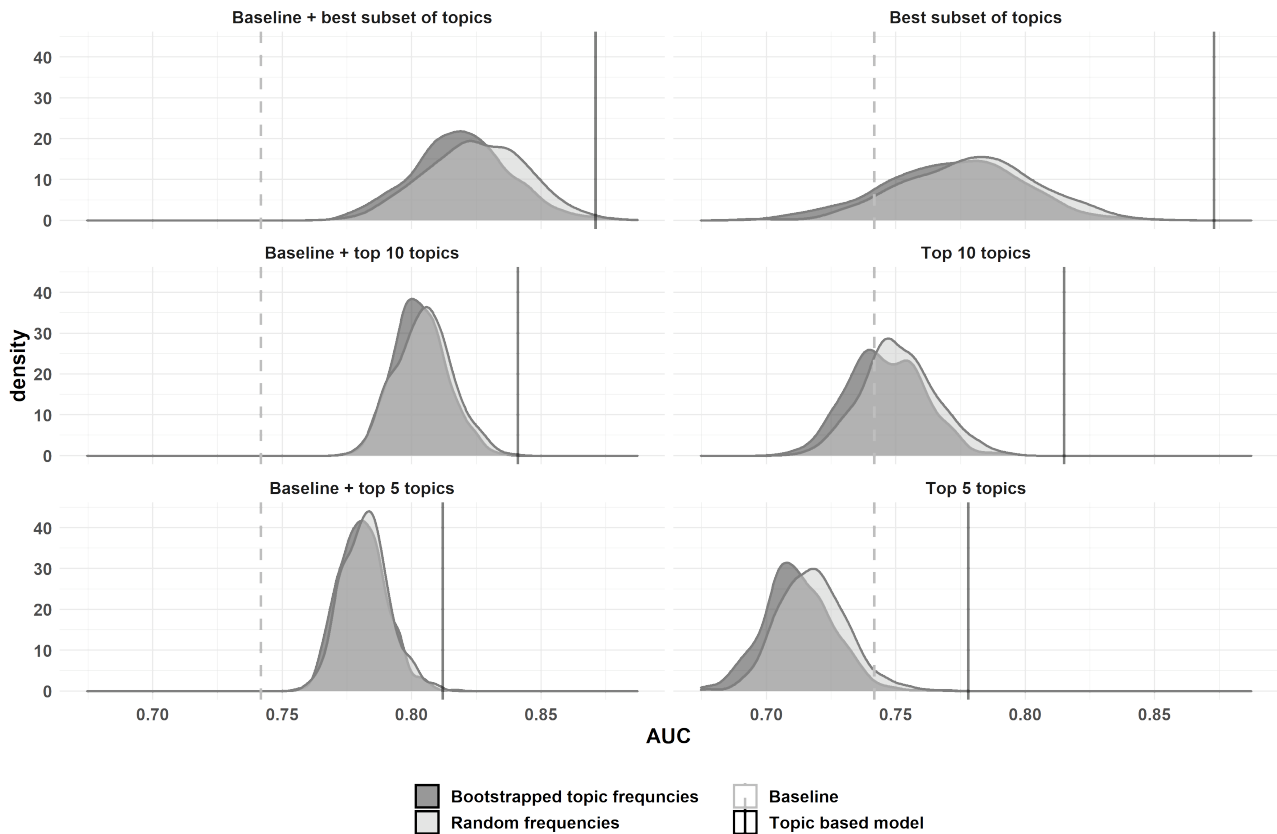


Figure 10: The distributions of AUC values for the topic based models estimated with randomly generated (light grey) or bootstrapped (dark grey) topic series. The distributions of AUC values are from 1000 crisis prediction models that are formed with the same forward-stepwise model selection procedure as the 'true' model, but with uninformative (random or bootstrapped) time series instead of the true topic frequency series. The dashed vertical line presents the AUC value of the baseline model with macroeconomic and financial variables together with country fixed effects. The solid vertical line presents the AUC values of the different topic series based models.

We estimate all the six prediction model specifications of the main analysis presented in the previous section with either bootstrapped topic frequencies⁹ or completely randomly generated topic frequencies. The purpose is to analyse how likely it is that the classification ability of a crisis prediction model improves statistically significantly with a combination of uninformative variables. When the random or the bootstrapped topic frequency series have been generated, we use the same model selection procedure that was used in the previous section. We generate 1000 new datasets where a half use bootstrapped topic frequencies and a half use completely random numbers between zero and one as topic frequencies. Although both procedures generate uninformative datasets from the perspective of topic frequencies and financial crisis, the former is more likely to have series that are more similar to the original ones.

Table 2: Robustness analysis: Comparing improvements in prediction ability between models estimated with original topic frequencies and uninformative time series.

Topic data	Model	$\frac{\#AUC_U > AUC_{BL}}{\#Total}$	$\frac{\#p < 0.05: AUC_U > AUC_{BL}}{\#Total}$	$\overline{AUC_U - AUC_{BL}}$	$\frac{\#AUC_U > AUC_{TR}}{\#Total}$	$\overline{AUC_U - AUC_{TR}}$
Bootstrapped	BL + best subset	1.000	0.151	0.077	0.006	-0.052
Random	BL + best subset	1.000	0.156	0.083	0.004	-0.046
Bootstrapped	BL + top 10	0.932	0.155	0.061	0.000	-0.038
Random	BL + top 10	0.978	0.158	0.063	0.000	-0.036
Bootstrapped	BL + top 5	1.000	0.091	0.039	0.002	-0.031
Random	BL + top 5	1.000	0.087	0.040	0.000	-0.030
Bootstrapped	Best subset	0.876	0.086	0.030	0.000	-0.101
Random	Best subset	0.932	0.090	0.037	0.000	-0.094
Bootstrapped	Top 10	0.556	0.082	0.004	0.000	-0.069
Random	Top 10	0.724	0.076	0.009	0.000	-0.064
Bootstrapped	Top 5	0.016	0.000	-0.030	0.000	-0.066
Random	Top 5	0.042	0.000	-0.024	0.000	-0.060

Notes: Abbreviations: *BL* (baseline model), *TR* (true model estimated with the original topic frequency data) and *U* (model estimated with data where original topic frequency series is replaced with uninformative random frequencies or bootstrapped original topic frequencies). Column 1: procedure of creating uninformative replacement series for topic frequencies. Column 2: the crisis prediction model specification. Column 3: share of prediction models estimated with uninformative topic frequencies whose AUC value was higher than the corresponding value of the baseline model. Column 4: share of prediction models estimated with uninformative topic frequencies whose AUC value was statistically significantly higher at a 5% significance level than the corresponding value of the baseline model. Column 5: average difference of AUC values between the models estimated with uninformative topic frequency data and the baseline model. Column 6: share of prediction models estimated with uninformative topic frequencies whose AUC value was higher than the corresponding value of a model estimated with the original topic frequency data. Column 7: share of prediction models estimated with uninformative topic frequencies whose AUC value was statistically significantly higher at a 5% significance level than the corresponding value of a model estimated with the original topic frequency data.

Figure 10 plots the distributions of the AUC values corresponding to the prediction models estimated with the new datasets. The figures show that using enough uninformative variables results into statistically significantly higher AUC values and hence improved prediction ability relative to the baseline model estimated from the original data. This seems especially clear in the case where uninformative series are added into a model that already includes the baseline macroeconomic and financial variables. On average, these models achieve AUC values that

⁹Stationary bootstrap by Politis and Romano (1994) with mean block length b .

are close to the middle point between the AUC values of the 'true' model estimated with the original data and the baseline model.

Table 2 gives a more detailed description of these differences. Around 8.7 to 15.8 percent of these models achieve a statistically significantly higher AUC value than the baseline model, but the corresponding AUC values are much lower than the AUC values of the models estimated from the original true topic series data. In the case with the baseline indicators and the best subset of topic series only 2-3 (or 0.4%-0.6%) out of 500 models estimated from the uninformative topic series reach a higher AUC value than the true model. In the case with baseline variables and top 10 topic indicators, none of the models outperform the true model and with top 5 indicators only 0-1 (or 0.0%-0.2%) of the models achieves the same prediction ability depending on the sample (bootstrap or random weights).

When macroeconomic and financial variables are not included in the model and we start to add topic variables to a model with only country fixed effects, the classification ability of the 'fake' models is even more seldom statistically significantly higher relative to the baseline model and these models never achieve a higher prediction accuracy than the true prediction model. In addition, the differences in the fake models' AUC values relative to the original model's AUC value are around twice as large relative to the cases where baseline macroeconomic and financial variables are included in the model with topic variables.

To conclude, these results imply that adding enough uninformative variables to a prediction model that already includes macroeconomic and financial variables improve prediction ability significantly, but it is highly unlikely that the improvement is of similar magnitude that our original topic indicators achieve. When using only uninformative variables in the model together with country fixed effects, it is even rarer to beat the baseline model and the likelihood of reaching prediction accuracy equal to the true models is close to zero. As reaching a crisis classification ability of a model using topic frequency series is highly unlikely with uninformative series, the improved prediction ability of the former is highly unlikely to result from poor luck.

3.2.2 Pseudo out-of-sample predictive performance

In addition to the analysis in the previous subsection, we analyse possible overfitting issues by performing a pseudo out-of-sample prediction exercise where the models are re-estimated every year from 1950 onward with only observations that were available until that specific year, with given constructed text-based series. Then, each of these models is used to predict the state of the economy in the following year since 1950. This provides an additional perspective on the predictive power of the text topics similar to a forecasting situation where the full sample is not used to estimate parameters.

As expected and typically seen in different applications, the AUC values are not as high in this case as they are with the in-sample predictions. In any case, the topic-based models seemingly outperform the baseline model by 6 to 14 percentage points. This strengthens the previous in-sample results on the importance of the text-based topic series as predictors substantially.

3.2.3 Different crisis definitions and the impact of major crises

In the next set of robustness tests, we use the crisis definitions by Reinhart and Rogoff (2011) and Jordà et al. (2016a) (JST) instead of the definition by Baron et al. (2020). The timing or even the occurrence of a financial crisis can differ between these different datasets as crises are defined in varying ways and most commonly with narrative information.

Baron et al. (2020)'s crisis chronology used in the main analysis is the most extensive to date as it also uses information about episodes of large bank equity declines to find crises that may not have been associated with a panic at all. The authors also use all narrative information based earlier crisis chronologies to update their crisis chronology with the information from bank equity declines. As Baron et al. (2020) include new episodes and exclude episodes that they label "spurious" crises, the chronologies differ considerably. In the vast survey on the crisis literature as well as different crisis chronologies, Sufi and Taylor (2021) conclude that the use of equity decline data together with narrative crisis information is feasible.

The results of these different crisis definitions are combined in Table 3. In addition to just different definitions, we want to make sure that the results are not specific to any subset of crises due to different ways of defining these events. Columns (1) and (2) Panel A of Table 3 show that this is clearly not the case, as the results remain seemingly the same as in the main analysis in Section 3.1. Only with the JST crisis definition and the model including only five topic frequency indicators can we not reject the null that the classification accuracy is the same as in the baseline model that includes only the usual macroeconomic and financial indicators. However, in this specification we have almost close to 20% fewer crisis episodes in the dataset, potentially affecting the testing results.

Another important point about view to think of is that the attention to different topics might be capturing information related to only crises that were related to a panic? This would imply that the attention paid to these topics affects public beliefs directly or indirectly through factors that are discussed in some time period and finally tip the scale toward a state where a panic starts easily. In Panel A (column (3)), we repeat the analysis so that crises with panics are excluded from the dataset. Again, the results are the same, except the results of the model with only five topics. These results confirm that the results are not specific to only crises with panics, but include panicless episodes.

Another robustness perspective is based on the fact that crises are rare events as they occur, on average, every 28 years according to Jordà et al. (2010). Furthermore, around one half of the crisis episodes occur in clusters at the same time in different countries. Examples of these clusters include the global financial crisis episodes in the 1930s and in 2008. Therefore, we estimate the models again so that we exclude these crisis years and the surrounding years to make sure that the results are not driven by these specific two worldwide episodes. The results shown in Panel B in Table 3 (columns (1) and (2)) confirm that this is not the case, as the AUC values, their differences, and the corresponding p-values are very similar to the cases above including these crisis periods. These results imply that the topic-based leading indicators really capture some general factors that are not specific only to these major episodes.

