Computational Economics
https://doi.org/10.1007/s10614-024-10594-5

Can Text-Based Statistical Models Reveal Impending Banking Crises?

Emile du Plessis[1]

Accepted: 18 March 2024

Abstract

This paper introduces statistical models Wordscores and Wordfish to study and predict banking crises. While Wordscores is akin to supervised learning, Wordfish is analogous to unsupervised learning. Both methods estimate the position of banking distress on a tranquil-to-crisis spectrum. Findings suggest that the two statistical methods signal banking crisis up to two-years in advance, with robust results from AUROC, Granger causality and VAR impulse responses. Both methods outperform random forests in predicting crises using textual data. The Wordscores index highlights increased usage of banking sector nomenclature two years preceding a crisis, and Granger causes a crisis series with one and two lag lengths. Results from the Wordfish technique, a statistical model with Poisson distribution, show the index spikes before and during the Global Financial Crisis, when a large share of the countries in the world encountered banking crises. This paper contributes to literature on text-based models of banking crises by bolstering the preemptive policy responses available to policy makers. Given their early warning signals, both Wordscores and Wordfish can be considered a part of the toolset to monitor the stability and resilience of the banking sector.

Keywords Quantitative analysis of textual data · Banking crises · Text-based models · Early warning signal

JEL Classification C49 · C53 · C54 · C55 · G21

I would like to thank Ulrich Fritsche for his guidance, and conference participants at the Annual Meeting 2022 held by the American Economic Association and Allied Social Sciences Association, 22nd Conference on International Economics held by the Spanish Association of International Economics and Finance, Annual Congress 2022 of the Swiss Society of Economics and Statistics, 18th Augustin Cournot Doctoral Days held by the University of Strasbourg, PhD Seminar Series at the University of Hamburg and Sophie Béreau-Gnabo for valuable suggestions and comments.

✉ Emile du Plessis
grensduplessis@gmail.com

[1] Department of Economics, University of Hamburg, Hamburg, Germany

Published online: 13 April 2024

1 Introduction

As the Global Financial Crisis surfaced vulnerabilities within financial markets, concerns on the sustainability of the banking sector led to renewed inquiries into the balance sheet strength and risk profiles associated with financial intermediaries. Given that banks continue to play a vital role in the allocation of scarce resources, including through their asset and liability mix, and maturity transformation capabilities, access to an expeditious policy framework moves to the forefront of contemporaneous discourse. With qualitative research methods gaining prominence as policy instruments over recent years, and elucidating a deeper understanding of economic phenomena, this paper introduces statistical modelling approaches to study banking crises.

In this regard, two text-based indices are developed, which are both centred on the measurement of similarity between texts, albeit through unique operational methods. For both indices, samples of texts of a rich and untapped dataset spanning over 19 million articles are selected based on a lexicon informed by banking crisis literature, and then pre-processed and converted into textual data. Furthermore, the content of articles is analysed to verify its relevance to the theme of banking sector performance. Thereafter, the two statistical models are utilised to study the intensity of word usage, in order to indicate shifts in nomenclature over time, which could indicate an associated position on a tranquil-to-crisis spectrum. The statistical modelling frameworks behind both indices have been employed within political science literature to accurately indicate policy positions of texts such as party speeches and manifestos, on a scale from a left wing to a right wing. The relevance of locating positions on such a spectrum to a banking crisis, relates to the oscillation of banking sector health, between a tranquil period to that of deepening levels of distress. In a stable economic environment, free of a recession, and with positive growth, the banking sector is strengthened through higher deposit inflows and low-risk lending. However, an overstimulated economy and unsustainable increase in asset growth and fixed capital formation, contributing to a perilous bubble, and eventually followed by a bust, result in banks confronted by liquidity constraints, a bank run, or deteriorating asset quality. This pendulum between a low-risk to high-risk environment continues to swing within both developed economies and emerging markets.

This paper contributes to the literature by employing statistical models Wordscores and Wordfish to study and forecast the binary conditions of banking crises, from a tranquil environment to a crisis state. To our knowledge, this is the first attempt to study banking crises through these statistical tools, both of which have demonstrated robust performance for text analysis in the policy making domain. Wordscores is based on a supervised learning approach, where the directional movements in the index, as generated by the model, is based on learning from a pre-specified anchor. Through the Wordscores technique, virgin texts are thereby compared to reference texts, to determine the position of the virgin texts on the tranquil-crisis scale. Wordfish constitutes an unsupervised learning approach, where word intensity informs the position of a text on the tranquil-crisis scale.

According to the findings, the Wordscores index highlights increased usage of banking sector nomenclature two years preceding a crisis, and Granger causes a share-of-world-in-crisis series and Global Financial Crisis series, with one and two lag lengths. The second index, as based on the Wordfish technique, is a statistical model with Poisson distribution. Results show the index spikes before and during the Global Financial Crisis, when a large share of the countries in the world encountered banking crises, and Granger causes the Global Financial Crisis series at one lag length. Both statistical methods outperform the commonly used random forests algorithm in predicting a crisis using textual data.

One of the challenges of text mining is centred on its ability to detect contextual meaning from the phrases in the corpus. This is addressed through the selection of relevant news articles as based on precursors to banking crises; pre-processing of textual data by implementing hand-coding to deal with the high dimension of written material, distribution of words, and degrees of analytical value of the words; and lastly validated through content inspection in order to verify the relevance of each text.

The structure of the paper is as follows. Section 2 provides an overview of the empirical literature. Section 3 describes the methodology, data, and variable selection. Section 4 evaluates the selected text through hand-coding, while Sects. 5 and 6 explain statistical applications Wordscores and Wordfish. Section 7 comments on comparable performance, Sect. 8 on a robustness test and Sect. 8 on policy implications. Section 10 concludes.

2 Empirical Literature

2.1 Text Analysis

The use of text analysis has been gradually increasing over recent years with the expediency provided by computer-based algorithms (Bholat et al., 2015). Considered an important resource for monitoring the evolution of discourse and policy positions over time, initial attempts were highly labour and time intensive. Manually, human coders have been analysing large volumes of text such as the Comparative Manifestos Project (Budge et al., 1987; Laver & Budge, 1992; Volkens, 2001). The democratisation of computer programmes, and contemporaneous improvements in text mining capabilities, have elevated the interest in qualitative data formats to investigate the contribution of unstructured information to the understanding of explanatory factors, and the oscillation of political, social, and economic developments.

To date, relatively few studies have focused on the use of text analysis in economics, partly due to the abundance of quantitative data sources. Most of the existing literature on text-based analyses pertains to the extraction of policy positions from texts, inclusive of party manifestos, legislative speeches, campaign speeches, constitutional negotiations, and judicial decisions (Benoit et al., 2005; Giannetti & Laver, 2005; Hug & Schulz, 2007; Klüver, 2009; Laver & Benoit, 2002; Laver et al., 2006; Proksch & Slapin, 2006). In particular, Benoit and Laver (2003) and Laver et al. (2003) advance a statistical model for analysing data in the form of words, based on reference texts. On an economic and social policy dimension, their Wordscores approach replicates published policy estimates in a more expedient manner. In a related study on the influence of interest groups on European Union policy making, Klüver (2009) finds a high correlation of actual policy positions to those computed from texts. To improve on the need for a priori policy positions, Proksch and Slapin (2008, 2009) advance Wordfish, a computer-based model employing the unique words as unit of analysis to compare texts based on relative word use. The study shows the technique to be highly robust across three languages, English, French and German, to hold for national and regional texts, and with comparable results irrespective of distributional assumptions and document selection. Research by Klüver (2009) underscores the expediency of the Wordfish model and reliability of results compared to policy positions identified through hand-coding.

Studied have also expanded to financial markets, and the use of alternative tools. In comparing similarities between texts, Deerwester et al. (1990) apply Latent Semantic Analysis to estimate linear combinations of terms and documents as based on highest variation. In particular, Acosta (2015) studies transparency in the minutes from the US Federal Reserve Open Market Committee to show increased conformity of members following the recording of deliberations. For the same research question and rendering similar results, Hansen et al. (2018) apply Latent Dirichlet Allocation in which terms and documents are allocated a probabilistic weight of pertaining to a specific topic. Kloptchenko et al. (2004) employ Euclidean distance to identify groups of financial reports, whereas Hoberg and Phillips (2010) make use of vector space modelling to compute the cosine commonality between company statements. More recently, sentiment analysis has become a popular approach for risk analysis in financial markets (Kou et al., 2019), including to measure the oscillation between positive and negative emotions in texts to predict stock markets and financial crises (Calomiris & Mamaysky, 2018; Heston & Sinha, 2017; Püttmann, 2018), and assess its impact on asset prices (García, 2013). Shapiro et al. (2020) show that news sentiment correlates with the business cycle, González-Fernández and González-Velasco (2020) find a sentiment index based on Google data to be useful to predict financial distress, while Nyman et al. (2021) demonstrate how sentiment changes before the Global Financial Crisis. Chen et al. (2023) use machine learning methods to analyse texts and identify financial crises. Other approaches include regression and ranking methods (Tsai & Wang, 2017), lexicon and sentiment analysis (Meyer et al., 2017), and Rule-Based Emission Model algorithm (Tromp et al., 2017).

To our knowledge, no other studies have used the Wordscores and Wordfish methods to study banking crises, so this would be a key contribution of the paper. These two statistical techniques are introduced to evaluate the performance strength of text as data, in order to categorise and translate the changing phases of economic discourse and disposition in signalling ensuing banking crises. Secondly, while recent techniques such as machine learning methods are also being used in the literature, this paper focusses on established statistical methods with an extensive track record, and which comes with a deeper understanding of the strengths, drawbacks and performance of these tools. The accuracy of these statistical tools has been demonstrated on a consist basis in analysing and classifying proposals for policy making. Yet, a robustness test is implemented using random forests to compare predictive strength. Thirdly, while studies on the alternative methods are based on supervised learning, the two statistical approaches in this paper feature both supervised and unsupervised learning, to determine optimal outcomes and fit for purpose models. Fourthly, the paper makes use of a rich, large, and mostly untapped data source, and which is specifically tailored to the banking sector. News articles are used instead of commentary and opinion pieces, as the latter could skew the results. Fifthly, this study focusses specifically on banking crises and furthermore contributes to the prediction of banking crises. Sixthly, in contrast to several text mining studies employing sentiment analysis, which is frequently shown as concurrent indicator of economic events, generally rising and declining in tandem, this paper analyses thematic reporting, in order to generate leading indicators of banking crises.

2.2 Precursors to Banking Crises

Literature on precursors to banking crises are used to develop a lexicon on banking crises, in turn to narrow down news articles for purposes of building the statistical models, and forecasting banking crises. In summary, findings feature leading indicators from the real, banking and external sectors.