Table 3: Robustness analyses: Different crises definitions, crisis types, impact of major crises, and out-of-sample predictive performance.

Panel A	Reinhart-Rogoff			Jorda-Schularick-Taylor			Excl. panics		
	AUC	p	p	AUC	p	p	AUC	p	p
Country FE only	0.591			0.616			0.616		
Macro/financial	0.673			0.741			0.893		
Macro/financial + top 5 topics	0.778	0.000		0.806	0.002		0.931	0.022	
Macro/financial + top 10 topics	0.790	0.000	0.097	0.819	0.001	0.039	0.973	0.002	0.006
Macro/financial + best subset topics	0.816	0.000	0.008	0.859	0.000	0.002	0.990	0.001	0.020
Top 5 topics	0.732	0.025		0.752	0.356		0.901	0.384	
Top 10 topics	0.789	0.000	0.000	0.802	0.023	0.002	0.963	0.016	0.008
Best subset topics	0.827	0.000	0.005	0.869	0.000	0.000	0.995	0.001	0.018
Crises		77			61			13	
Obs		1680			1751			1587	

Panel B	Excl. GFC			Excl. Great Depression			Out-of-sample		
	AUC	p	p	AUC	p	p	AUC	p	p
Country FE only	0.658			0.661			0.520		
Macro/financial	0.757			0.762			0.622		
Macro/financial + top 5 topics	0.825	0.000		0.822	0.001		0.748	0.000	
Macro/financial + top 10 topics	0.849	0.000	0.007	0.842	0.000	0.032	0.759	0.000	0.176
Macro/financial + best subset topics	0.886	0.000	0.002	0.875	0.000	0.007	0.760	0.000	0.460
Top 5 topics	0.797	0.057		0.790	0.119		0.679	0.068	
Top 10 topics	0.824	0.005	0.035	0.830	0.006	0.014	0.735	0.002	0.003
Best subset topics	0.894	0.000	0.000	0.879	0.000	0.001	0.733	0.005	0.525
Crises		55			65			36	
Obs		1550			1632			1086	

Notes: *Reinhart-Rogoff*: We use the crisis definition by Reinhart and Rogoff (2011). *Jorda-Schularick-Taylor*: We use the crisis definition by Jordà et al. (2016a). *Excl. panics*: We exclude the crisis episodes that co-occurred with a panic. Information on panics taken from Baron et al. (2020). *Excl. GFC*: We exclude the global financial crisis from the analysis by including only years before 2005. *Excl. Great Depression*: We exclude the years 1929–1936 from the analysis. *House prices*: We include the lagged three-year moving average of the change in the house price index as a potential crisis indicator. *Out-of-sample*: We estimate the models each year from 1950 onward by using observations that were available only up to that specific year. Each model is used to predict the following year state of the economy out of sample.

3.2.4 Additional macroeconomic and financial control variables

Next, we include new potential common or promising additional macroeconomic and financial indicators in the baseline model in addition to the variables that were included in the main analysis. We include the changes in house prices¹⁰ and bank balance sheet variables (bank capital, loans to deposits, and non-core funding ratios) used in Jordà et al. (2020). In one specification, we also replace the baseline model with the so-called red zones for credit and asset price growth in business and the household sector by Greenwood et al. (2020)¹¹. The results for these robustness perspectives are reported and compiled in Table 4. Columns (1)–(3) of Panel A in Table 4 show that the classification ability of the models that include these additional indicators in the baseline model are very similar to the ones in the main analysis. The differences in the AUCs are all statistically significant (except between the model with the top five topics and the baseline model) at the 5% significance level, favoring the use of topics as predictors.

3.2.5 Excluding some text information

In Panel B of Table 4, we also report the results of three model specifications where we estimate the topic frequencies in alternative ways. It is possible that the news article titles at the very end of the year might hold very specific information about an unfolding crisis that has been officially dated in the crisis chronology for the next year although the event was already sort of common knowledge at the end of the previous year. To assess that our results are not driven by this kind of endogeneity issue, we use annual topic frequencies that have been composed of each year’s news titles that were not published in December. We also take this even further and exclude news article titles that were published in the last quarter. In both cases, the major results hold.

Next, we exclude all country-year observations for which there were fewer than 10 article titles. This way, we can check that the topic frequencies are useful when they are composed of a large number of article titles and when there are only a few article titles responsible for the change in topic frequency. The results show that the improvements in prediction accuracy are still significant in all model specifications with topic frequencies when observations with fewer than 10 articles titles are removed¹². The absolute values of the AUC statistics are several percentage points smaller than in the main analysis, but they are still significantly larger for topic-based models relative to the baseline model.

¹⁰More specifically, we use the change in the house price index taken from the Jordà et al. (2016a) dataset and include it in the model also as a lagged three-year moving average.

¹¹According to Greenwood et al. (2020), the periods when the 3-year growth of house prices and household credit to GDP are in the top quintile and tercile, the household sector credit market is in an overheating state, and this predicts crises. Similarly, the business sector credit markets are in an overheating state or R-zone, when the three-year growth of credit to businesses to GDP and stock prices are in the top quintile and tercile. Greenwood et al. (2020) show that if a country experiences overheating in both sectors, a crisis is much more likely in the following years. We replicate these variables as closely as we can and include indicator variables for both R-zones and also their interaction term as the baseline model.

¹²This excludes more than 20% of the data.

Table 4: Robustness analyses: Additional leading indicator variables and excluding parts of text information.

Panel A	R-zone			House			Balance sheet		
	AUC	p	p	AUC	p	p	AUC	p	p
Country FE only	0.646			0.635			0.639		
Macro/financial	0.761			0.742			0.768		
Macro/financial + top 5 topics	0.857	0.000		0.812	0.001		0.829	0.000	
Macro/financial + top 10 topics	0.883	0.000	0.011	0.847	0.000	0.002	0.854	0.000	0.004
Macro/financial + best subset topics	0.916	0.000	0.012	0.873	0.000	0.021	0.876	0.000	0.004
Top 5 topics	0.823	0.070		0.776	0.092		0.789	0.178	
Top 10 topics	0.868	0.009	0.015	0.823	0.002	0.017	0.824	0.012	0.016
Best subset topics	0.897	0.001	0.025	0.872	0.000	0.004	0.874	0.000	0.000
Crises		36			55			68	
Obs		949			1408			1608	
Panel B	Excl. Dec. news			Excl. Q4 news			Min. 10 news per year		
	AUC	p	p	AUC	p	p	AUC	p	p
Country FE only	0.638			0.638			0.650		
Macro/financial	0.740			0.739			0.732		
Macro/financial + top 5 topics	0.809	0.000		0.804	0.000		0.802	0.001	
Macro/financial + top 10 topics	0.833	0.000	0.002	0.825	0.000	0.004	0.823	0.000	0.025
Macro/financial + best subset topics	0.857	0.000	0.004	0.852	0.000	0.001	0.866	0.000	0.001
Top 5 topics	0.778	0.050		0.772	0.081		0.773	0.045	
Top 10 topics	0.810	0.003	0.015	0.808	0.005	0.011	0.800	0.007	0.042
Best subset topics	0.856	0.000	0.000	0.852	0.000	0.002	0.865	0.000	0.000
Crises		75			74			58	
Obs		1748			1304			x	

Notes: *R-zone*: We include the so-called red zones of credit growth and asset prices by Greenwood et al. (2020). *Min. 10 news per year*: We include only country-years for which there were at least 10 news articles. *Excl. Q4 news*: We form the annual topic frequency series without news from the last quarter of each year. *Excl. Dec. news*: We form the annual topic frequency series without news from December.

3.2.6 Newspaper sentiment vs. topic attention

A possible concern related to the prediction process is that the topic frequencies might only capture the general economic sentiment contained in the economic news article titles and that the improved crisis prediction performance is mostly due to adding this sentiment measure (indirectly measured) to the prediction models. This would imply that the variables measuring the attention to specific general narratives (topics) are not likely to be important for crisis prediction performance. Instead adding a news title sentiment to a prediction model would be sufficient to improve the crisis prediction ability.

We test this possibility by re-estimating the prediction models so that the specification that includes macroeconomic and financial variables also includes a measure of average newspaper title sentiment during that year for titles mentioning a specific country. As a sentiment measure we exploit a commonly used procedure that utilizes the number of positive and negative words to form a polarity score for a title. It is common to only count the positive and negative words of a text and then subtract the sum of the latter from the sum of the former to comprise a very simple measure of sentiment.¹³ However, this kind of a naive way of calculating the sentiment score does not take into account so called valence shifter or amplifying/de-amplifying words. A previous word or group of words can switch the sentiment of a word to the opposite direction. For example, the word *good* on its own is clearly a positive word, but if it is preceded by the word *not*, then its sentiment turns to negative. Amplifying words like the word *very* increase the sentiment of a word. Again, the word *good* on its own has a positive sentiment, but if it is preceded by the word *very*, then it becomes even more positive. We use the R package `qdap` Rinker (2020) to calculate a sentiment score that takes into account these factors. The algorithm counts the polarity of a single word by taking account the context e.g. surrounding words and then adds all individual word polarity scores together and divides this sum with the square root of the total number of words in the newspaper title. This results into an unbounded polarity score where more negative values correspond to more negative sentiment for a newspaper title and vice versa for positive polarity score values.

The results in Table 5 reveal that adding this sentiment score as a possible crisis indicator variable to the baseline model does not improve crisis prediction ability of a model as the AUC value is basically identical to the baseline model with only macroeconomic and financial variables. In addition, a model with topic frequencies has statistically higher prediction power over these models as the AUC values are significantly higher after adding these to the baseline model or when they are used without a sentiment measure and macroeconomic - and financial variables. These results hold in both cases where the sentiment measure is included in levels or in differences (lagged 3-year moving averages). These results seem to confirm the finding that the improved prediction ability is driven by shifts in attention to specific topics, or narratives, and not just by the changes in the sentiment of the newspaper titles.

Overall, this extensive set of robustness tests confirms that the topic-based prediction models significantly outperform the baseline model that contains only the usual macroeconomic and

¹³A normalisation is often performed by dividing this measure with the total number of words in the text.

Table 5: Robustness analyses: Newspaper title sentiment and topic frequencies.

Panel A	Robustness analyses: Control for newspaper title average sentiment				
	AUC	Diff.to.baseline	p	Diff.upper	p
Macro/financial	0.742				
Macro/financial + sentiment	0.747			0.005	0.260
Macro/financial + sentiment + top 5 topics	0.812	0.065	0.000		
Macro/financial + sentiment + top 10 topics	0.843	0.096	0.000	0.031	0.003
Macro/financial + sentiment + best subset topics	0.873	0.126	0.000	0.030	0.001
Top 5 topics	0.778	0.031	0.089		
Top 10 topics	0.815	0.068	0.004	0.038	0.010
Best subset topics	0.873	0.126	0.000	0.058	0.000

Panel B	Control for change in newspaper title average sentiment				
	AUC	Diff.to.baseline	p	Diff.upper	p
Macro/financial	0.740				
Macro/financial + change in sentiment	0.740			0.000	0.638
Macro/financial + change in sentiment + top 5 topics	0.809	0.070	0.000		
Macro/financial + change in sentiment + top 10 topics	0.836	0.096	0.000	0.026	0.002
Macro/financial + change in sentiment + best subset topics	0.868	0.128	0.000	0.032	0.001
Top 5 topics	0.776	0.036	0.056		
Top 10 topics	0.814	0.074	0.002	0.038	0.009
Best subset topics	0.868	0.129	0.000	0.055	0.000

Notes: The annual average sentiment of newspaper titles the mention a specific country and the change in sentiment from previous year are included as lagged 3-year moving averages.

financial leading indicators. Although the prediction accuracy of the model with only the top five topic frequency indicators does not improve statistically significantly in some model specifications, the overall evidence suggests that the attention to specific topics in newspaper article titles has very relevant information about the probability of a crisis in the near future. In the next section, we take a deeper look into the interpretation of these topics and how they might contribute to different crisis episodes.

4 Text-based leading indicators

As the text-based newspaper article topic frequencies have been shown to add predictive power to financial crisis prediction models very substantially, it is important to understand, or at least hypothesize, what these "economic" linkages between the variables and financial crises might be. To examine this, we first look at the labels and the signs of the estimated coefficients of the topics that were selected to predict the financial crises and improve the classification ability of the baseline model. Next, we compare the selected topic labels to events and developments that have been shown to relate to or cause specific historical crisis episodes. We do this to see whether the new text-based topic labels and their content can be matched with the discussed causes of the historical crises.

4.1 Interpreting estimated coefficients

Table 6 shows all of the topic-based leading indicators selected by the forward stepwise selection procedure. In Panel A, it can be seen that all of the topic-based leading indicators (with macroeconomic and financial variables, i.e., the best subset of 22 topics) have labels that can be

seen to relate to issues that are relevant to the (macro)economy and financial stability. According to our models, the relative increase in the texts with titles discussing *commodities, changes in international trade barriers, political proposal lifecycles, foreign investment, promises, diplomatic relations, firm-level sales and purchases, matches and fights, high-level and government economists, plans and ideas, noble metals, investigations and charges, periodic production numbers, and jobs* increase the probability of a crisis in the future. All of these topics can be seen to include factors that might break the camel's back and shift public opinion about the path of the macroeconomy or later decrease trust in the resilience of the financial system. This might sow the seeds of a banking panic in the near future. On the other hand, as the model also predicts crises without panics, these topics can be seen to have relevant, more detailed, and surprising information about economic activity that could relate to the current and near future solvency of the banking system. For example, negative surprises in production, commodity markets, foreign investments, and firm sales can clearly cause economic turmoil. Sudden shifts in diplomatic relations, laws, restrictions, and agreements can have the same effect.

The titles that were the most representative for the topic we labeled as *promises* were mostly about situation where major, often government- or international-level promises were either made or broken by countries, firms, or highly relevant people for both. This might also indicate large changes in the economic activity. The *high-level and government economists* topic included numerous newspaper article titles stating or speculating about the appointments of economists to high-level government positions and the possible effects on future economic policies. The *matches and fights* topic represented most of the titles talking about issues like war, confrontation, and tensions between countries.¹⁴ It is clear that financial markets and the economy react to changes in international trade barriers, foreign investments, or shifts in politics. It is interesting that even when the selected model includes the equity capital gain and real GDP growth as macroeconomic and financial variables, these different topic series still add information to the model. This result might imply that more detailed information on financial markets and production than just the numerical (statistical) information on annual real GDP growth and the increase in stock prices is needed to assess the likelihood of a crisis.

The topics that have a negative coefficient such as *international trade restrictions (exports or imports), industrial production, relief and help, money markets, crises and dangers, economic policy agreements and cooperation, national supply shortages, factory problems, analysis and outlook, commodity producers, positive quarterly corporate reports, reports on business conditions, marine ships and other big purchases, seasonal reporting and bookkeeping, and factory problems* are in some cases a bit harder to interpret. The decrease (increase) in the relative frequency of these topics in newspaper article titles increases (decreases) the probability of a crisis in the future. Thus, these topics are discussed in the news during normal times, but when they are more absent, we are more likely to be in a "pre-crisis" period.

Importantly, and showing the robustness of some completely different additional predictive information over the conventional leading indicators, the selected model without the macroe-

¹⁴This topic also included some titles about major sports events between countries.

Table 6: Text-based leading indicator topics

Panel A: Model with macro/financial and the best subset of 22 topics

Topic	Label	Coefficient	Significance	Most common and exclusive words
30	International trade restrictions - Imports	-0.725	**	curbs-import-restrictions-barriers-retaliation-quota-eased
31	International trade restrictions - Exports	-0.778	***	drive-export-prepares-uranium-doubtful-world-wide-approved
92	Commodities	0.623	**	product-domestic-gross-butter-enlarged-dairy-type
3	Industrial production	-0.582	***	output-production-industrial-speed-newsprint-aluminum-copper
83	Change in international trade barriers	0.356	*	duty-materials-items-cloth-bans-autos-wool
97	Relief and help	-1.288	**	relief-busy-unemployed-fund-cross-idle-glad
23	Money markets	-0.328	*	money-current-call-events-tight-developments-speculation
11	Political proposal lifecycle	1.3	***	proposals-representatives-resistance-pledges-restored-freedom-cease
99	Crises and dangers	-0.496	*	coal-fuel-acute-nears-miners-middle-east
88	Foreign investment	0.646	***	like-potential-climate-investment-foreseen-shape-welcome
25	Economic policy agreements and cooperation	-0.215	**	economic-ties-problems-impact-policies-stability-summit
58	National supplies shortage	-1.15	*	needs-food-supply-requirements-adequate-assure-material
77	Promises	0.662	*	promise-given-failing-breaks-authorized-drawn-advice
82	Diplomatic relations	0.462	**	reply-grants-granted-powers-issued-encouragement-recognize
44	Firm level sales and purchases	0.329	***	importer-shut-firm-option-buys-sold-owner
96	Factory problems	-0.555	*	resources-factories-wealth-rich-natural-vast-development
101	Matches and fights	0.439		prospect-hill-sports-invasion-cheered-complications-dark
45	High level and government economists	0.193	*	economics-politics-sights-lesson-lessons-search-nobel
40	Analysis and outlook	-1.105		improved-avoided-alarm-tumble-hurting-term-response
104	Commodity producers	-1.429	*	producers-pleased-longer-tactics-importers-regret-dealers
21	Plans and ideas	0.512	*	idea-whose-forward-importing-stress-revived-accept
55	Noble metals	0.259		gold-metal-standard-silver-reserves-reserve-shipment

Panel B: Model with the best subset of 27 topics

Topic	Label	Coefficient	Significance	Most common and exclusive words
30	International trade restrictions - Imports	-0.705	**	curbs-import-restrictions-barriers-retaliation-quota-eased
31	International trade restrictions - Exports	-0.923	***	drive-export-prepares-uranium-doubtful-world-wide-approved
44	Firm level sales and purchases	0.432	***	importer-shut-firm-option-buys-sold-owner
92	Commodities	1.118	***	product-domestic-gross-butter-enlarged-dairy-type
67	Periodic production numbers	0.29	***	international-annual-journal-cent-manufacturing-awaits-climb
3	Industrial production	-0.423	**	output-production-industrial-speed-newsprint-aluminum-copper
83	Change in international trade barriers	0.416	*	duty-materials-items-cloth-bans-autos-wool
88	Foreign investment	0.82	***	like-potential-climate-investment-foreseen-shape-welcome
25	Economic policy agreements/cooperation	-0.354	**	economic-ties-problems-impact-policies-stability-summit
97	Relief and help	-1.26	**	relief-busy-unemployed-fund-cross-idle-glad
99	Crises and dangers	-0.43	*	coal-fuel-acute-nears-miners-middle-east
23	Money markets	-0.24		money-current-call-events-tight-developments-speculation
65	Positive quarterly corporate reports	-0.377		profits-earnings-corporate-dividend-reports-dividends-corporation
45	High level and government economists	0.273	*	economics-politics-sights-lesson-lessons-search-nobel
72	reports on business conditions	-0.33	*	looking-ways-throughout-world-talking-means-shaken
16	Investigations and charges	0.243	*	inquiry-trader-involves-investigating-charges-probe-bribery
82	Diplomatic relations	0.516	**	reply-grants-granted-powers-issued-encouragement-recognize
106	Marine ships and other big purchases	-0.434		ships-instead-running-size-potash-pictures-voted
101	Matches and combats	0.521		prospect-hill-sports-invasion-cheered-complications-dark
58	National supplies shortage	-1.324	*	needs-food-supply-requirements-adequate-assure-material
66	Jobs	0.305	*	jobs-lose-create-women-creates-employment-eliminate
55	Noble metals	0.33	*	gold-metal-standard-silver-reserves-reserve-shipment
104	Commodity producers	-1.538	*	producers-pleased-longer-tactics-importers-regret-dealers
11	Political proposal lifecycle	1.012	*	proposals-representatives-resistance-pledges-restored-freedom-cease
21	Plans and ideas	0.524	*	idea-whose-forward-importing-stress-revived-accept
2	Seasonal reporting and bookkeeping	-0.283		review-weekly-spring-reviews-books-features-failures
96	Factory problems	-0.367		resources-factories-wealth-rich-natural-vast-development

Notes: “Coefficient” refers to the estimated coefficient in the panel logit model (see Eq. (4)). “Significance” is based on the corresponding (robust) t-test statistics of Driscoll and Kraay (1998). The statistical significance is denoted in the usual way (referring to the p-values of the robust t-test statistics) by *p<0.1; **p<0.05; ***p<0.01.

conomic and financial leading indicators shows that the best subset model includes almost the same set of topic indicators as the one with the macroeconomic and financial variables. The *promises* and *analysis and outlook* topics are missing from the former model, but it includes five extra topics: *positive quarterly corporate reports*, *reports on business conditions*, *investigations and charges*, *marine ships and other big purchases*, and *seasonal reporting and bookkeeping*. The models with and without macroeconomic and financial leading indicators have very similar prediction abilities and the coefficients of the mutual text-based indicators have similar signs and magnitudes.

Table 7: Representative news titles for the top 5 leading topic indicators. The top 5 text-based leading indicators are from the model that uses the best subset of topic indicators in addition to the macroeconomic and financial indicators.

Topic 30 *International trade restrictions - Imports*

Year	Country	θ	Title
1934	ITA	0.86	<i>italy to raise import duties on some goods</i>
1975	USA, ESP	0.84	<i>u.s. imposes an embargo on spanish tuna import</i>
1981	USA	0.65	<i>u.s. given warning on import curbs: u.s. warned import curbs may trigger trade war</i>
1988	JPN	0.50	<i>hill study urges japan to drop import curbs: study says japan must end import curbs</i>
1991	USA, FRA	0.45	<i>request to import u.s.-made hondas creates new trade dilemma for france</i>

Topic 31 *International trade restrictions - Exports*

Year	Country	θ	Title
1987	USA	0.93	<i>u.s. lifts controls on export of drilling gear to soviets: u.s. permits export of drilling gear</i>
1996	JPN, USA	0.91	<i>japanese chips may scramble u.s. export ban: new japanese chips may scramble u.s. export regulation</i>
1988	SWE	0.40	<i>sweden about to tighten its laws on arms export</i>
1993	FRA	0.35	<i>french set for 'crisis' on trade: farmers press government to keep export subsidies</i>
1907	JPN, GBR	0.34	<i>japan stops coal export.: british naval officers returning from orient tell of embargo.</i>

Topic 92 *Commodities*

Year	Country	θ	Title
1920	AUS, USA, GBR, CAN	0.75	<i>enormous wool stocks deaden world market: australia, united states, great britain, canada and argentina surfeited with raw product, for which there is no demand—canadian co-operative wool growers' executive will issue interesting statement good type of clyde mare</i>
1930	CAN, DNK	0.56	<i>food animal export drops dangerously in past nine years: importance of foreign markets for canadian live stock sadly underestimated—denmark assuming control of british bacon demand, while dominion figures are reduced yearly—domestic packing houses market food products totalling \$300,000,000 yearly</i>
1927	USA	0.43	<i>u.s. uses 80 per cent of world auto output: improved product and decrease in cost is held largely responsible.</i>
2004	USA	0.40	<i>gross domestic product: revision shows u.s. economy grew faster in second quarter</i>
1981	CAN	0.32	<i>other cheese varieties are after canadian cheddar market slice</i>