Banking failures are commonly the result of weak macroeconomic fundamentals (Calomiris & Mason, 2003; Gorton, 1988; Lindgren et al., 1996; Minsky, 1982; Temin, 1976; Wicker, 1980), as well as credit booms and asset bubbles (Kindleberger, 1978; Reinhart & Rogoff, 2009), recently due to financial deregulation, expansionary fiscal policies, and overinvestment in real estate (Drees & Pazarbasioglu, 1998; Laeven & Valencia, 2010; Ranciere et al., 2008). Kaminsky and Reinhart (1999) show that banking crises follow after economic recessions, diminished terms of trade, escalating interest rates, a stock market collapse, and a currency crisis. Demirgüç-Kunt and Detragiache (1998) show that accelerating inflation also increases the probability of failure. In addition, Hardy and Pazarbasioglu (1998) suggest that a decreasing capital output ratio, a sizeable depreciation in the exchange rate, and trade shocks are associated with banking crises. Joy et al. (2015) highlight tight interest rate spreads, inverted yield curves, and housing prices as leading indicators. Alessi et al. (2015) further underscore depressed bank profitability and government debt as precursors. Samitas et al. (2020) demonstrate the impact of contagion on the propagation of financial crises.

As validation, content analysis also scrutinises the relevance of the selected texts, by verifying the presence of these leading indicators, in order to support the empirical strength of the models in studying and predicting banking crises.

3 Text Data and Approach

3.1 Textual Data

To measure the similarity between texts, this paper makes use of a vast number of articles that have sparsely been studied to date. Thomson Reuters News Archive (2020) consists of over 19 million news articles at the time of this study. Its international coverage spanning several developed economies and emerging markets benefits the study of the Global Financial Crisis (GFC), which similarly affected many economies. Articles are composed by journalists, and newsfeeds are in turn then consumed by broadcasters, publishers, financial and policy analysts, and the general public. Certain customisations are applied to the archive, and these include source, date, language, and text. Source is constrained to the News Archive, the date ranges from 2004 to 2012 while language is limited to English. Serving as contextually relevant news feeds, the texts feature business and financial markets reporting as opposed to opinion pieces and editorial commentary. Given the exclusive focus on banking crises, only news articles containing information on all three sectors inclusive of real, banking, and external sector developments are considered. In studying precursors to banking crises, the relationships between these three integrated economic sectors are frequently confirmed in literature (Du Plessis, 2022a, 2022b; Kindleberger, 1978; Reinhart & Rogoff, 2009). Specifically, González-Hermosillo et al. (1997) find that banking sector predictors describe the likelihood of a banking failure, whereas real sector predictors determine its timing. The lexicon based on the precursors to banking crises and used to select the relevant articles is summarised in Table 11 (Appendix 1).

To efficiently manage the large number of articles, and to compare it to the annual crisis classifications from the Reinhart and Rogoff (2014) database, all textual data is aggregated and analysed on an annual basis. In comparison to other sources, including Laeven and Valencia (2013), and Romer and Romer (2017), crisis classification is broadly similar. Reinhart and Rogoff (2014) feature crisis identification over an extended long-term horizon, thereby serving as one of the most comprehensive datasets to date. The benefits of annual series are emphasised on predictive strength by Calomiris and Mamaysky (2018), whereas Sinha (2016) and Heston and Sinha (2017) exemplify the usefulness of aggregating news over extended horizons. Due to the limited number of articles available in previous decades, the focus is on the 2000s period, and which also coincides with the Global Financial Crisis, which impacted most countries between 2008 and 2009.

In support of macroprudential policy, two banking crisis indices are constructed, as based on financial newspaper text analysis. Both indices examine intensity of word utilisation within the corpora of a sample of texts. Essentially, it conveys degrees of importance of banking specific terms. Whereas the first index makes use of the Wordscores approach, the second focusses on the Wordfish technique. The two indices are developed as classification methodologies, generally used for political texts, in order to locate policy views on the left-centre-right political spectrum. By inculcating these techniques into a banking crisis domain, it serves to locate a position of a latent banking distress indicator, ranging from a vulnerable banking sector experiencing a crisis on the one side of the spectrum, through stasis to a more robust banking sector on the other edge, without a crisis. While the two sides are akin to the left and right wings of the political spectrum, a central position would reflect a mid-point, where a direction in movement could indicate a shift towards or away from a crisis.

Plotted in Table 1, the banking crisis spectrum is overlaid onto the left–right political axes. Speeches by politicians from political parties are shown to emphasise

Table 1 Plotting Banking Crises on Political Spectrum

Left-wing	**Right-wing**
'Socialism'	'Capitalism'
'Equality'	'Private sector'
'Wealth tax'	'Tax cuts'
'Social programmes'	'Conservative'
'Regulation'	'Red tape'
'…'	'…'
No-crisis	**Banking crisis**
'Stable'	'Distress'
'Certain'	'Uncertain'
'Low rates'	'Volatile'
'Growth'	'Stagnate'
'Robust'	'Weakness'
'Tranquil'	'Crisis'
'…'	'…'

For illustrative purposes, the banking crisis spectrum of no-crisis to crisis is overlaid onto the left-wing to right-wing political axes. The dimensions for both domains featuring banking crisis and political disposition can be characterised by having opposite poles, however, these two domains are separate and are not connected to each other based on their assigned poles. The banking crisis outcomes are agnostic to their position on the left–right scale, and no-crisis and crisis can be placed either left or right on the scale, without changing results

their ideological positioning through word selection and usage. Left-wing manifestos, speeches and publications are likely to feature declarations on socialism, policies to reduce income inequality, higher tax on the wealthy, strategies for government spending on social programmes, and proposals on stronger regulation of businesses. In contrast, statements and transcriptions by right-wing politicians would emphasise the role of capitalism in the economy, propositions for conservative government spending, philosophies on lower taxes, and intentions to remove red tape.

As both left-wing and right-wing terminology signals distinct environments, distinguishable to readers, observers, and policy analysts, so too is there a distinction between environments grappling with a crisis, and one with stable growth and a healthy operating landscape, and the terminology used by journalists, analysts and policy makers to describe these environments. During the formation of a crisis, reporting would contain more concerning narratives on economic events such as distress, uncertainty, volatility, stagnation, weakness, and crisis. In comparison, absent an emergency, descriptions would feature an account of a stable environment, with terms such as certainty, stable interest rates and low inflation, mild unemployment, with positive domestic growth, robust performance of the banking sector and a tranquil operating landscape. Similar to the inclusion of a left–right dimension, the no-crisis to crisis spectrum is also implemented through the use of Wordscores and Wordfish.

One shortfall of statistical-based models such as Wordscores and Wordfish is the impact that decreasing document length could have on lowering performance (Egerod & Klemmensen, 2019). This is addressed in the methodological approach by selecting random samples of comparably sized articles, which in the case of business news is frequently consistent in setting lower bound and upper bound limits to article length.

The dating and definition of a banking crisis is based on Reinhart and Rogoff (2009, 2014), which define it as diminishing banking sector deposits that lead to (1) the closure, merger, or takeover of a financial institution by the government, or (2) the provision of financial assistance to a financial institution by the government. As the GFC was experienced by several countries and spread through trade links and spillovers to other countries, two crisis series are constructed. The first series comprises the share of countries in the world encountering banking crises, which increased from 3 percent in 2006 to around 50 percent in 2009. It can be expected that as more countries experience a crisis, it would coincide with more extensive crisis related news reporting. The second series applies the GFC as the crisis episode to forecast. GFC started in the US in 2007, and most countries experienced related banking distress in 2008 and 2009. These years are implemented in dummy format, with assigned values of $y = 1$ for the GFC, and the balance of the years in the sample taking the value of $y = 0$. This index enables the assessment of predictive strength, as a discrete data format is required, with variability in the outcome series, and with actual crisis episodes. Crisis classification to inform the share-of-world-in-crisis series, and Global Financial Crisis series is obtained from Reinhart and Rogoff (2014).

3.2 Performance Measures

The performance of Wordscores and Wordfish is assessed through three main measures, namely Granger causality, receiver operating characteristics with area under curve (AUROC) estimates, and impulse responses applied to a vector autoregression and captured through Cholesky decomposition.

Performance is firstly assessed through Granger causality (Granger, 1969), which is a method that is regularly used to establish the direction of influence between two series, and a useful measure employed for textual analysis (Nyman et al., 2021; Tucket, 2017). The aim of Granger causality is to determine if one series predicts another series, as estimated by utilising the historical values of the one series to predict the future values of the other time series. Lagged values of the outcome variable are determined through a univariate autoregression of y, where $y_t = c + A_1 y_{t-1} + \cdots + A_p y_{t-p} + e_t$. Consequently, the autoregression is enhanced by adding lagged values of x, so that $y_t = c + A_1 y_{t-1} + B_1 x_{t-1} + \cdots + A_p y_{t-p} + B_p x_{t-m} + e_t$. Through a sorting process, only significant values of x are kept in the equation, as measured by t-statistics, and conditional on x contributing explanatory power to the regression equation as based on F-test measures. Null hypothesis that x does not Granger cause y cannot be rejected if any values of x remain in the final regression equation. Granger causality is operationalised on both share-of-world-in-crisis and GFC series.

The second performance measure incorporates receiver operating characteristics (ROC) with area under curve (AUC) estimates to evaluate predictive strength. Predicted values from the Wordscores and Wordfish indices are assessed through the use of the Global Financial Crisis series. AUROC curves are commonly used to assess predictions, including in studies on business cycle (Berge & Jordà, 2011; Döpke et al., 2017) and banking crises (Beutel et al., 2019; Du Plessis, 2022a, 2022b; Wang et al., 2021). AUROC plots true and false positive rates at different cut-off points, where true positive rates (TPR) constitute the ratio of correct predictions to the combined total of both correct predictions (TP) and false negatives (FN), denoted $TPR = \frac{TP}{TP+FN}$, with outcomes from 0 to 1, the latter reflecting a flawless prediction. One constraint inherent in the AUROC measure is the requirement of a dependent variable with binary dimension. Resultantly, as the share of the world in crisis constitutes a continuous series with values below unity, only the GFC series, with two outcomes being a crisis or no-crisis, can be evaluated for predictive strength.