Topic 3 *Industrial production*

Year	Country	θ	Title
1988	USA	0.89	<i>u.s. industrial production figures record biggest gain in 6 months</i>
1929	SWE	0.84	<i>sweden may increase newsprint production</i>
1929	BEL, USA	0.62	<i>shortage of workmen in belgium is laid to importers of american automobiles</i>
1990	JPN, USA	0.59	<i>japan could surpass the u.s. in output of electronics soon, new report says</i>
1927	DEU	0.55	<i>talk of a 'saturation point' in german industrial output</i>

Topic 83 *Change in international trade barriers*

Year	Country	θ	Title
1920	AUS, CAN	0.93	<i>higher duties on implements: australia tariff changes affect products supplied by canada want reciprocal trade</i>
1982	USA	0.89	<i>u.s. farm export aid considered: farm export aid considered</i>
1927	FRA	0.82	<i>volume of trade with france may be cut by higher duties</i>
1990	USA, JPN	0.81	<i>u.s. persuades allied-signal to end unfair trade complaint against japan</i>
1928	BEL, CHE, PRT	0.71	<i>report on changes in foreign tariffs: commerce agents send word on new rules and duties affecting exports. bleached flour banned belgium takes action—swiss regulate meat imports—portugal puts on surtaxes.</i>

Furthermore, we want to check in more detail what kind of an information these leading text based indicators actually include. We do this by looking at the most representative news article titles from the years close to the crisis beginning. The most representative news article titles for the text based indicators and their corresponding share of that particular topic (θ) are presented in Table 7. These are the 5 leading indicators from the model that uses the baseline macroeconomic and financial variables together with the best subset of topic frequency series selected via forward stepwise selection procedure. These news titles display the fact that the topic measure general narratives in news that include more detailed information on the usual crisis indicators and also economy related information that is not likely captured in macroeconomic and financial indicators. For example, the title *talk of a 'saturation point' in german industrial output* is a good example of the former case for the *industrial production topic* as it clearly has a more forward looking information as basic industrial production numbers. The title also describes the expectations and economic atmosphere related to future production. Similarly for the same topic the title *sweden may increase newsprint production* has a similar more forward looking information in it. Stock prices may capture these kind of news if markets see them as being relevant for future dividend flow. However, in the cases where markets are 'wrong' and detailed news titles as the ones described previously are relevant to crisis occurrence and hence future dividend flow, the topic frequencies are able to generalize and measure this information in an efficient manner. In the next section we will study how these topics might relate to actual historical crisis occurrences.

4.2 Crisis case studies

To understand these indicators better, we next present a number of case studies of historical financial crisis episodes for which the model using topic indicators gave a clearly higher crisis probability than the baseline model. Our hypothesis is that the topic-based leading indicators represent factors that have been discussed as possible causes of specific crisis episodes in general, but these "known" causes have not been identified generally before, or more specifically, used as indicators in empirical analyses. To interpret the contribution of different indicators to this increasing crisis probability, we use a bar chart to plot each crisis probability and its decomposition over different leading indicators. That is, we report the share of the j th indicator's index value $I_{crisis,j} = \beta_j X_{crisis,j}$ in that particular crisis (episode) for the the total index value $TI_{crisis,s} = \sum_{j=0}^J I_{crisis,j}$ of the J indicators whose index value has the same sign s in that year. With this notation, the crisis probability (cf. Eq. (4)) can be represented in the following way:

$$Pr(Y_{i,t} = 1|X_{i,t-1}) = \sigma(TI_+ - TI_-),$$

where $\sigma(\cdot)$ denotes a logistic function, as in Eq. (4). In each case study, we first roughly describe the historical context and possible crisis causes with a number of historical literature sources. Next, we identify topics that have increased financial crisis probability in the crisis year according to our prediction model. Finally, we discuss possible links between crisis explanations

given by historical sources and the topic indicators in our model.

4.2.1 Portugal in 1890

We start with a financial crisis case study from the late 19th century. The Portuguese financial crisis in 1890 is a clear example of an event where the topic-based model outperformed the baseline model in terms of crisis classification ability and the interpretability and informativeness of the given crisis probability from the perspective of a policy maker. From the upper panel of Figure 11, it can be seen that the baseline model gives basically a flat probability during the previous decade, the crisis year, and the following couple of years, whereas the crisis probability estimated by the topic-based model starts to rise two years before the crisis, and peaks during the event year, and in all other years stays flat.



Figure 11: Case: Portugal in 1890: crisis probability and indicator contributions.

The 1890 financial crisis in Portugal was mainly due to a massive accumulation of public debt. This led to a fragile state of the economy that resulted in a crisis in 1890, when the main Portuguese export (wine) decreased. In addition, a financial crisis in Brazil contributed

to these events as it decreased the remittances of Portuguese emigrants. These two factors were the main sources of public income. In addition, the financial crisis shook other Latin American countries and led to problems in international financial markets, and especially affected the Baring Brothers, the main financial intermediary of the state of Portugal. However, the roots of the fragile state of Portugal's public sector and the subsequent crisis were not just in the crises in Latin American countries. The ultimate reasons behind the crisis and the accumulated public debt were in the policies driven by the financial minister Mariano de Carvalho and the government of *Partido Regenerado*. De Carvalho and *Partido Regenerado* revolutionized the way public financing was conducted in Portugal by, for example, nationalizing the tobacco trade. However, at that time, optimism prevailed in the markets, which meant good creditworthiness and the pricing of Portuguese assets (Lains, 2021).

From the lower right panel of Figure 11, it can be seen that the main factors and topics that increased the crisis probability were *high-level and government economists* and *promises*. Actually, the crisis probability and the share of these two topics increased two years before the actual crisis, indicating that the Portuguese economy was probably already in a risky state. These two topics most likely capture the public discussion about the financial policies driven by the *Partido Regenerado*: The appointments and public statements of some high-level and government economists increased before the crisis and possibly the party or politicians had also made many public promises to enhance their political agenda, which ultimately led to the crisis. The figure also shows that the share of the topic *diplomatic relations* clearly increased during the crisis period and then remained at a higher level than before the crisis. The revolution in the Republic of Brazil might be the cause of the increase in the share of *diplomatic relations*-related news, which, however, was also one of the actual triggers of the crisis.

In addition, the three topics together are in line with the political and economic trajectory in Portugal during that time and possibly are also related to Portugal's role in the international financial markets, as the main sources of the Portuguese debt were the London and Paris markets. As a result of the crisis, Portugal abandoned the gold standard and defaulted partially as the interest payments for domestic and foreign debts were cut. Consequently, Portugal was banned from international lending markets until 1902.

The main topic that, before and after the crisis, *decreased* the crisis probability was *factories and resources*. The drop in the share of this topic and its crisis-decreasing effect might be due to the fact that wine production and exports fell drastically. Consequently, there was no comprehensive news about factory production, which in this case was a bad sign.

4.2.2 United States in 1930

After the end of World War I and the Spanish flu pandemic, the so-called Roaring Twenties was a period of very high economic growth. If the recession of 1920–1921 is excluded, then the compounded annual per capita output growth for the decade was as high as 3.5% (Harrison and Weder, 2009). Harrison and Weder (2009) argue that the main driver of this remarkable growth was notable growth in productivity. According to the authors, process innovation and adaptation

of electricity in manufacturing were some of the most important examples of factors contributing to the increased productivity. Friedman and Schwartz (1965) state that the deterioration in the quality of loans during this optimistic decade might have partly caused the banking crisis that occurred in October 1930, but that the two subsequent crises in 1930 and 1931 could not be linked to the same phenomenon.

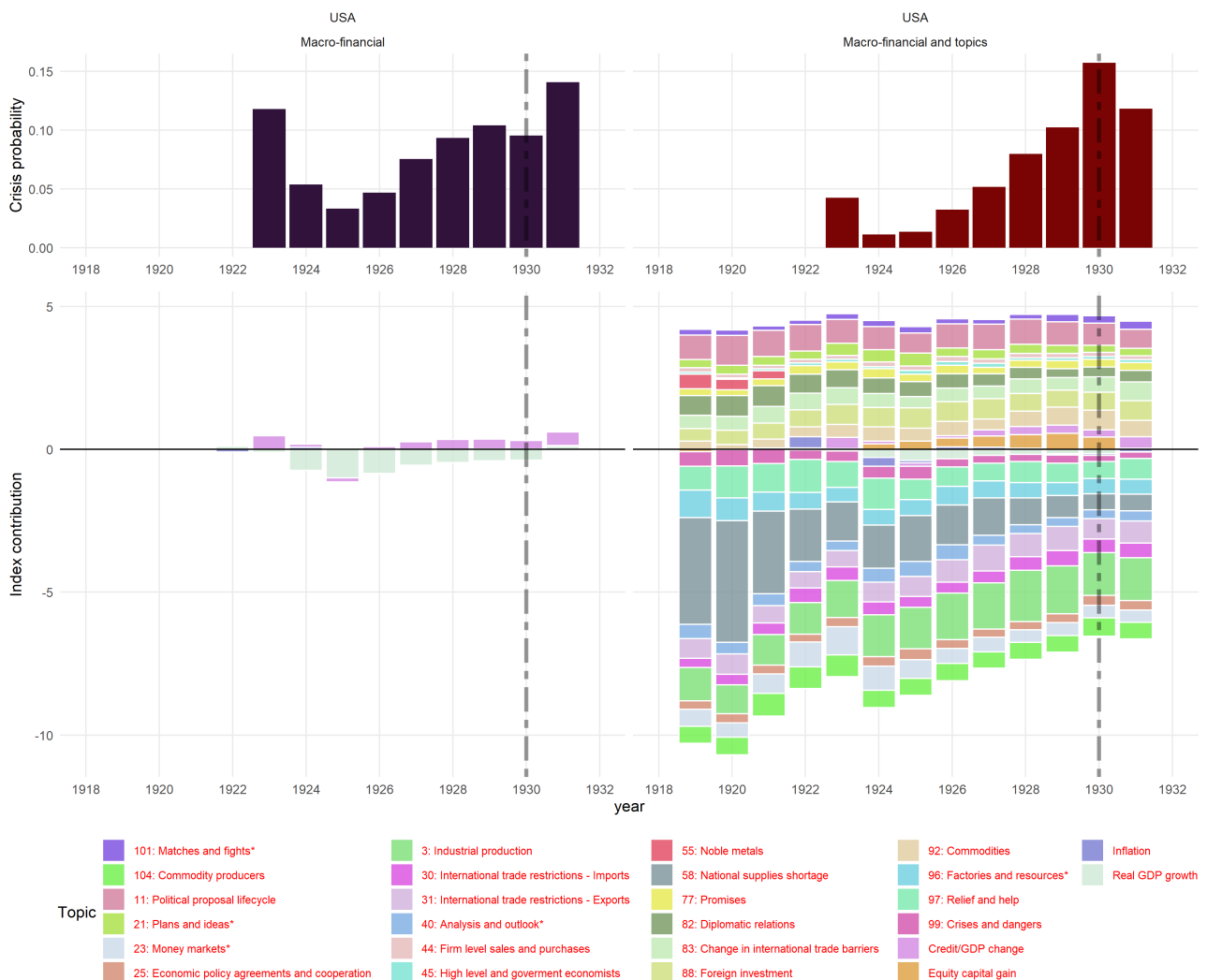


Figure 12: Case: United States in 1930: crisis probability and indicator contributions to index.

The better predictive performance of the topic-based model is again evident in the results in Figure 12. Although the macroeconomic and financial benchmark also implies an elevated crisis risk for the last years of the Roaring Twenties, it peaks a year too late in 1931 and in 1923. On the other hand, in the topic-based models crisis probability increases gradually toward the actual crisis year and peaks in 1930. When we look at the contribution of each indicator to the rising crisis probability, it can be seen that the increasing factor was mainly the increase in equity capital gain. However, the significant decrease in the national supplies shortage topic in newspapers seems to certainly contribute the most to the gradual increase in crisis probability. This might capture the euphoric general mindset and atmosphere of economic excess that most likely contributed to the increase in speculation in the financial markets and the deterioration

in loan quality.

4.2.3 Finland in 1931

In the 1920s, Finland took on a lot of foreign currency debt, which amounted to 25% of the Finnish GDP in the middle of the decade. Foreign bond loans nearly doubled from 1927 to 1930. According to Hjerpe et al. (1993), the balance of trade turned very negative in 1928, which was mostly driven by imports of house building materials and consumer goods (due to the failure of crops in the same year) at the peak of a boom phase. The capital Helsinki and some big banks took somewhat large long-term loans during the next two years, but in 1931, there was a sudden end to the supply of these loans.

During the 1920s, one third of the Finnish GDP came from primary production, mainly agriculture and forestry. The decline of the price level by 25% from mid-1925 to 1931 pushed the revenues of this sector smaller and smaller. This decrease together with the increasing indebtedness of farmers resulted in the compulsory auction of thousands of farms from 1928 to 1936. Through the 1920s until 1929, the volume of industrial production doubled, and sawmill production increased with loaned capital due to the absence of competition from the Soviet Union. The production volume of the industry dropped by almost half in 1928, when the Soviets returned to the market, and at the same time, the global demand and domestic house building declined. Together with the overheating of the housing market and its collapse during the same three-year window, these developments and problems of the real economy resulted in major loan losses for banks. A total of 11 banks went bankrupt (Hjerpe et al., 1993).

The upper panel of Figure 13 displays how the text-based model clearly outperforms the baseline model by giving a higher crisis probability for the crisis year of 1931 and for the two previous years. In contrast, for the latter model, the crisis probability is basically flat during the entire 12-year-window. The topics that contributed the most to the increase in the crisis probability for the crisis year were the *diplomatic relations*, *change in international trade barriers*, and *plans and ideas* topics. The second factor most likely captures the many changes that occurred in industrial production, such as the Soviet Union entering the sawmill industry and the collapse of global demand that led to the decrease in prices that also hurt other important sectors like agricultural production. The increase of this topic's frequency might also be related to the fact that Finland joined the gold standard in 1926 and abandoned it in 1931. At the end of this period, the Markka was significantly overvalued relative to the pound.

The increase in the *plans and ideas* topic in the newspapers three years before the crisis is expected to capture factors related to the overheating of the housing market and the surge of foreign funding into the country. On the right side of the bar chart of the topic-based model, it can be seen that the frequencies of the *money markets* and *national supply shortages* topics were smaller for the three years before the crisis, on average, relative to earlier years. This implies that the lower frequencies of these topics were not dampening the crisis probability in the same way as they did before. The quieting down of money market news might be related

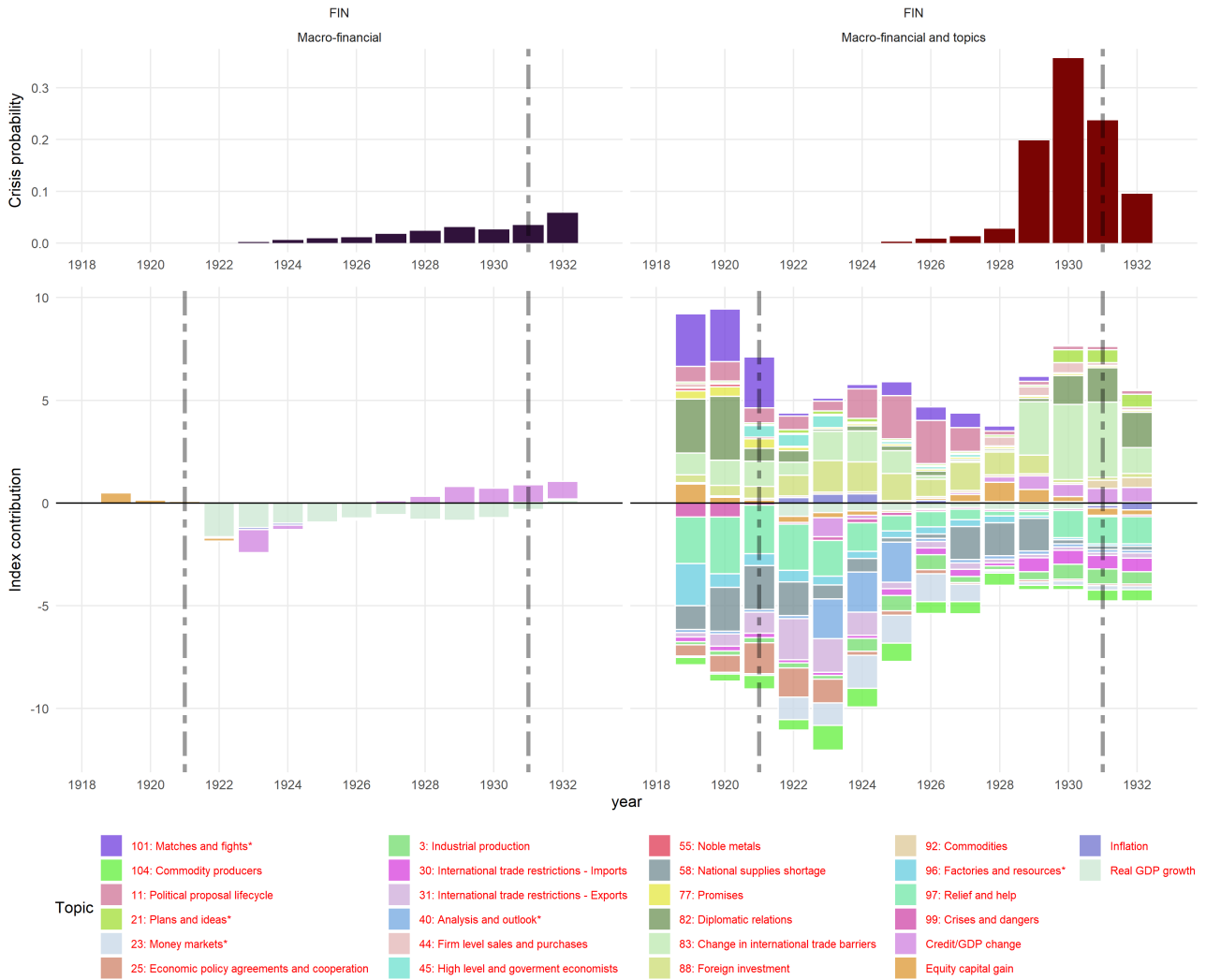


Figure 13: Case: Finland in 1931: crisis probability and indicator contributions to index.

to the fact that the Finnish government did not take on any debt from 1928 to 1930, whereas in the first half of the decade, the government took out quite a lot of foreign loans to fund its budget. Finally, the absence of the *national supply shortages* topic might relate to the deficit of the trade balance and the vast importing of consumer goods and building materials during the end of this boom phase.

4.2.4 Finland in 1990

After 1931, the next major financial crisis in Finland occurred six decades later in 1990. Financial liberalization in the early 1980s started a large inflow of foreign capital to the country. This resulted in credit expansion and then increased spending, investments, and asset prices. Due to the loss of competitiveness and the central bank defending the pegged exchange rate, interest rates increased significantly in 1990. The high rates were unbearable for many households and firms as they had become highly indebted during the boom. In addition, trade between the Soviet Union and Finland collapsed when the former fell apart. Finally, as the economic output declined, and asset prices collapsed, a large number of bankruptcies occurred, and loan losses

began to accumulate for banks. The government ended up taking control of the Savings Bank Group, and other banks were recapitalized (Kiander and Vartia, 2011).

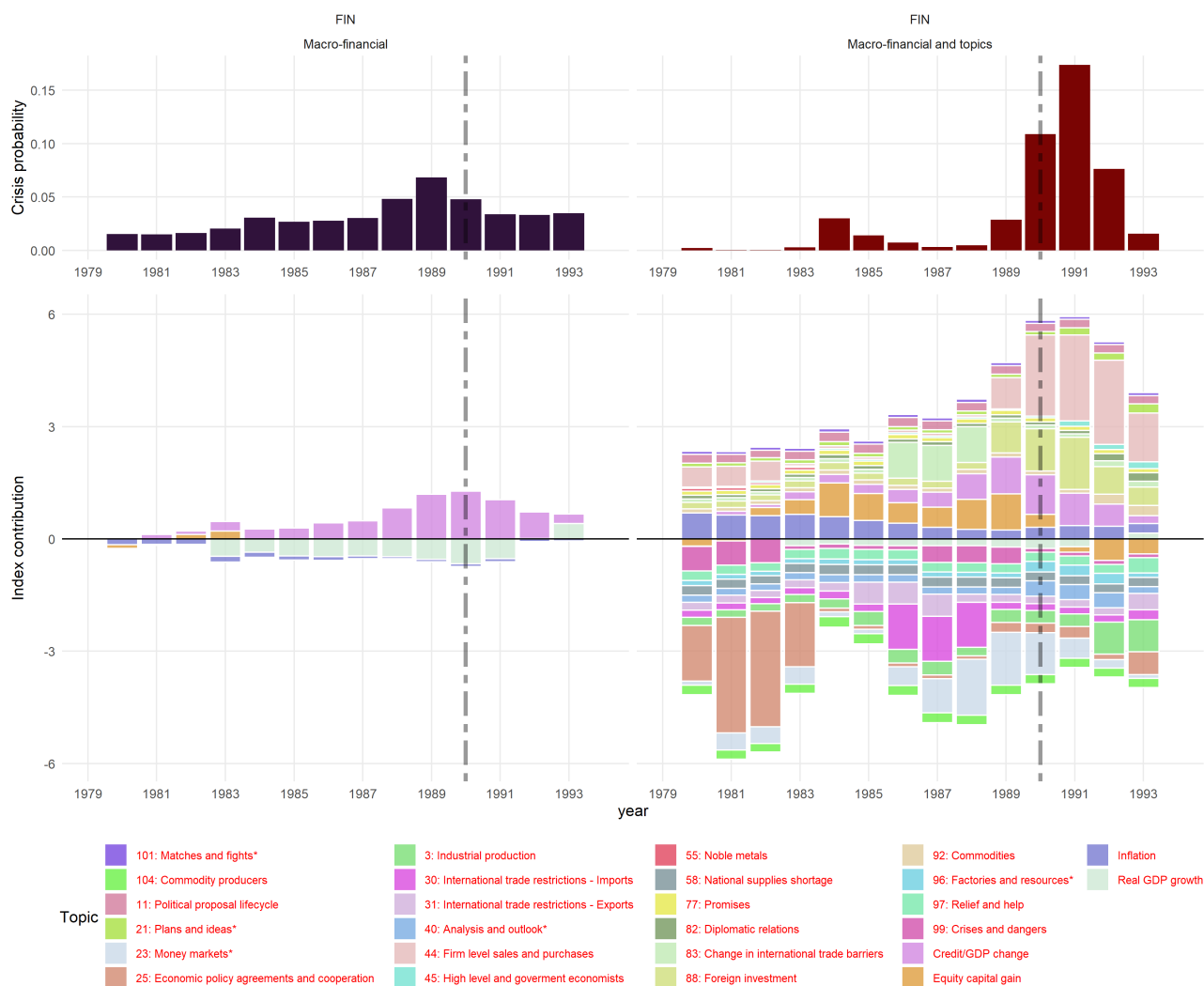


Figure 14: Case: Finland in 1990: crisis probability and indicator contributions to index.

Figure 14 shows that the baseline model gives an elevated crisis probability for the crisis year and several preceding years, but for the topic-based model, the increase in crisis probability is much larger at the end of the 1980s. The increase of *firm-level sales and purchases* and *foreign investments* topics and the change in credit-to-GDP are the largest contributors to the increase in crisis probability for that year. In addition, relative to previous years, the *crises and dangers*, *industrial production*, and *international trade restrictions on imports* topics were less frequent in the newspaper article titles, which also increased crisis probability for this year. These factors likely capture the collapse of trade with the Soviet Union and problems in the corporate sector that ultimately partly caused the banks to have massive amounts of defaulted loans.