As third performance measure, a vector autoregression (VAR) is instated to assess the impact of an innovation in Wordscores and Wordfish on banking crises, and also employs the GFC series. Formally, the p th-order VAR can be stated as $y_t = c + A_1 y_{t-1} + \cdots + A_p y_{t-p} + e_t$, with A signifying either of the index generated by Wordscores and Wordfish, and p the lags. Orthogonal shocks are captured through a Cholesky decomposition of respective Wordscores and Wordfish in conjunction with the GFC series. The VAR is estimated with two lags for both Wordscores and Wordfish. The lag length is the median across the Akaike, Schwarz and Hannan-Quinn information criteria measures, as identified through VAR lag selection for Wordscores and Wordfish in Table 12 (Appendix 1). Furthermore, by applying a shock to the VAR system, the resultant change in the forecast error variance can be estimated, in order to determine the variation in the GFC series that can be explained by the Wordscores and Wordfish indices. Formally, the amount of forecast error variance due to innovations in variable j can be described by $w_{kj,h}$ where $w_{kj,h} = \frac{\sum_{i=0}^{h-1} \theta_{kj,i}^2}{MSE_k(h)}$ and $w_{kj,h} = \frac{\sum_{i=0}^{h-1} (\mathrm{e}_k\prime\Theta_i e_j)^2}{MSE_k(h)}$ with $\theta_{kj,i}$ the kjth element of Θ_i, e_k the kth column of I_K, and the forecast error variance of h-step ahead forecast of explanatory variable k denoted as $MSE_k(h) = \sum_{i=0}^{h-1} \mathrm{e}_k\prime\phi_i\Sigma_u\phi_i\prime e_k$ (Seyman, 2008).

4 Content Analysis

In preparing the textual data for incorporation into the two statistical models, content analysis is employed to construct the corpora into textual data, and to verify the relevance of the textual data. Involving a number of steps, it entails a verification of the textual data through hand-coding, so that contextual meaning and tone can be derived and validated for relevance. Indeed, textual data encounters a spatiotemporal challenge. The spatial dimension faced by an analysis of an article includes the possibility of a reference made to a different type of crisis such as a pandemic, as opposed to a banking crisis. A temporal dimension entails reporting on a historical crisis episode instead of the specific event under scrutiny. Furthermore, given the duration of the period under review, word variability and the introduction of new vernacular phrases could also alter the meaning of expressions. To limit the spatiotemporal challenges faced by text-mining, only news articles are used, which excludes opinion pieces and are generally centred on contemporaneous statistical releases, thereby restricting the temporal dimension mostly to present-day events. Secondly, news articles are constrained to economic events, which reduces the likelihood of references surfacing from non-economic sectors. Pre-processing techniques are further applied to ensure relevance of spatial information. To reduce the impact of word variability over time, the period is constrained to 2004 to 2012.

Therefore, to assess the preponderance of these dimensions still present in the textual data, and to limit its impact on the modelling framework, we thoroughly inspect a large sample of the textual data. In utilising the English language text source, but to balance analytical efficacy in scrutinising high dimensional corpora, a random sample of articles is systematically selected. For the period 2004 to 2012, the month of February constitutes the base for comparative analysis across the Thomson Reuters News Archive. While arbitrarily selected, for many banks the time frame represents a midpoint between reporting fourth quarter results and full-year earnings. For consistency, a random selection of 100 articles from each February month is extracted, inspected, analytical pre-processed and subsequently incorporated in the construction of the two content-based indices, namely Wordfish and Wordscores.

The dual-purpose of the content analysis is first, to ensure the relevance of the news articles to banking crises, and to limit spatiotemporal challenges, the latter subsequently verified through content inspection, and secondly, then to prepare the text for inclusion in the statistical models.

4.1 Content Inspection

Hand-coding is applied to confirm the meaning expressed by articles and the topics under discussion. This is achieved by the author studying and summarising the samples of articles and emphasising the salient themes, as reflective of the year from which it has been extracted. This analysis features three main categories, namely banking, real and external sectors. Through this approach, the drift across discourse is evident and focus is directed to literature-based causes of banking crises. A content summary of the randomly selected economic news article is contained in Appendix 2, which reflects growing weaknesses and early warnings from 2004 such as a heated real estate market, concerns about sub-prime lending, high levels of debt and rising global interest rates preceding the GFC. This serves to verify the relevance of the textual data to banking crises, which is subsequently employed to conduct pre-processing, an interim step to prepare the data for adoption by the statistical models.

4.2 Pre-Processing

Various research studies have shown how the type of words and the frequency of their occurrences can serve as gauge of the economic discourse or sentiment (Bholat et al., 2015). While the intuition behind this approach is straightforward, its implementation requires an extended degree of analytical pre-processing. In order to process the text for mining, a balance needs to be maintained between dimensionality reduction and analytical value. There are various modes of text pre-processing in the literature. The main aim is to prepare the text data for the model. Some words add little analytical value such as prepositions, and their removal ensure improved workings of the model, and a focus only on terms that can convey specific meaning, and that are relevant to the analysis. Although other modes could include topic modelling to surface the salient themes in a large corpus, this paper applies pre-processing so that the two statistical models Wordscores and Wordfish could be operationalised on the frequency distribution of word usage. While salient themes are useful for identifying the direction and content of discourse, and which are highlighted through pre-processing in the form of a term document matrix, the latter does not preclude or influence the systematic workings of the two statistical models.

This paper makes use of three main pre-processing techniques, namely removing stop words and applying case folding and stemming. Common words such as pronouns, articles and auxiliaries are excluded from the texts as it contributes limited analytical value. Numbers, standards of measures, mentions of time and date, and publisher details are omitted. Case folding is selectively employed, in order to account for common references such as 'Banking' and 'banking', yet to maintain a separation between proper nouns and collective nouns in the case of 'Central Bank' and 'bank'. Stemming allows the removal of affixes in order to derive the root form of a word where for instance 'bank' and 'banking' both refer to the same sector.

Before any pre-processing, 27,408 individual terms are detected across the full sample. At this stage, the term-document matrix features high-dimensionality and sparsity. To address the latter, and given that most words are used infrequently, only those featuring more than 0.5 percent within each annual sample are retained. This ensure that common terms mostly feature, while very infrequent words don't exert an undue influence on the outcome. Through pre-processing techniques, low-dimensional information is extracted, resultantly contained in a term-document matrix in Table 2. Most common terms include 'bank', 'rate' and 'market', with 'shares' and 'stock' prominent during the Global Financial Crisis. Given their relevance to the banking sector, these terms can be considered leading indicators for inclusion in the model development as well as for ongoing monitoring of news reporting based on these themes.

Visually, word clouds in Fig. 11 (Appendix 1) illustrate the evolutionary nature of key words over this period. With closer inspection and chronologically, interest rates dominate journalistic discourse in 2004 and 2005, before shares, stocks and reserves grow in prominence in the following two years. In particular, 2007 and 2008 emphasise themes on banking, in the same context as risk related adverbs such as 'rose' and'hit' and adjectives inclusive of 'higher'. Key phrases during 2009 revolve around recession, banking, and debt, followed by troublesome descriptors such as 'low' and 'issues' in 2010, with reporting the following two years on 'debt' and 'rates', depicted in the context of 'rise' and 'low'.

The following two sections incorporate the pre-processed textual data to build and operationalise the two statistical text-based indices.

Table 2 Term-Document Matrix

Top 20	2004	2005	2006	2007	2008	2009	2010	2011	2012
1	Bank	Rate	Shares	U.S.	Shares	Bank	Bank	Bank	Bank
2	Rate	Bank	Market	Market	High	Economy	Market	Market	Economy
3	Reserve	Up	Bank	Prices	Bank	U.S.	Rate	Rate	Market
4	Federal	Market	European	Economy	Price	Shares	Economy	U.S.	rate
5	Economy	Prices	Up	Rates	Up	Market	U.S.	Up	Inflation
6	U.S.	Foreign	Stocks	Bank	Market	Up	China	Prices	U.S.
7	Earnings	U.S.	U.S.	Inflation	Rise	Financial	Financial	Inflation	Policy
8	Market	Bonds	Rate	Up	European	Credit	Reserve	Oil	Central
9	Issues	Debt	Federal	Growth	U.S.	Rate	Up	China	Up
10	Dollar	Balance	Inflation	Credit	Euros	Federal	Federal	Policy	Credit
11	Policy	Interest	Reserve	Financial	Stocks	Policy	Inflation	Interest	Growth
12	Interest	Federal	Earnings	Dollar	Federal	Central	Central	Financial	Housing
13	Prices	Inflation	Prices	Interest	Inflation	Debt	Growth	Central	Financial
14	Monetary	Central	Euros	Shares	Earnings	Crisis	Investors	Down	Monetary
15	Trade	Trade	Economic	Federal	Ftse	Down	Policy	Investors	Debt
16	Balance	Reserve	Growth	Policy	Rate	Reserve	Credit	Global	Interest
17	Foreign	Shares	Investors	High	Outlook	Interest	Euro	Federal	Federal
18	House	Monetary	Rise	Oil	Reserve	Global	Interest	Reserve	Prices
19	Sales	Currency	Price	Fed	Investors	Inflation	Debt	Monetary	Reserve
20	European	Plans	Oil	Reserve	Down	Govern-ment	Govern-ment	Stocks	Data

5 Wordscores

Advanced by Laver et al. (2003), Wordscores represents an expedient mechanism to measure policy positions. By introducing this approach to the banking sector context, it can analogously serve to measure positions of appropriate articles related to their coverage between crisis and tranquil states. Positioned across two different poles, news articles containing crisis terminology are expected to be dissimilar in content to articles using non-crisis terminology. Intuitively, virgin texts are compared to reference texts with predetermined policy positions in order to score and locate the position of the new texts.

Starting with a set of reference texts R, policy positions on dimension d, are known or determined a priori as A_{rd}. The relative frequency of each word w in the reference text r can be computed as a proportion of the total number of words, formalised as F_{wr}. Following construction of F_{wr} for each of the reference texts, a matrix of relative word frequencies allows the computation of conditional probabilities. Accordingly, the matrix provides the probability of observing reference text r, given the appearance of word w, illustrated as $P_{wr} = \frac{F_{wr}}{\sum_r F_{wr}}$. In turn, the probability matrix can be used to assign a score for each word w, on dimension d, defined as

$S_{wd} = \sum_r (P_{wr} A_{rd})$. The latter score is an average of the known reference text scores A_{rd}, weighted by probabilities P_{wr}.

Following the computation of scores for all words in the reference texts, a set of virgin texts V can be analysed and compared. Correspondingly, the relative frequency of each word in the text as proportional to the total count of words, denoted by F_{wv}, is used in conjunction with the reference text score S_{wd} to compute S_{vd}, the latter an associated score of virgin text v on dimension d, expressed as $S_{vd} = \sum_w (F_{wv} S_{wd})$. This virgin text score represents the mean value of all its scored words, scaled by the frequency of the latter. Such scores are used to assign positions to virgin texts on the same dimension as the reference texts. In order to directly compare the virgin texts to the reference texts, scores from virgin texts are transformed to achieve the same dispersion metric. In $S^*_{vd} = \left(S_{vd} - S_{\widetilde{vd}}\right)\left(\frac{SD_{rd}}{SD_{vd}}\right) + S_{\widetilde{vd}}$, SD_{rd} and SD_{vd} are added as respectively the standard deviations of the reference and virgin texts, so that the mean of the virgin texts remains the same, but the variance equals that of the reference texts. As S_{vd} is a weighted score, variance of each word score V_{vd}, likewise can be weighted around the total text score by the frequency of the word in the text, stated as $V_{vd} = \sum_w F_{wv}(S_{wd} - S_{vd})^2$.

To construct a Wordscores index, the reference texts constitute the sample of combined news articles from the immediate pre-crisis period, spanning the years 2006 and 2007. For these reference texts, a priori scores are assigned as 1.00 for all articles from 2006 and 1.10 for articles from 2007. As the samples were taken from February of each year, it would still be considered pre-crisis as the GFC only commenced in August of 2007, and would therefore be reflected in the 2008 sample for the first time. Based on the fast growth in sub-prime mortgages and contemporaneous sharp increases in interest rates to slow the unsustainable pace of the housing market, 2006 takes a standard value of unity, followed by a slightly higher number the following year, concomitant with the higher risks expressed by the US Federal Reserve Bank, as highlighted through content inspection (in Appendix 2). Subsequently, the Wordscores methodology is operationalised to compute the scores of the virgin texts spanning the balance of the years between 2004 to 2012. Results highlighted in Fig. 1 demonstrate a sharp rise in the derived score in 2008, which could serve as leading indicator of ensuing weakness. Following a decrease in 2009 and moderation until 2012, the index remains in parallel to the share of countries in crisis during this period.

In a robustness test, the reference texts are varied, given the reliance of the virgin texts on the reference texts to determine their own positions on a left–right scale, and the partial subjectivity where scores are a priori allocated to these reference texts. Accordingly, the articles from two different periods are selected to serve as reference texts, namely the crisis period from 2008 to 2009 and the post-crisis period from 2010 to 2011. The only modification entails allocating the same scores of 1.0 for all articles in the sample making up the first reference text (2008) and 1.10 for the second reference text (2009), when the crisis period constitutes the selected reference texts, in line with a worsening in the period under review. In contrast, for the post-crisis period as reference texts, the first year of 2010 likewise assumes the value of 1.0, whereas

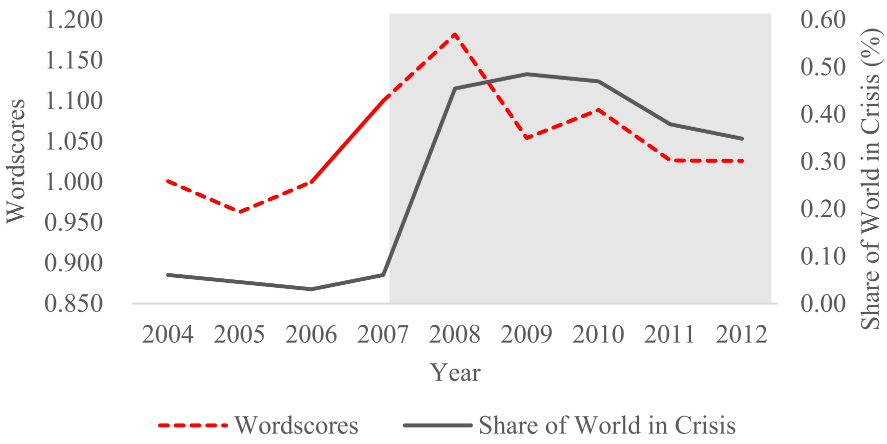


Fig. 1 Wordscores and Banking Crises

the second year (2011) takes the value of 0.9 as a recovery was observed through the content inspection (in Appendix 2). While the Wordscores approach allows a partially subjective input, it should still be adequately substantiated to ensure robustness of the model output. The initial pre-crisis Wordscores series serves as benchmark.

Highlighted in Fig. 2, the trajectory of the Wordscores series with the reference texts based on the crisis period signifies an initial increase in the index from 2004 to 2006, which would correspond to the subprime crisis and build-up of vulnerabilities, and following a short dip in 2007, the series reverts to higher levels during the subsequent years, in line with the increase in countries experiencing a banking crisis, yet spiking around 2012 with the onset of the European sovereign debt crisis.

The post-crisis series as reference texts initially increases from 2004 to 2005, before another gradual increase from 2006 to 2008 with a sharp spike in 2009 during the crisis, and subsequently declines until 2011. While the selection of the reference texts should be based on sound economic judgement, and informed by empirical literature, the output from picking different starting points as reference texts shown in this analysis, demonstrates comparable trajectories of the derived Wordscores index, the latter highlighting consistency and robustness.

To verify the strength of the Wordscores series in signalling a banking crisis, two Granger causality approaches are implemented. The first utilises the Wordscores index in conjunction with the share-of-world-in-crisis series, the latter increasing until the GFC and remaining elevated thereafter. Here the focus is on the degree to which Wordscores can signal the propagation of the crises globally. The second Granger causality approach applies the GFC as a global systemic crisis event to forecast.

Summarised in Table 3, Wordscores Granger causes both one and two lags of the share-of-world-in-crisis series. This confirms its strength as leading signal, one to two years in advance of a crisis. At neither lag length does the crisis series Granger cause Wordscores. More than two lags are not possible to operationalise given the limited sample size, which also precludes disaggregating the period under review into separate

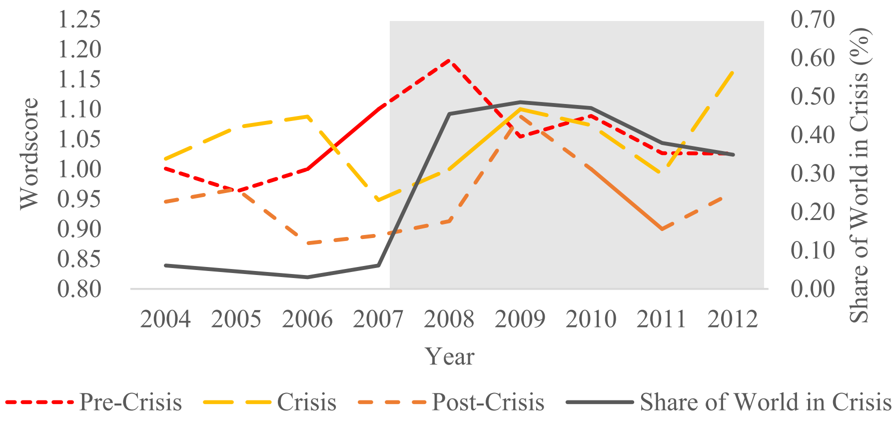


Fig. 2 Wordscores and Banking Crises: Variation in Selection of Reference Texts

time periods for further evaluation. However, by employing a lagged correlation coefficient, confined crisis phases can further be scrutinised in the following analysis.

So, to further examine the strength of the Wordscores series in correlating with the share-of-world-in-crisis series, Table 4 encapsulates the cumulative correlation coefficient of the Wordscores index with a one-year lag, and shows a strong relationship with the crisis indicator pre-crisis (93 percent) and during the crisis period (96 percent) when the number of countries experiencing a crisis peaks. In a separate analysis, Wordscores with a two-year lag is shown to remain near 70 percent across most periods, denoting higher strength over a one-year period.

For the second Granger causality approach, Wordscores Granger causes the GFC series at one lag length, whereas the GFC series is not shown to Granger cause the Wordscores series, as shown in Table 5, thereby affirming the direction of causality from the Wordscores index. The Granger causality results from both the share-of-world-in-crisis series and GFC series confirm the leading indicator strength of Wordscores.

To evaluate predictive strength, receiver operating characteristics with area under the curve estimates are used. Highlighted in Table 6, Wordscores has a 94.4 percent mean accuracy score. ROC curve in Fig. 3 further demonstrates high levels of sensitivity or correctly predicting crisis and non-crisis events, and specificity, so not sending false signals.

Table 3 Wordscores and World in Banking Crisis: Granger Causality

Lag	Wordscores	World-in-Crisis
1	5.633*	0.988
2	106.220***	4.469

*** (**, *) denotes significance at 1%, (5%, 10%)

Table 4 Wordscores and Banking Crisis: Correlation Coefficient

		Crisis	
		Pre-crisis until start of Crisis (2006–2008)	Pre-crisis until end of Crisis (2006–2009)
Wordscores	Pre-crisis (2004–2006)	0.939	
	Pre-crisis (2004–2007)		0.969

Table 5 Wordscores and Global Financial Crisis: Granger Causality

Lag	Wordscores	GFC
1	3.891*	2.114
2	3.132	0.062

***(**, *) denotes significance at 1%, (5%, 10%)

Table 6 AUROC Results for Wordscores and GFC

AUROC Results	Mean	CI lower bound	CI upper bound	Standard error
Wordscores	0.944	0.790	1.000	0.078

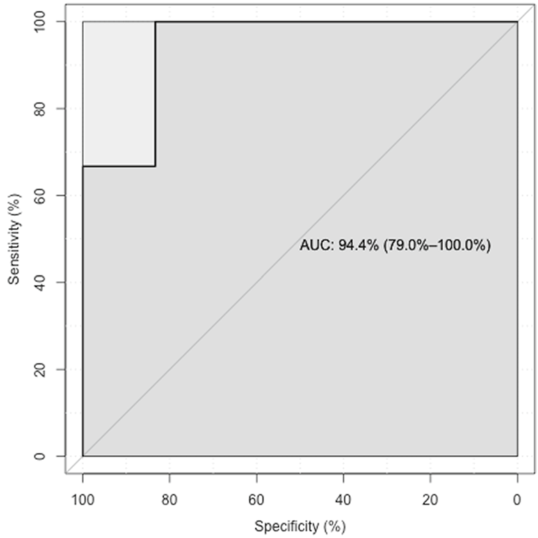
Upper and lower bounds are based on 95 percent confidence intervals. Variance of AUC is defined by DeLong et al. (1988) and estimated with algorithm specified by Sun and Xu (2014)

A vector autoregression (VAR) comprising Wordscores and the GFC series is constructed in Table 13 (Appendix 1), and shows an R-squared of 83 percent, highlighting the large share of variation in the crisis series explained by Wordscores. Impulse responses captured through Cholesky decomposition further demonstrate the direction and strength of influence between the two series. Illustrated in Fig. 4, with a one standard deviation band of observations, an innovation in the Wordscores indicator exhibits a sharp rise in the GFC series which peaks after three years, and reverts to the zero bound after five years. This emphasises the strength of Wordscores in signalling an impending crisis. The forecast error variance in Fig. 5 emphasises the forward-looking nature of Wordcores, with the latter index rising in influence and explaining around 80 percent of the variation in the GFC series from three years onwards.