4.2.5 Japan in 1990

After remarkably fast economic growth between 1953 and 1973, Japan was becoming the US's biggest rival as the biggest economic superpower. Japan had changed from an agricultural country into a leader in steel and car production. Later, the country took over the video recorder and other high electronic industries as well. Many different explanations, such as superior education levels and high savings rates, have been provided as explanations for this never-before-seen economic growth. In addition, the strategic policies of Japan's government, the Ministry of International Trade and Industry (METI), and the Ministry of Finance together with the so-called *keiretsu* way of forming company networks have been argued to be some of the key factors responsible for the country's economic success. The fast growth increased stock and real estate prices to unforeseen levels in the late 1980s. This increase was accompanied by high levels of speculative activity, and the bubble collapsed soon after the Bank of Japan started increasing the interest rate in 1990 (Krugman, 2009).

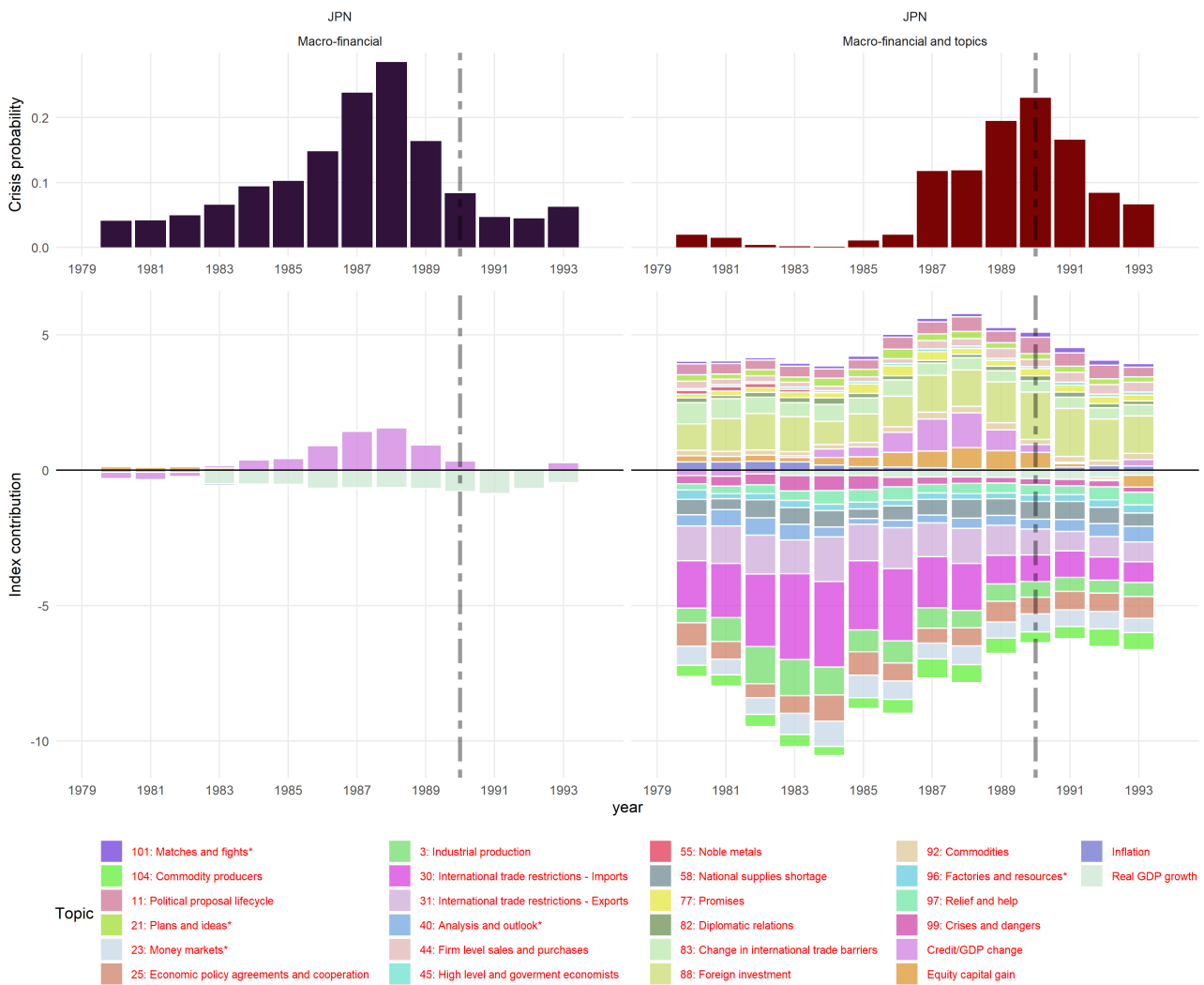


Figure 15: Case: Japan in 1990: crisis probability and indicator contributions to index.

In Japan's case, a purely macroeconomic and financial indicators-based prediction model

would have warned of a heightened financial crisis risk as can be seen from the upper-left corner of Figure 15. This increase was driven mainly by the increased credit-to-GDP ratios, but the crisis probability seems to peak two years too early, and then the model signals a much less riskier state of the economy for the actual year of the crisis. However, the crisis probability from the prediction model utilizing information about newspaper topic frequencies also rises several years before the crisis and actually peaks during the crisis year. The credit-to-GDP ratio and rising equity capital gains contribute to the early rise of the crisis probability in this case, but the biggest contributor to the crisis probability during the peak year relative to the preceding years is the decreased frequency or the absence of the *international trade restrictions on imports* topic.

The development of the frequency of this topic in the newspapers might capture the U.S.-Japan trade conflicts of the 1980s. The US accused Japan of keeping the yen undervalued relative to the dollar, running a persistent trade surplus with the US, and using unique Japanese organizations (*keirutz*) that, according to competing U.S. exporting firms, create trade barriers and unfair competition. The so-called Plaza Accord was signed in 1985 to depreciate the value of the dollar relative to the currencies of France, West Germany, the United Kingdom and Japan. This agreement led to significant appreciation of the yen with respect to the dollar, which later negatively affected the competitiveness of Japanese exporters. The Bank of Japan reacted to the appreciation of yen by raising interest rates, which has been argued to be the trigger of the collapse of asset prices (Fung et al., 2020).

4.3 Discussion

The selected leading indicators and the crisis case studies in the previous sections suggest that the topic series capture factors whose frequency in the news titles provide useful information about the risk of a financial crisis. As hypothesized earlier, many of the topics seem to relate content wise and by the way they evolve over time with phenomena and factors that have been discussed in public and in individual studies to have caused or contributed at least partly to a specific historical crisis episode. Examples of these factors are important changes in economic policies (*Promises* and *high level government economists* - Portugal 1890; *money markets*- Finland 1931; *international trade restrictions* - Japan 1991), external shocks (*diplomatic relations* - Portugal 1890; *changes in international trade barriers*- Finland 1931; *international trade restrictions on imports* and *foreign investments* - Finland 1990), detailed information on production or output (*factories and resources* - Portugal 1890; *firm level sales and purchases* and *industrial production* - Finland 1990), and economic atmosphere (*national supplies shortages* - United States 1930; *plans and ideas*- Finland 1931).

Although it is not straightforward to use newspaper article titles as real-time early warning indicators, it is still noteworthy that the financial markets and the macroeconomic key figures do not reflect all the relevant public information that is highly relevant for the risk of a financial crisis. These results bring a new, more detailed perspective on the view that crises are all the same but still different as we find many general leading indicators implying that the "stories"

of crises are still more complicated than just a rise in credit and asset prices before a bust. These results confirm that these different stories or paths to crisis can be generalized in more detail than before in the previous research with the newspaper title topic series. As a large portion of the most important topic based crisis indicators are highly related to international trade, it would be important to investigate these factors and their possible relationship to crisis in more detail in future research.

5 Conclusions

We utilize a topic model to identify different economy-related topics found in a group of global newspapers in 17 developed countries for the past 150 years. The topics of the final model are coherent and understandable, and their time series frequencies capture historical events that can be connected to the specific economic topics. When this text-based information is included with the usual financial and macroeconomic crisis indicators, the best predictive model that uses information from the 22 most relevant different topics in addition to the macroeconomic and financial variables outperforms the baseline model with only the latter indicator group statistically significantly. A prediction model with only the five most useful topic series still clearly outperforms the baseline model. In addition, different topic-based models with 5 to 27 different topics outperform the baseline model even when the usual macroeconomic and financial indicators are excluded from the former model. These results imply that text data, and more specifically, the topics of the newspaper article titles, contain highly relevant information about the risk of future financial crisis in addition to the usual macroeconomic and financial indicators.

The actual text-based leading indicators include topics that capture, for example, changes in economic policies, external shocks, detailed information on production or output, and the economic atmosphere. A number of the examined crisis case studies show that the topics can be used to identify (generalizable) events that precede a crisis in more detail than by just using the usual macroeconomic and financial variables, such as credit and asset prices. The topics capture factors that have often been discussed as causes of specific crisis episodes, but which have not been used to in empirical models. A large number of presented robustness tests confirm these main results.

References

- Airoldi, E. M. and J. M. Bischof (2016). Improving and evaluating topic models and other models of text. *Journal of the American Statistical Association* 111(516), 1381–1403.
- Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics* 131(4), 1593–1636.
- Baron, M., E. Verner, and W. Xiong (2020). Banking crises without panics. *The Quarterly Journal of Economics* 136(1), 51–113.

- Berge, T. and O. Jordà (2011). Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics* 3(2), 246–77.
- Blei, D. M., A. Y. Ng, and M. J. Jordan (2003, March). Latent dirichlet allocation. *The Journal of Machine Learning Research* 3, 993–1022.
- Bybee, L., B. Kelly, and Y. Su (2021, July). Narrative asset pricing: Interpretable systematic risk factors from news text. Working Paper 21-09.
- Bybee, L., B. T. Kelly, A. Manela, and D. Xiu (2020, Jan). The Structure of Economic News. NBER Working Papers 26648, National Bureau of Economic Research, Inc.
- DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44, 837–845.
- Driscoll, J. C. and A. C. Kraay (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* 80(4), 549–560.
- Friedman, M. and A. Schwartz (1965). *The Great Contraction, 1929–33*. National Bureau of Economic Research, Inc.
- Fung, K. C., N. Aminian, X. Fu, J. H. Rhee, and C. Y. Tung (2020, 01). "silicon trade policies" and a comparative analysis of china-u.s. and japan-u.s. trade conflicts. *Journal of Chinese Economic and Business Studies* 18(1), 29–50.
- Greenwood, R., S. G. Hanson, A. Shleifer, and J. A. Sørensen (2020, Jun). Predictable Financial Crises. NBER Working Papers 27396, National Bureau of Economic Research, Inc.
- Harrison, S. and M. Weder (2009, 09). Technological change and the roaring twenties: A neoclassical perspective. *Journal of Macroeconomics* 31(3), 363–375.
- Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun (2019). Firm-Level Political Risk: Measurement and Effects. *The Quarterly Journal of Economics* 134(4), 2135–2202.
- Hjerppe, R., V. Ikonen, and P. Valkama (1993). 1930-luvun lama ja suomen velkaantumien. *The Finnish Economic Journal* 89(1), 5–21.
- Jordà, O., B. Richter, M. Schularick, and A. M. Taylor (2020, 08). Bank capital redux: Solvency, liquidity, and crisis. *The Review of Economic Studies* 88(1), 260–286.
- Jordà, O., M. Schularick, and A. M. Taylor (2010, December). Financial crises, credit booms, and external imbalances: 140 years of lessons. Working Paper 16567, National Bureau of Economic Research.
- Jordà, O., M. Schularick, and A. M. Taylor (2016a, April). Macrofinancial History and the New Business Cycle Facts. In *NBER Macroeconomics Annual 2016, Volume 31*, NBER Chapters, pp. 213–263. National Bureau of Economic Research, Inc.