6 Wordfish

Whereas Wordscores can be classified as supervised learning, Wordfish would be deemed unsupervised learning. In contrast to Wordscores, there are no reference texts. Developed by Proksch and Slapin (2008), Wordfish is a practical quantitative content analyser, which operationalises a statistical scaling model to estimate policy positions of texts through relative word usage. Benefits of this approach include the generation of time-series estimates, which is appropriate to monitor banking sector developments over a period of time. Moreover, this methodology is not bounded by a requirement for a priori policy dimensions or reference texts as it implements

Fig. 3 ROC Curve for Wordscores and GFC



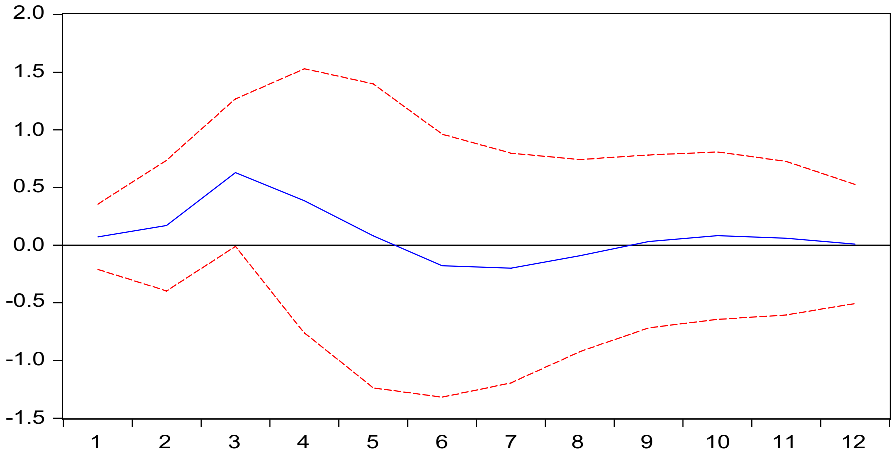


Fig. 4 Wordscores: Impulse Responses with One Standard Deviation Band

a statistical distribution of word frequencies. This could also be a shortfall of the Wordfish approach in that expert judgement is not considered in the model. Yet, it overcomes subjective judgments. Through this process, all words across every document are used and their importance estimated. A discrete probability Poisson distribution is applied to estimate the likelihood of an event occurring, which is commonly employed to model banking crises (Eichengreen, 2001; Gresnigt et al., 2014).

The model is assumed to follow a Poisson distribution, where only one parameter is estimated, serving as both mean and variance, and formally stated as:

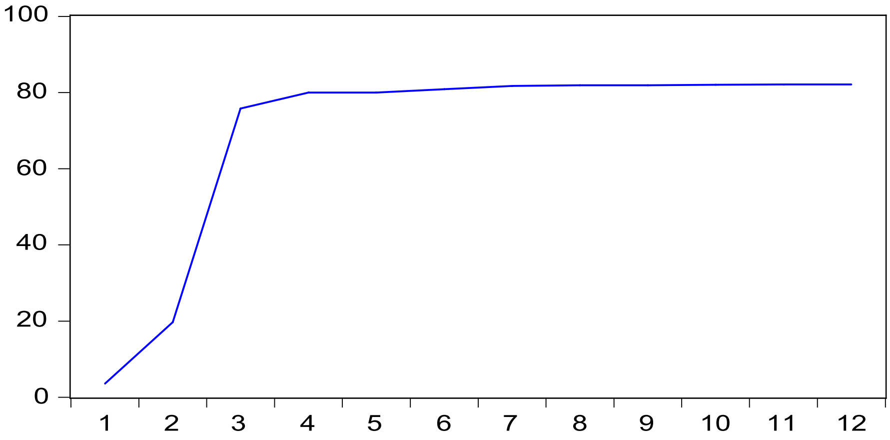


Fig. 5 Wordscores: Forecast Error Variance

$$Wordcount_{ij} \sim Poisson(\lambda_{ij})$$

$$\lambda_{ij} = \exp(\alpha_i + \psi_j + \beta_j \omega_i)$$

Wordcount is the count of word j in text i, with α a set of document fixed effects, ψ as word fixed effects, β represents an estimated word specific weight to highlight the importance of word j in distinguishing between positions, and ω is an estimate of document i's position. Word fixed effects take into account the relative intensity in word usage between documents, while document fixed effects control for differences in length of articles. An expectation maximisation algorithm is employed to compute the parameters of the model, through iteratively oscillating between deriving the word-specific parameters by maintaining the document parameters fixed, and in turn estimating the document specific parameters while keeping the word specific parameters unchanged. Convergence is reached when the log-likelihoods cease to change between iterations. The likelihood function is estimated by scaling the dimensions of all documents to a mean of zero and standard deviation of one. Resultant policy positions are subsequently located on a standardised dimension.

Illustrated in the shape of an Eiffel Tower, Fig. 6 plots the estimated word effects against their word weights. Accordingly, words with high fixed effects such as 'economy', 'financial' and 'banking', have low weights, whereas words with low fixed effects have negative weights for instance 'bubbles', 'pre-crisis', 'decrease' and 'over-leveraged' or positive weights such as 'increase', 'strengthened', 'recovery' and 'rebounding'.

This paper introduces Wordfish to model and predict banking crises. Intuitively, policy positions are akin to the $y = 0$ (no-crisis) and $y = 1$ (crisis) spectrum of banking sector distress. Thereby, the aim is to scrutinise if the model can detect a change in terminology between crisis and non-crisis articles. The period under review spans from 2004 to 2012, by using the same annually extracted text data samples. Illustrated in Fig. 7, the Wordfish index rises sharply from 2004 to 2006, to indicate

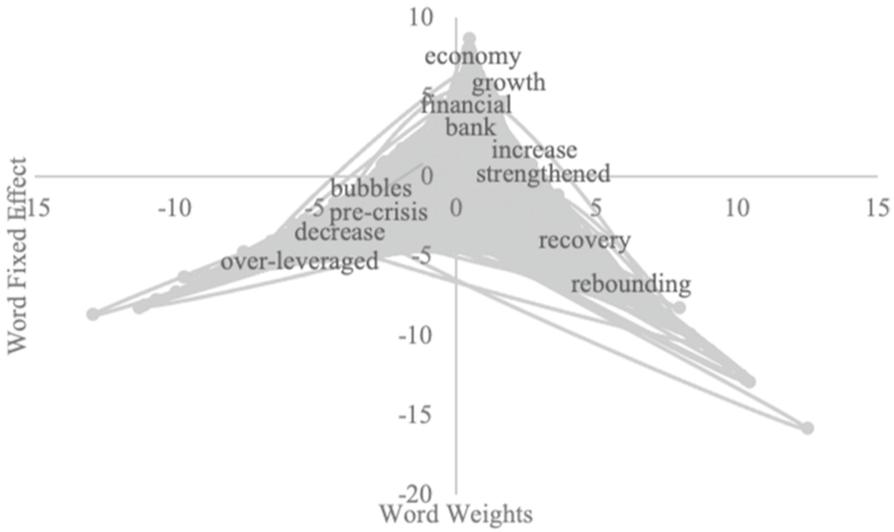


Fig. 6 Wordfish: Word Weights and Fixed Effects

ensuing pressure in the lead up to the sub-prime mortgage crisis, followed by a second spike from 2007 to 2008 to signal the subsequent Global Financial Crisis.

The Granger causality results for the share-of-world-in-crisis series are shown in Table 7. The world-in-crisis series is found to Granger cause Wordfish at one and two lag lengths. From the graphical assessment, this observation could be a result of the more volatile nature of the Wordfish series, which appears to be more reactive to the crisis indicator during the second half of the period under review.

To further examine this phenomenon, the cumulative correlation coefficients of Wordfish and the share-of-world-in-crisis series are summarised using a two-year lag for Wordfish in Table 8. The one-year lag was also considered, but given the volatility in the trajectory of the series, two-years presents a more stable period. According to the results, the first three years of the sample, and which corresponds to the pre-crisis phase exhibits a 91 percent correlation with the share of countries experiencing a banking crisis. While gradually reducing over time, the relationship strength remains above 60 percent during the crisis period. The results strengthen the applicability of Wordfish as early warning signal, by providing a two-year lead time during a period of heightened concern, albeit tapering off towards the end of the crisis period. While Granger causality is not from the direction of the Wordfish series to the share-of-world-in-crisis series across the period under review, the staggered correlation coefficient nonetheless highlights the sensitivity of the statistical indicator to signal an early warning in the lead up to the crisis.

Table 9 features Wordfish and the GFC series. In forecasting the GFC series, Wordfish is shown to Granger cause the GFC series at one lag length. The

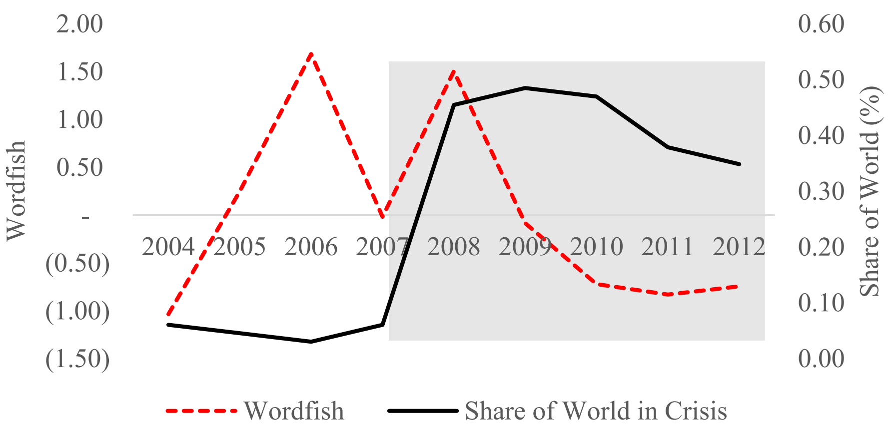


Fig. 7 Wordfish and Banking Crises

Table 7 Wordfish and World in Banking Crisis: Granger Causality

Lag	Wordfish	World-in-Crisis
1	0.052	6.809**
2	5.896	14.185*

*** (**, *) denotes significance at 1%, (5%, 10%)

Table 8 Wordfish and World in Banking Crisis: Correlation Coefficient

		Crisis	
		Pre-crisis until start of Crisis (2006–2008)	Pre-crisis until end of Crisis (2006–2009)
Wordfish	Pre-crisis (2004–2006)	0.910	
	Pre-crisis (2004–2007)		0.622

direction of causality does not flow from the GFC series to Wordfish. The opposite results to the share-of-world-in-crisis series could be due to the volatility in the Wordfish indicator which tends to converge to upper and lower bounds, rather than stabilising between these opposing poles.

AUROC results are summarised in Table 10, and ROC curve plotted in Fig. 8. Accordingly, Wordfish reaches a mean accuracy score of 72 percent over the sample, which spans 2004 to 2012. The spread between the confidence intervals of the lower bound and upper bound encapsulates the more volatile nature of the Wordfish series, which could further be ascribed to its role in signalling two events, first, the sub-prime crisis, and second, the Global Financial Crisis. Weaknesses inherent in

Table 9 Wordfish and Global Financial Crisis: Granger Causality

Lag	Wordfish	GFC
1	6.546**	0.008
2	1.824	1.018

*** (**, *) denotes significance at 1%, (5%, 10%)

the U.S. subprime housing market, observable from 2004, gradually evolved into a financial crisis within the subsequent three to four years.

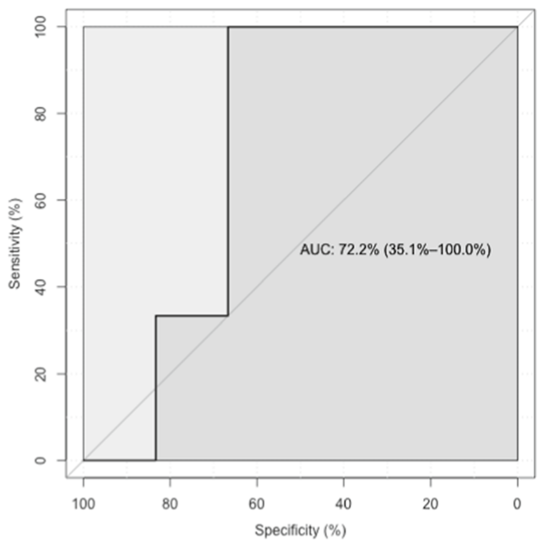
VAR results as shown in Table 14 (Appendix 1), with R-squared of 75 percent underscore the value of Wordfish in explaining the crisis trajectory. Figure 9 plots the impulse response with Cholesky decomposition of a VAR model based on Wordfish and the GFC series. Initially, a shock to Wordfish results in a decrease in the GFC series, turning positive after four to five years, remaining elevated until year nine. This trajectory could be due to the influence of the sub-prime mortgage crisis. Figure 10 highlights the forecast error variance, showing Wordfish to explain around 60 percent of the movements in the GFC series.

Table 10 AUROC Results for Wordfish and GFC

AUROC Results	Mean	CI lower bound	CI upper bound	Standard error
Wordfish	0.722	0.351	1.000	0.189

Upper and lower bounds are based on 95 percent confidence intervals. Variance of AUC is defined by DeLong et al. (1988) and estimated with algorithm specified by Sun and Xu (2014)

Fig. 8 ROC Curve for Wordfish and GFC



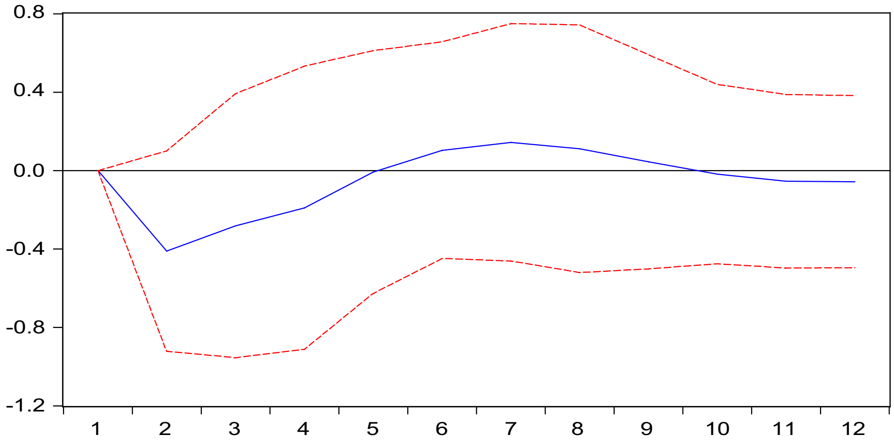


Fig. 9 Wordfish: Impulse Responses with One Standard Deviation Band

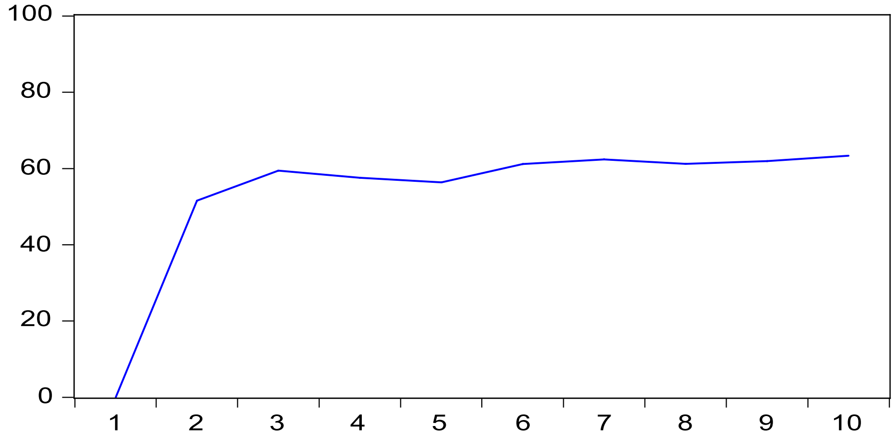


Fig. 10 Wordfish: Forecast Error Variance

7 Comparative Performance

Based on the graphical assessment, Wordscores spikes with a two-year lead time, allowing an adequate response period for policy makers to intervene. In a disaggregated analysis using the cumulative one-year correlation coefficient, Wordscores exhibits a solid relationship to the crisis indicator during periods of imminent banking distress, which would be a valuable tool for policy makers needing to anticipate and prevent forthcoming catastrophes. Granger causality further demonstrates the value of Wordscores in predicting the share-of-world-in-crisis series, with one and two lag lengths, and the GFC series with one lag length.

The initial early warning signal provided by Wordfish with a two-year lag is one of its advantages, which becomes more apparent through a medium-term correlation coefficient when deconstructing the period under review, but its higher degree of volatility impacts interpretability, and results in an opposite direction for Granger causality across the share-of-world-in-crisis series. Yet, Wordfish Granger causes the GFC series at one lag length, and potentially fair better explaining movements in a binary dependent variable, particularly if convergence to higher and upper bounds are more important. Notwithstanding, Wordfish is a series more sensitive to changes in economic discourse, which could be useful for studying sudden shocks such as a looming banking crisis, and also represents an objective approach to analyse the underlying textual data.

Overall, Wordscores reaches 94 percent accuracy against Wordfish on 72 percent, and Wordscores explains the variation in the GFC crisis series at 80 percent compared to Wordfish on 60 percent. VAR results are more comparable with R-squared for Wordscores on 83 percent to Wordfish on 75 percent. While comparatively lower influence, however Wordfish signals both the sub-prime crisis and GFC, which holds benefits for predicting related and diverse economic environments.

In terms of the comparison between supervised and unsupervised models, the slightly stronger performance of the former highlights the value of a hybrid learning approach where models are validated and enhanced through expert input, instead of a standalone operation.

8 Robustness Test

Random forests algorithm is a popular approach to classify text (Fernández-Delgado et al., 2014). In a robustness test, the same textual dataset used for Wordscores and Wordfish is implemented using the random forests algorithm. Essentially the same pre-processing techniques are employed to generate a term-document matrix, after which the random forests algorithm is trained on a subset of the term document matrix, with the remaining part of the term-document matrix held out for new predictions to be done by the algorithm. During training, the random forests algorithm analyses the features of the term document matrix that best explain the outcome classes. This is done by constructing several decision trees using a random selection of words from the term document matrix as based on their frequency of use and given the most optimal split amongst those words. Every decision tree is used for predictions, and these predictions are then averaged to determine the final prediction model, in turn employed to predict new observations.

One modelling limitation is that both classes of outcomes, namely crisis and non-crisis episodes are required for the algorithm to be operationalised. Given that there are only two crisis years in the sample from 2004 to 2012, a recursive out-of-sample approach would be limited by sample size as the training data spans until the first crisis year of 2008 and then leaving only one year, 2009 available to assess an actual crisis prediction. While more data points follow until 2012, the latter would instead be used to gauge false signals. Summarised in Table 15 (Appendix 1), predictions generated based on this approach indicate all four remaining years to be non-crisis periods, thereby scoring a 75 percent overall accuracy rate, albeit missing the crisis episode. Subsequently, native machine learning forecasting are employed where 50 percent of the sample is randomly allocated to the training set and 50 percent to the testing set. While it does not follow a timeseries dimension, observations are taken from across the sample and thereby provide broad exposure to varying environments. Accuracy rates computed as the share of correct predictions, spanning both crisis and non-crisis periods, expressed as a share of all predictions likewise amount to 75 percent out-of-bag rate. Both recursive out-of-sample and native forecasting predictions equate to a mean AUROC accuracy rate of 50 percent. AUROC results are comparatively lower compared to the 72 percent for Wordfish and 94 percent for Wordscores.

9 Policy Approach

The two statistical models Wordscores and Wordfish both convey early warning signals in the lead up to the Global Financial Crisis. As a supervised method, Wordscores requires a degree of initial subjectivity to implement, which could have an effect on the classification of the new texts. Yet, a robustness test demonstrated comparable outcomes when varying the reference texts. For a policy maker to rely exclusively on a model for direction might not always be preferable, and the ability to overlay human intelligence would make a stronger case for adoption, calibration and ongoing improvement as based on new economic findings and developments. In contrast, the Wordfish approach is unsupervised, and it estimates the positions of discourse on banking sector weaknesses independently of expert opinion. As a result, it could be considered a more objective mechanism to monitor banking distress. A key benefit of the Wordfish approach is that results are shown to remain robust in different languages, across national and regional texts, and irrespective of distributional assumptions and document selection (Proksch & Slapin, 2008, 2009).

Through the term-document matrix, several leading indicators are highlighted that could be used for ongoing monitoring of reporting on such topics, where themes around debt, interest rates, and stock market become more associated with deteriorating conditions, exemplified by rising levels in these leading indicators, such as unsustainable credit growth that could bring about more pressure, or defaults resulting in weaker bank earnings.

Although the prediction of the Global Financial Crisis is the main objective of this study, as more digital content is generated and becomes available, tools such as Wordscores and Wordfish, commonly used in the political domain, could become more frequently utilised for monitoring the stability, risks and resilience of financial markets, and banking crises.