- Jordà, O., M. Schularick, and A. M. Taylor (2016b). Sovereigns Versus Banks: Credit, Crises, and Consequences. *Journal of the European Economic Association* 14(1), 45–79.
- Jordà, O., M. Schularick, and A. M. Taylor (2017). Macrofinancial history and the new business cycle facts. *NBER Macroeconomics Annual* 31(1), 213–263.
- Kiander, J. and P. Vartia (2011, 02). Lessons from the crisis in finland and sweden in the 1990s. *Empirica* 38(1), 53–69.
- Krugman, P. (2009). *The return of depression economics and the crisis of 2008*. W. W. Norton and Company.
- Lains, P. (2021). *A History of Public Banking in Portugal in the 19th and 20th Centuries: The Caixa Geral de Depósitos*. Routledge.
- Larsen, V. H., L. A. Thorsrud, and J. Zhulanova (2021). News-driven inflation expectations and information rigidities. *Journal of Monetary Economics* 117, 507–520.
- Mimno, D., H. M. Wallach, E. Talley, M. Leenders, and A. McCallum (2011). Optimizing semantic coherence in topic models. EMNLP '11, USA, pp. 262–272. Association for Computational Linguistics.
- Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association* 89(428), 1303–1313.
- Reinhart, C. M. and K. S. Rogoff (2011, August). From financial crash to debt crisis. *American Economic Review* 101(5), 1676–1706.
- Rinker, T. W. (2020). *qdap: Quantitative Discourse Analysis Package*. Buffalo, New York. 2.4.2.
- Schularick, M. and A. M. Taylor (2012, April). Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008. *American Economic Review* 102(2), 1029–1061.
- Shiller, R. J. (2017, April). Narrative Economics. *American Economic Review* 107(4), 967–1004.
- Sievert, C. and K. Shirley (2014). Ldavis: A method for visualizing and interpreting topics. In *Proceedings of the workshop on interactive language learning, visualization, and interfaces*, pp. 63–70.
- Sufi, A. and A. M. Taylor (2021, August). Financial crises: A survey. Working Paper 29155, National Bureau of Economic Research.
- Tetlock, P. C. (2007, June). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance* 62(3), 1139–1168.

Appendix

Text collection procedure

The news article texts used in the paper were collected from Proquest Historical newspapers and Proquest Newstream databases. The articles were downloaded from the year 1870 to the most recent year that was available. The four newspapers from the historical newspapers database had articles from 1870 to 2010 (*Times of India*) or from 1870 to 2016 (*The Washington Post*, *The New York Times*, and *The Globe and Mail*), and the remaining two newspapers had articles from 1982 to 2017 (*The Wall Street Journal*) or from 1989 to 2017 (*The Guardian*). The purpose was to use search keywords that would maximize the likelihood of selecting articles related to the economy. The list of keywords is as follows: *economy, economic, economics, economist, finance, financial, financed, financing, business, production, product, producing, productive, productivity, producer, trade, trading, trader, traded, export, exporting, exported, exporter, import, importing, imported, importer, employment, employed, employee, employer, unemployment, unemployed, jobless, job, industry, industrial, output, depression, recession, deficit, forecast, forecasted, forecasting, forecaster, market, GNP, gross-national-product, GDP, gross-domestic-product, investment, investing, investor, inflation, consumer, consumption, asset, stock, CPI, consumer-price-index, bank, banking, banker, interest-rate, currency, projection, manufacturing, manufacture, manufactured, manufacturer, exchange, boom, boost, devaluation, outlook, prospect, price-level, company, corporation, enterprise, retail, retailer, factory, labor, work, works, worked, working, worker, fiscal, supplier, commerce, commercial, sale, income, profit, revenue, earning, expenditure, expense, spending, resource, dividend, and purchase*). Articles were searched so that the results had to include at least one of the mentioned "economy" keywords and a keyword related to a country name. The keywords for the 17 countries were *Australia, Australian, Belgium, Belgian, Canada, Canadian, Denmark, Danish, Finland, Finnish, France, French, Germany, German, Italy, Italian, Japan, Japanese, Netherlands, Dutch, Norway, Norwegian, Portugal OR Portuguese, Spain, Spanish, Sweden, Swedish, United Kingdom, UK, U.K., Great Britain, England, British, United States, USA, U.S., Switzerland, and Swiss*.

In addition to these two keyword groups, the searched articles had to be from a specific year from 1870 to 2016. The metadata of the 200 most relevant articles with these keywords for each year and country combination were downloaded to get a panel of articles (200 economy-related texts for every newspaper-year-country combinations). The metadata includes the title, abstract, publication date, start page, and document type of each text article. We narrow the texts to front page articles, articles, and editorials. For the two newspapers with articles with shorter time spans, we also include cover stories, news, features, and commentaries. The searches were conducted between February 2020 and May 2020.¹⁵ There were a few newspaper-country-year combinations that did not have any results, but the majority of the combinations had the full 200 results.

¹⁵The specifics can be provided on request from the authors if needed.

Text preprocessing

We describe the necessary text preprocessing procedures conducted on the full set of articles originally gathered from the Proquest database before we fit the topic model or perform any further analyses.

Manual audition of the texts

As the first step before the articles were passed to an extensive text cleaning phase, we glanced at some randomly picked articles. We immediately noticed that the corpus included articles that did not exactly describe the content that was intended by containing the specific keyword or country. This was mostly because the articles might include *some* form of the specific keyword or country name, e.g., "france" - "franceskos", "company" - "accompany". As a result, we decided to change some of the keywords and country-specific indicators to more *precisely* defined, that is, no other forms of a certain word were allowed, but the exact form of the word itself. In addition, we noticed that words related to the base word "work" were frequently not related to the economy, as those articles included a large share of news about some artistic and cultural piece of work or performance; thus, these keywords were removed from the title keyword search.

Consequently, we decided to do a systematic manual audit of the articles included in the original corpus to see if some of the keywords still resulted in articles that are frequently related to something other than the economy. In many cases, these "bad" articles included news about culture and sports. For the audit, we took all the articles from a certain decade and country and manually checked them one by one to see if they were economy related to at least some level and what the keyword was if the article was considered "bad". We tried to include as many decades and countries as possible to avoid timing or location bias, as some topics (e.g., cultural) might be more relevant during some times and in some places than others. The audit sample of articles that we checked manually consisted of a total 5727 articles. We kept track of if the article was "bad" according to our criterion and what keyword was found in that article. As a result, the share of articles that we labeled "bad" was 6.1%, and we did not exclude any further keywords.

Furthermore, during the audition, we noticed that the length of the abstracts varied considerably, especially across years. Thus, we decided to use the titles of the news articles instead of abstracts. This was also because that usually the title includes the main information of the news, and the actual text might bring up more random noise in the data. We decided to form our corpus by using the keyword and word list and that those were have to appear in the title of the article; that is, for the inclusion of an article, it had to mention at least one of the keywords and countries in the title. With this definition, we wanted to increase the probability that the included titles "make sense".

Text cleaning

Before we can actually use the historical newspaper data for the analysis purposes, the text data must be "cleaned" thoroughly. This is done before forming the document-term-matrix (DTM), which is used in the fitting of the topic model. It is particularly important to remove all the unnecessary signs and words as we want to identify interesting and distinct patterns across the titles apart from random noise. We did the text preparation and cleaning with the following steps:

- 1) We use so-called uni-grams, one-word terms, to capture the essential information regarding the economic news.

- 2) We remove any special characters, numbers, URLs and a list of common English stop-words (e.g., "about", "each", "have", "very").

- 3) We remove a custom list of proper nouns, such as names of people, places, companies, etc.

- 4) We remove text written numbers, months, weekdays, and currencies.

- 5) We remove all terms of three signs or less.

- 6) We include only words that occur in all three of the 50-year periods (1870–1919, 1920–1969, 1970–2016) and at least in 8 out of the 17 sample countries, thus, implying vocabulary that is independent across different time periods and countries.

- 7) We remove all articles that contain no words that are included in the final corpus (technical issue as the zero rows in the DTM do not add any information to the topic model).

Topics: Descriptive tables

Table A1: Labels and most common words of topics 1–55.

Topic	Label	Title
1	Positive market movements*	gain-total-record-purchases-larger-farm-jump
2	Seasonal reporting/bookkeeping	review-weekly-spring-reviews-books-features-failures
3	Industrial production	output-production-industrial-speed-newsprint-aluminum-copper
4	Economic growth	economy-growth-spur-signs-grows-slowness-pace
5	Central bank	central-bank-governor-institution-raises-eyes-credit
6	Fiscal policy*	cuts-defense-plans-spending-worried-tourists-program
7	Surveys	finds-survey-study-profitable-manufacturer-edge-poll
8	Economic atmosphere	confidence-sentiment-game-farmer-mood-caution-spree
9	Commercial banking	banks-savings-accounts-cash-lending-institutions-fail
10	Specific Industry	industry-textile-competition-silk-film-iron-electrical
11	Political proposal lifecycle	proposals-representatives-resistance-pledges-restored-freedom-cease
12	Investors' economic outlook	investors-briefs-look-wary-feel-yields-interest
13	Culture*	asset-dilemma-times-management-cultural-weighting-consequences
14	Commercial treaties	commercial-relations-treaty-reciprocity-treaties-interests-arrangement
15	Colonial issues	colony-floor-cutting-pearson-assembly-resented-armed
16	Investigations and charges	inquiry-trader-involves-investigating-charges-probe-bribery
17	Competition in international trade*	revival-dispute-losing-scope-returns-pushing-consul
18	Important market events in specific location	topics-markets-incident-wall-comment-elsewhere-local
19	Crime or abuse revealed	banker-wife-book-admits-leave-charged-arrested
20	Telegraphs and cables	latest-telegraph-transatlantic-prince-cables-cable-burned
21	Plans and ideas*	idea-whose-forward-importing-stress-revived-accept
22	Suppliers*	major-appears-supplier-becomes-bought-pork-steadily
23	Money markets*	money-current-call-events-tight-developments-speculation
24	Unemployment rate	rate-jobless-unemployment-highest-lowest-discount-drops
25	Economic policy agreements and cooperation	economic-ties-problems-impact-policies-stability-summit
26	Prospects	prospects-mail-bright-letter-poor-brighter-good
27	Economists' expectations	expect-economists-predict-believe-optimistic-forecast-weighed
28	Banking crises	banking-system-giant-scandal-beats-sector-crises
29	Bond markets, IPOs and deals	enter-watch-offering-bacon-enters-expanding-bond
30	International trade restrictions - Imports	curbs-import-restrictions-barriers-retaliation-quota-eased
31	International trade restrictions - Exports	drive-export-prepares-uranium-doubtful-world-wide-approved
32	Strifes and confictions*	satisfaction-source-supports-indications-seeds-renews-dictator
33	High officials and representatives	prime-minister-dinner-resigns-foster-quits-ambassador
34	Taxation	taxes-income-fixed-taxation-affects-levies-billions
35	Counseling and lecturing*	extent-bound-lakes-seizures-mountains-engage-enterprises
36	Subsidiaries and business actions*	foundation-popular-disposal-initiative-loyal-completes-opposed
37	Labour, unions and strikes	labor-party-unions-leader-strikes-reds-vote
38	Inflation	consumer-inflation-check-leaving-wage-restraint-perils
39	Company acquisitions	company-maker-phone-telephone-electric-venture-trust
40	Analysis and outlook*	improved-avoided-alarm-tumble-hurting-term-response
41	Movements in trade	expects-hopeful-drop-benefit-wars-district-agency
42	Battles, hostilities and war	fight-style-winning-battle-supremacy-hostile-breaking
43	Foreign investments	investments-lift-foreigners-changing-shift-safe-welcomes
44	Firm level sales and purchases	importer-shut-firm-option-buys-sold-owner
45	High level and government economists	economics-politics-sights-lesson-lessons-search-nobel
46	Financial market movements	trading-issues-advances-quiet-list-rails-mixed
47	Alcohol	wine-wines-liquor-grievances-spirits-beer-whiskey
48	Government bills, resolutions and statements	colonies-bill-senate-measure-islands-senator-passed
49	Government level conflicts and wars	story-hear-false-rumored-prisoners-disasters-disappointed
50	Reports on international trade	exports-imports-exceed-curbed-exceeded-beef-rice
51	Countries' economic relations	large-finance-postwar-retains-responsibility-thus-autonomy
52	Balance of trade	deficit-billion-widens-payments-budget-surplus-widened
53	Trade policy	policy-fiscal-industries-defends-adopt-technical-fair
54	Rearrangement of firm operations and mergers	firms-companies-concerns-insurance-invest-chain-mergers
55	Noble metals	gold-metal-standard-silver-reserves-reserve-shipment

Notes: In the table, * indicates that the labelling a specific topic was not as clear as for the majority of topics.