10 Conclusion

The growth in digital news feeds constitutes rich data sources to elucidate economic phenomena, in general, and banking crises, specifically. In this context, the paper makes a number of contributions, firstly, through the introduction of statistical text-based models to scrutinise banking sector weaknesses; and secondly, through the use of a rich textual data source with over 19 million articles, which has been sparsely studied to date; and thirdly, in an international study spanning the Global Financial Crisis, and with generalisable outcomes. Fourthly, supervised and unsupervised learning methods are compared to determine fit for purpose modelling approaches. Fifthly, this study focusses on banking crises, and specifically the prediction of banking crises. The aim is to improve the toolkit available to policy makers to institute preventative measures in curbing the impact of a banking crisis.

Through annually extracted samples of textual data, based on precursors to banking crises, and following pre-processing techniques, two banking crisis indices are constructed by analysing the content of articles and intensity of word usages. In comparing virgin texts to reference texts, with a priori positions, the Wordscores index indicates heightened usage of banking sector nomenclature two years preceding a crisis. Using a statistical model with Poisson distribution, the Wordfish index signifies spikes before both the sub-prime housing crisis and Global Financial Crisis. Through correlation coefficient measures, Wordfish provide a strong initial early warning signal, while Wordscores predicts the expansion of crises to other countries more consistently. Both Wordscores and Wordfish are shown to Granger cause a GFC series. Content analysis verifies the veracity of the indices across the period under investigation. Both statistical models outperform the commonly used random forests algorithm and make a contribution to the detection of banking sector weaknesses, and provide policy makers an objective and partially subjective method to reveal the trajectory of banking distress through text-based models. Future studies could consider country-specific models to further reveal idiosyncratic nuances and risk factors on a more granular level, as well as newsfeeds in the vernacular where English-based reporting remains limited.

Appendix 1

See Fig. 11 and Tables 11, 12, 13, 14, 15.

Fig. 11 Word Cloud Annual Samples

Table 11 Banking Crisis Lexicon

Banking sector	Real sector	External sector
Bank or banking and	Consum or invest or product	Export or import or trade or terms of trade
Deposit or credit or debt,		
Interest rate or		
Inflation or cpi,		
Reserve or gold,		
Liquid or contract or eas or tight or monetary or boom or bust or crisis,		
Fraud or earning or hous		

Table 12 VAR Lag Order Selection Criteria

Lag	LogL	LR	FPE	AIC	SC	HQ
Wordscore: Optimal Lag Length						
0	4.162968	NA*	0.001855	−0.617991	−0.633445	−0.809002
1	9.209216	5.767141	0.001542*	−0.916919	−0.963282	−1.489954
2	13.272840	2.322069	0.002782	−0.935096*	−1.012368*	−1.890154*
Wordfish: Optimal Lag Length						
1	−11.81021	NA	0.324382*	4.517204*	4.486296*	4.135181*
2	−9.561063	1.927844	0.707898	5.017447	4.955630	4.253401

*indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

Table 13 Wordscores: VAR Results

Lag	Y = Crisis
Wordscores (-1)	2.070424
(3.31962)	
[0.62369]	
Wordscores (-2)	6.116339
(3.15936)	
[1.93594]	
Crisis (-1)	− 0.042873
(0.41665)	
[-0.10290]	
Crisis (-2)	0.063935
(0.42328)	
[0.15105]	
C	− 8.321041
(3.75484)	
[−2.21609]	
R-Squared	0.833919

Standard errors in () & t-statistics in []

Table 14 Wordfish: VAR Results

Lag	Y = Crisis
Wordscores (−1)	− 0.375736
(0.19697)	
[− 1.90761]	
Wordscores (− 2)	− 0.132834
(0.28253)	
[− 0.47016]	
Crisis (− 1)	0.207503
(0.60086)	
[0.34534]	
Crisis (−2)	− 0.295693
(0.39244)	
[−0.75347]	
C	0.341030
(0.26535)	
[1.28523]	
R-Squared	0.757008

Standard errors in () & t-statistics in []

Table 15 Random Forests and Global Financial Crisis: Robustness Test

	Recursive out-of-sample	Out-of-bag
Accuracy rate	0.750	0.750
AUROC	0.500	0.500

Appendix 2

Content Inspection Summary

As extracted from Thomson Reuters News Archive (2020), and summarised by the author with reference to the real, banking and external sectors, accordingly in the 2004 sample of articles, many countries embarked on a tightening interest rate cycle, with exception of countries in the European Union. Policy makers became concerned that lax monetary policy was encouraging a consumer borrowing boom, particularly against property, that could leave the economy unbalanced and vulnerable. Higher credit growth in the housing market was responsible for inflationary pressures. Real sector experienced maturing investments and a cooling economy, while currencies were appreciating, gold price remained low and imports cheap as applicable to the external sector.

Banking sector in 2005 observed lower bank profit, with concerns about inflation, high debt, strong credit, and central bank tightening, which was down from double digit housing debt growth a year ago. Japan reduced guarantees for bank deposits. In the real sector, world growth was easing, with rising concerns about jobs. Exports and investments remained strong in the United States, while Asia continued with higher investments in infrastructure. Externally, negative net imports led to growing trade deficits, the gold price was on the increase and oil prices started to escalate.

The follow year coincided with a slowdown in the housing market, high global interest rates and tighter liquidity. Fast credit growth featured in emerging markets such as 20.4 percent in South Africa, mostly due to high mortgage growth of 27.4 percent, and accompanied by high inflation. A consumer spending spree fueled greater demand for credit. In the financial markets, bond future contracts increased, while debt default swaps tightened, in the context of slightly elevated risk sentiment towards shares. As such, investors showed appetite for bonds, with 36 percent of investors investing in credit default swops for the first time. With expectation that subordinated bank debt would outperform other investments, investors were selling put options, worldwide demand for top-rated government bonds increased, and shares in Europe traded near a four-and-a-half-year record. In the real economy, higher consumption was observed in emerging markets such as Turkey and China. Risks mentioned which relate to the external sector include speculation of a hard landing for China's economic growth, in particular due to a slowdown from higher oil prices, bird flu, collapse in consumption, and dependence on net exports.

In 2007, a significant amount of coverage was dedicated to the banking sector. Specifically, the Senate Banking Committee in the United States was drawing up regulation to guide banks on underwriting non-traditional mortgages by including loans made to subprime borrowers. This came at a time when rising delinquencies have reverberated throughout U.S. financial markets. A growing number of subprime borrowers faced foreclosure and their lenders faced insolvency. Subprime loans are so-called '2–28 mortgages' which feature a fixed resetting as much as 6 percentage points higher. Fremont General Corporation, the fifth-largest U.S. subprime lender, closed down 24.5 percent after the company postponed its release of fourth-quarter results and its annual report. In China, the central bank was expected to continue raising banks' required reserves to consolidate excess cash generated by an outsized external payment surplus. Fears of an overheating economy due to fast growth in investment, credit and money, slowed, yet retail sales have been robust. Exports remained strong and corporate profits have been rising quickly. Higher interest rates coincided with a nine percent drop in the Shanghai stock market. In contrast, European stocks were moderating following a six-year high. Banking and corporate earnings were slowing. Related to the real sector, a warning from former Federal Reserve Chairman Alan Greenspan that the U.S. could be in recession by the end of the year has driven a global stock market sell-off. Together with downward revisions to fourth quarter gross domestic product, a rally in the United States government bond market was observed. In the external sector, gold was nearing a doubling over the recent two years and oil prices were rising.

Beginning 2008, banking failures were anticipated. According to Federal Reserve Chairman Ben Bernanke, some small U.S. banks could fail from the financial stress of housing market troubles and mortgage losses. A comment which drove stock prices lower. U.S. Treasury prices surged as investors fled for safety after signals that the economy may be descending into a recession, and fears that the housing sector was still worsening. The deep housing market downturn and tightening in the availability of credit were driving the economy into a recession. Based on a statement from the Federal Reserve, the central bank would act as required to ensure the housing and credit markets do not further undermine the subdued U.S. economy. Markets interpreted this move as a growing likelihood of an interest rate cut, which weakened the U.S. dollar. The Federal Reserve emphasised the importance for policy makers and the mortgage industry to find long-term solutions. In emerging markets, rising food and energy prices were expected to force their central banks to raise interest rates. In the context of the real sector, major economies were expected to further slow, coupled with higher inflation. In the United States, first quarter economic growth was stagnating, while growth in the eurozone was anticipated to slow sharply, with prospects of a recession. Likewise, Japan's economy was slowing and no increase in interest rates were anticipated, with sluggish UK growth expected. Externally, the flight to safety resulted in record high gold prices.

In 2009, high interest rates and tight market liquidity were providing a complicated prognosis about the economic recovery. In countries with low interest rates, consumer spending and credit could take a long time to rebound. To stimulate economy activity, government spending plans were aimed at replacing private spending, and through monetary expansion to boost credit markets. Deliberations focused on the financial crisis and investigations started to uncover fraudulent activity. Wall Street shrunk to a 12-year low with uncertainty about the efforts of the government to stabilise the financial system. Related to the real sector, economic contractions, large budget deficits, government intervention to buy toxic assets, and bankruptcy of corporates comprise the main themes. Regarding the external sector, exports declined for many countries, including Japan and Germany, while trade protectionism rose. Gold prices remained elevated after reaching an 11-month high above $1000.

The following year, banking sector topics involved loose monetary policy, and the investigation of irregularities regarding derivative contracts, and role of credit ratings agencies. High debt levels in Greece featured as consideration for necessary reforms. Lower demand for credit, softer inflation, a poor housing market, and consumers struggling to repay loans impacted many countries. Articles on the real economy expressed wildly mixed messages about the pace of the recovery, while the jobless rate was rising. In the external sector, the gold price set a new record.

The focus in 2011 shifted to the banking bailouts in Europe including from the IMF, tougher banking regulations and identification of systemically important banks. World renowned investor Warren Buffet disclosed an intention to increase share holdings. Monetary tightening was expected in China and Europe, due to higher fuel prices and inflationary pressures. Despite a 2010 rebound, global growth was slowing in the real sector. Less affected by the Global Financial Crisis, China's growth increased given a massive appetite for raw materials, leading to higher gold production impacting the external sector.