Table A2: Labels and most common words of topics 56–110.

Topic	Label	Title
56	Exchange rate movements	moves-currency-falls-devaluation-single-traders-closes
57	Transactions	horses-class-purchased-water-bread-kind-chiefly
58	National supplies shortage	needs-food-supply-requirements-adequate-assure-material
59	Agriculture and nature	live-disease-timber-sweeping-breeders-poultry-diamonds
60	Pacts and treaties	pact-accord-agreement-sign-chamber-signed-pacts
61	Freezing and overtaking of assets	assets-sell-blocked-claim-freeze-property-blocks
62	Recession*	felt-recession-unlikely-predicted-toll-worry-mild
63	Rumours, stories and propaganda	spirit-stories-caused-spots-streets-sick-ended
64	Highly specialised products and producers	enterprise-route-developing-road-conversion-success-things
65	Positive quarterly corporate reports	profits-earnings-corporate-dividend-reports-dividends-corporation
66	Jobs	jobs-lose-create-women-creates-employment-eliminate
67	Periodic production numbers	international-annual-journal-cent-manufacturing-awaits-climb
68	Loans and aid*	dependent-less-gasoline-novel-resort-owing-mainly
69	Sovereign seeking economic aid or cooperation	seeks-halt-bigger-group-access-avert-branch
70	Industrial manufacturing	manufacture-drugs-process-project-cars-engines-fine
71	Stamps	placed-stamp-stamps-penalties-adopts-freely-stations
72	Reports on business conditions*	looking-ways-throughout-world-talking-means-shaken
73	Law, court and bills	supreme-court-case-judge-decisions-decision-settles
74	Antique and design	built-design-furniture-unexpected-submarines-produces-castle
75	Public appointments	department-town-division-navy-vessel-officers-corps
76	Economic outlook and planning	post-war-model-later-tension-proves-carry-believes
77	Promises	promise-given-failing-breaks-authorized-drawn-advice
78	Free trade difficulties and barriers	freer-empire-free-merchants-discuss-extend-dominions
79	Wheat, crops and harvesting	effects-estimates-drought-widespread-corn-acreage-wheat
80	Common market	back-entry-common-commonwealth-door-bloc-beat
81	Economic or political influence and control	mine-critics-fate-learned-influence-retreat-summer
82	Diplomatic relations	reply-grants-granted-powers-issued-encouragement-recognize
83	Change in international trade barriers	duty-materials-items-cloth-bans-autos-wool
84	Firm financing	million-spend-link-distribution-mass-spent-least
85	Firm profits	report-profit-trio-revise-posts-perks-shows
86	Boost of something	boost-gives-gets-steps-hopes-green-defenses
87	Clash	holders-strategy-aims-clash-aggressive-pool-reveal
88	Foreign investment	like-potential-climate-investment-foreseen-shape-welcome
89	Trade and marketing	marketing-guarantee-differ-widely-advertising-differences-vary
90	Gains and ahead	ahead-without-puts-tough-keeping-prove-exporting
91	Planes	bars-planes-remove-blow-coffee-moral-virtually
92	Commodities	product-domestic-gross-butter-enlarged-dairy-type
93	Economists	economist-predicts-noted-leading-professor-criticizes-diplomat
94	Global business	business-around-wire-exporter-area-forum-third
95	Individual solutions and resolutions	solution-worked-agents-truth-rule-arrive-newspapers
96	Factories and resources*	resources-factories-wealth-rich-natural-vast-development
97	Relief and help	relief-busy-unemployed-fund-cross-idle-glad
98	Businessmen visits	businessmen-visit-mean-ability-visiting-businessman-lacking
99	Crises and dangers	coal-fuel-acute-nears-miners-middle-east
100	Royals and diplomats	unique-happy-appear-queen-indication-couple-fault
101	Matches and fights*	prospect-hill-sports-invasion-cheered-complications-dark
102	Trade control and regulation	laws-regulations-tighten-shippers-neutral-stiff-fought
103	Stock exchange	closing-tone-prices-boerse-quotations-trend-upward
104	Commodity producers	producers-pleased-longer-tactics-importers-regret-dealers
105	Communities and institutions	notes-school-city-church-society-college-association
106	Marine ships and other big purchases*	ships-instead-running-size-potash-pictures-voted
107	Government officials' speeches	praises-speaks-official-snap-practice-tribute-employees
108	Financial crises	position-difficulties-services-matters-panic-circles-condition
109	Investing*	economies-ending-test-revised-trials-pessimistic-fights
110	Miscellaneous*	board-effect-talk-still-realism-foreign-dominion

Notes: In the table, * indicates that the labelling a specific topic was not as clear as for the majority of topics.

Table A3: The most representative texts of selected topics.

Topic	Label	Proportion	Text
1	Positive market movements*	92 %	canada maintains balance in trade; favorable november figures show \$8,852,377 over imports.
3	Industrial production	96 %	newsprint output at record level: u. s. and canadian production in may totals 509,412 tons - consumption also up newsprint output at record level
4	Economic growth	93 %	british fear u. s. hurts recovery; effect of strong dollar, weak economy cited british fear u. s. hurts recovery
5	Central bank	93 %	addresses made at the annual meeting of shareholders: the royal bank of canada address of chairman and president
6	Fiscal policy*	95 %	denmark budgets defense increase: maps spending of \$57,100,000 in 2 years - french pledge set at 10% of revenue denmark budgets defense increase france pledges increase norway to double outlay
8	Economic atmosphere	84 %	canadian athletes feeling the spending pinch
9	Commercial banking	92 %	u. s. banks balk at long terms for polish debt: u. s. banks balk at extension of polish debt
10	Specific industry	93 %	japanese automobile industry now faces u. s. competition
17	Competition in international trade*	92 %	east germany sees trade obstacles; east germany sees trade obstacles
18	Important market event in specific location	96 %	wheat is lower: the western market dull and weaker canadian stocks easier wall street securities inactive and easier failures in canada - local breadstuffs and other markets - features in groceries
23	Money markets*	92 %	current affairs in england: events of the week nr. gladstone at greenwich money market state of trade
24	Unemployment rate	94 %	unemployment rate drops to 6.1 percent: u. s. jobless level is lowest in 7 1/2 years nation unemployment rate hits 7 1/2-year low
25	Economic policy agreements and cooperation	94 %	economic doubt on uniting korea: the problems of german unity are causing Seoul to turn cautious. economic doubts emerge on efforts to unite the 2 koreans
27	Economists' expectations	89 %	england sees u. s. boom as just begun: british economists assert they are in second prosperity year.
28	Banking crises	92 %	scandal could alter swiss banking: swiss banking facing changes in wake of scandal
29	Bond markets, IPOs and deals	90 %	on italy black market, dollar continues strong: dollar is strong on black market
30	International trade restrictions - Imports	93 %	france removes 200 import curbs: lifts quotas on goods from u. s. canada and europe france removes 200 import curbs
31	International trade restrictions - Exports	93 %	u. s. lifts controls on export of drilling gear to soviet: u. s. permits export of drilling gear
38	Inflation	94 %	u. s. fears 'runaway inflation': treasury hints at controls on wages, prices
40	Analysis and outlook*	84 %	near term outlook for u. s. called favorable
43	Foreign investments	92 %	u. s. investments upset canadians: ottawa eager to implement law curbing take-overs legislation is spurred canada upset over u. s. investments
44	Firm level sales and purchases	88 %	data sought on thomson deals with ihya, inq: u. s. probes sale of ihv unit to french firm
45	High level and government economists	93 %	norwegian awarded nobel for economics: trygve haavelmo honored for '48 work in science of econometrics norwegian awarded nobel prize
47	Alcohol	92 %	defeating prohibition: one month employed on the main border swedish settlers introduced skin- the cite yankee had them made hollow, and then brought in whiskey from canada
50	Report on international trade	93 %	heavy decrease in trade with europe: february imports from europe down \$1,000,000 below 1920. exports down \$142,000,000 german trade large exports to asia and south americaincrease heavily; also imports from them.
51	Countries: economic relations	87 %	finance at large: canada hears from germany last democratic chancellor how germans forged constitution
52	Balance of trade	95 %	'85 trade deficit is worst ever: \$148.5 billion gap expected to intensify pressures on congress '85 u. s. trade deficit hits \$148 billion; protectionist pressures seen intensifying
53	Trade policy	94 %	tariff advocated by british liberal: ardent free trade advocate, sir john simon, reverses own stand. tariff advocated by british liberal
54	Rearrangement of firm operations and mergers	92 %	set merger to cut costs: usn acquires last 20% of british subsidiary in \$34-million deal companies take merger actions
56	Exchange rate movements	94 %	dollar plagued by peso: dollar dips as the peso falls again mexico crisis hurts u. s. currency value dollar slips as the peso continues sharp fall
60	Pacts and treaties	96 %	italy, america sign stop-gap pact on trade: hill announces temporary treaty, effective wednesday. u. s. and italy sign trade pact pending draft of new treaty
62	Recession*	93 %	pain spreads to the heart of europe: euro zone remains mired in six-quarter slowdown as france slips into recession and germany totters on the brink of one
66	Jobs	91 %	civil service discloses 400 high-pay jobs: log jam of promotions broken. u. s. posts in \$11,200-14,000 range civil service discloses 400 high-pay jobs
67	Periodic production numbers	88 %	eurobonds decline on slumping dollar, u. s. retail sales keep - special to the wall street journal
69	Sovereign seeking economic aid or cooperation	92 %	danish concern is go-between for u. s. clients on china trade: danish concern is go-between for u. s. clients on china trade
70	Industrial manufacturing	92 %	remarkable activity in oil engine manufacture: three new british oil engines big oil engine plant for b. c. north regional station
72	Reports on business conditions*	88 %	the advertising world: good market-testing centres hard to find in canada
77	Promises	92 %	canada as pioneer: credit given us for the preferential tariff british manufacturers should take full advantage of the opening - the london outlook on the situation
80	Common market	93 %	snags still exist in kennedy round: u. s. and common market at odds on '3d countries' u. s. and the common market hit snags at the kennedy round
83	Change in international trade barriers	93 %	higher duties on imports: australia tariff changes affect products supplied by canada want reciprocal trade
85	Firm profits	86 %	u. s. corporate profit margins up in quarter
86	Boost of something	88 %	gilbert comes through for u. s.: gives americans lead in davis cup; agassi-becker tied gilbert gives u. s. a boost in davis cup; agassi-becker tied
88	Foreign investment	92 %	foreign investment is waxing australia: australia vexed by foreign funds one-quarter foreign ownership
89	Trade and marketing	92 %	session means end and toy cohorts prepare for fey: mt liberals are planning to enlarge canada trade relationship u. s. tariff reactions
90	Gains and ahead	87 %	manufacturing not the only thing being outsourced by the u. s.: ahead of the curve
92	Commodities	92 %	potato is everywhere: from bakery to feed bin, u. s. product is in demand: even farmers speculate in futures market
93	Economists	89 %	iron head hands canadian employer: hillman predicts faster clothing trade pick-up in canada british outlook here
94	Global business	86 %	business around the world: brewerries merging in norway
102	Trade control and regulation	92 %	quarantine abuses: still further changes against the quarantine officials: further serious changes-violation of the united states revenue laws-spanish vessels and their ballast.
107	Government officials' speeches	88 %	treasury chief defends outsourcing of u. s. work: snow calls the practice 'part of trade'
108	Financial crises	92 %	swiss are critical of german terms: bankers disappointed military government will not permit financial transfers
109	Investing*	90 %	investing in funds kampf, cfrs: a monthly analysis - global investing: swiss pros and cons - is the value still there for the loved country?

Notes: Column 3 shows how much of a specific article titles consists of a specific topic. We have selected here examples of articles where the proportion of an article title is over 80% of a specific topic and hence highly representative of that topic.