In the last year under scrutiny, 2012 featured quantitative easing programmes in the United States, Britain, and Japan, through increasing the money supply by buying securities from the market. European Central Bank was committed to support banks in providing liquidity to the corporate sector through accommodative monetary policy and were considering quantitative easing as intervention. In Europe, sovereign debt problems came to the fore together with a credit crunch. Bailout agreements and renegotiation featured in particular for Greece and Portugal. With a reduction in interest rates on the Greek bailout, pressure was rising in Portugal to renegotiate the terms of their bail-out agreement in order to ease the impact of austerity measures during a recession. Bank earnings were under pressure in the United Kingdom, while consumers were under-spending in the United States. China's central bank reduced banks' required reserve ratio in order to increase liquidity in the economy, given a slowdown in annual housing inflation. Interest rates were anticipated to remain low. In the real economy, surging oil prices and improving U.S. jobs figures provided a more optimistic outlook in North America in contrast to the eurozone. In China, growth was slowing with expectations for the weakest full-year expansion in a decade. In relation to the external sector, countries were increasing international reserves by buying the U.S. dollar, while higher gold prices were expected to dampen the demand for gold imports during the year.

Author contributions The author contributed to the study conception and design. Material preparation, data collection and analysis were performed by the author. The first draft of the manuscript was written by the author. The author read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

References

Acosta, M. (2015). FOMC responses to calls for transparency. *Finance and Economics Discussion Series, 2015*(60). Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2015.060.

Alessi, L., Antunes, A., Babecký, J., Baltussen, S., Behn, M., Bonfim, D., Bush, O., Detken, C., Frost, J., Guimarães, R., Havránek, T., Joy, M., Kauko, K., Matějů, J., Monteiro, N., Neudorfer, B., Peltonen, T., Rodrigues, P.M.M., Rusnák, M., Schudel, W., Sigmund, M., Stremmel, H., Šmídková, K., Van Tilburg, R., Vašíček, B., & Žigraiová, D. (2015). Comparing different early warning systems: Results from a horse race competition among members of the macro-prudential research network. *MPRA Paper*, *62194*.

Archive, T. R. N. (2020). *Thomson Reuters News Archive*. Reuters News & Media Inc.

Benoit, K., & Laver, M. (2003). Estimating Irish party policy positions using computer Wordscoring: The 2002 election–a research note. *Irish Political Studies, 18*(1), 97–107. https://doi.org/10.1080/07907180312331293249

Benoit, K., Laver, M., Arnold, C., Pennings, P., & Hosli, M. O. (2005). Measuring national delegate positions at the convention on the future of europe using computerized word scoring. *European Union Politics, 6*, 291–313.

Berge, T. J., & Jordà, O. (2011). Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal, Macroeconomics, 3*(2), 246–277.

Beutel, J., List, S., & Von Schweinitz, G. (2019). Does machine learning help us predict banking crises? *Journal of Financial Stability, 45*, 100693.

Bholat, D., Hansen, S., Santos, P., & Schonhardt-Bailey, C. (2015). Text mining for central banks. *CCBS Handbook, 33*. Bank of England.

Budge, I. (1987). The internal analysis of election programmes, in I. Budge, D. Robertson and D. J. Hearl (eds), *Ideology, strategy and party change: Spatial analyses of post-war election programmes in 19 democracies*. Cambridge: Cambridge University Press.

Calomiris C.W., & Mason, J. (2003). Fundamentals, panics and bank distress during the depression. *American Economic Review*, *93*.

Calomiris, C. W., & Mamaysky, H. (2018). How news and its context drive risk and returns around the world. *NBER Working Paper*, *24430*.

Chen, M., DeHaven, M., Kitschelt, I., Lee, S. J., & Sicilian, M. J. (2023). Identifying financial crises using machine learning on textual data. *Journal of Risk and Financial Management*, *16*(161).

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391–407.

DeLong, E. R., DeLong, D. M. and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics, 44*, 837–845.

Demirgüç-Kunt, A., & Detregiache, E. (1998). The determinants of banking crises in developing and developed countries. *IMF Staff Papers, 45*(1).

Döpke, J., Fritsche, U., & Pierdzioch, C. (2017). Predicting recessions with boosted regression trees. *International Journal of Forecasting, 33*, 745–759.

Drees, B., & Pazarbasioglu, C. (1998). The Nordic banking crisis: Pitfalls in financial liberalization. *International Monetary Fund Occasional Paper*, *161*.

Du Plessis, E. (2022a). Multinomial modeling methods: Predicting four decades of international banking crises. *Economic Systems, 46*(2), 100979. https://doi.org/10.1016/j.ecosys.2022.100979

Du Plessis, E. (2022b). Dynamic forecasting of banking crises with a Qual VAR. *Journal of Applied Economics, 25*(1), 477–503. https://doi.org/10.1080/15140326.2020.1816132

Egerod, B. C. K., & Klemmensen, R. (2019). Scaling political positions from text. *The SAGE Handbook of Research Methods in Political Science and International Relations*, Chapter 27.
Eichengreen, B. (2001). International financial crises: Is the problem growing? *Jahrbuch für Wirtschaftsgeschichte*, *43*(1).
Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research, 15*(1), 3133–3181.
García, D. (2013). Sentiment during recessions. *The Journal of Finance, 68*(3), 1267–1300. https://doi.org/10.1111/jofi.12027
Giannetti, D., & Laver, M. (2005). Policy positions and jobs in the government. *European Journal of Political Research, 44*, 91–120.
González-Fernández, M., & González-Velasco, C. (2020). An alternative approach to predicting bank credit risk in Europe with Google data. *Finance Research Letters*, *35*.
González-Hermosillo, B., Pazarbasioglu, C., & Billings, R. (1997). Determinants of banking sector fragility: A Case Study of Mexico. *IMF Staff Papers, 44*(3). Washington: International Monetary Fund.
Gorton, G. (1988). Banking panics and business cycles. *Oxford Economic Papers*, *40.*
Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross- spectral methods. *Econometrica, 37*(3), 424–438.
Gresnigt, F., Kole, E., & Franses, P. H. (2014). Interpreting financial market crashes as earthquakes: A new early warning system for medium term crashes. *Tinbergen Institute Discussion Paper, 2014* (067).
Hansen, S., McMahon, M., & Prat, A. (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *Quarterly Journal of Economics, 133*(2), 801–870.
Hardy, D. C., & Pazarbasioglu, C. (1998). Leading indicators of banking crises: Was Asia different? *IMF Working Paper*, *98*(91).
Heston, S., & Sinha, N. (2017). News vs. sentiment: Predicting Stock returns from news stories. *Financial Analysts Journal, 73*(3), 67–83.
Hoberg, G., & Phillips, G. M. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *The Review of Financial Studies, 23*(10), 3773–3811.
Hug, S., & Schulz, T. (2007). Left-Right positions of political parties in Switzerland. *Party Politics, 13*, 305–330.
Joy, M., Rusnák, M., Šmídková, K., & Vašíče, B. (2015). Banking and currency crises: Differential diagnostics for developed countries. *International Journal of Finance & Economics, 22*(1), 44–67.
Kaminsky, G., & Reinhart, C. (1999). The twin crises: The causes of banking and balance of payments problems. *American Economic Review, 89*, 473.
Kindleberger, C. (1978). *Manias, panics, and crashes: A history of financial crises*. Basic Books.
Kloptchenko, A., Magnusson, C., Back, B., Visa, A., & Vanharanta, H. (2004). Mining textual contents of financial reports. *The International Journal of Digital Accounting Research, 4*(7), 1–29.
Klüver, H. (2009). Measuring interest group influence using quantitative text analysis. *European Union Politics, 10*(4), 535–549. https://doi.org/10.1177/1465116509346782
Kou, G., Chao, X., Peng, Y., Alsaadi, F. E., & Herrera-Viedma, E. (2019). Machine learning methods for systemic risk analysis in financial sectors. *Technological and Economic Development of Economy, 25*(5), 716–742.
Laeven, L., & Valencia F. (2010), Resolution of banking crises: The good, the bad and the ugly. *IMF Working Paper*, *10*(146): 1.
Laeven, L., & Valencia, F. (2013). Systemic banking crises database. *IMF Economic Review, 61*, 225–270.
Laver, M., & Benoit, K. (2002). Locating TDs in policy spaces: Wordscoring dail speeches. *Irish Political Studies, 17*, 59–73.
Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *The American Political Science Review, 97*(2), 311–331.
Laver, M., & Budge, I. (1992). Measuring policy distances and modelling coalition formation. In M. Laver & I. Budge (Eds.), *Party policy and government coalitions* (pp. 15–40). Macmillan.
Laver, N., Benoit, K., & Sauger, N. (2006). Policy competition in the 2002 French legislative and presidential elections. *European Journal of Political Research, 45*, 667–697.
Lindgren, C.J., Garcia, G., & Saal, M.I. (1996). Bank soundness and macroeconomic policy. *IMF Working Paper, 215*.

Meyer, B., Bikdash, M., & Dai, X. (2017). Fine-grained financial news sentiment analysis. *SoutheastCon 2017* (pp. 1–8). IEEE. Chalotte, NC, USA. https://doi.org/10.1109/SECON.2017.7925378
Minsky, H. P. (1982). The financial instability hypothesis: Capitalistic processes and the behavior of the economy. In C. P. Kindleberger & J.-P. Laffargue (Eds.), *Financial crises: Theory, history and policy*. Cambridge University Press, NewYork.
Nyman, R., Kapadia, S., & Tuckett, D. (2021). News and narratives in financial systems: Exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, *127*.
Proksch, S. O., & Slapin, J. B. (2006). Institutions and coalition formation: The German election of 2005. *West European Politics, 29*, 540–559.
Proksch, S. O., & Slapin, J. B. (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science, 52*(3), 705–722.
Proksch, S. O., & Slapin, J. B. (2009). Position taking in European parliament speeches. *British Journal of Political Science, 40*(3), 587–611. https://doi.org/10.1017/S0007123409990299
Püttmann, L. (2018). Patterns of panic: Financial crisis language in historical newspapers, *SSRN*. https://ssrn.com/abstract=3156287
Ranciere, R., Tornell, A., & Westermann, F. (2008). Systemic crises and growth. *Quarterly Journal of Economics*, 123.
Reinhart, C. M., & Rogoff, K. S. (2009). *This time is different: Eight centuries of financial folly*. Princeton Press.
Reinhart, C. M., & Rogoff, K. S. (2014). Recovery from financial crises: Evidence from 100 episodes. *American Economic Review: Papers & Proceedings, 104*(5), 50–55. https://doi.org/10.1257/aer.104.5.50
Romer, C. D., & Romer, D. H. (2017). New evidence on the aftermath of financial crises in advanced countries. *American Economic Review, 107*(3), 72–118.
Samitas, A., Kampouris, E., & Kenourgios, D. (2020). Machine learning as an early warning system to predict financial crisis. *International Review of Financial Analysis*, 71.
Seyman, A. (2008). A critical note on the forecast error variance decomposition. *Centre for European Economic Research, Discussion Paper*, Number 08–065.
Shapiro, A. H., Sudhof, M., & Wilson, D. W. (2020). Measuring news sentiments. *Federal Bank of San Francisco*, *Working Paper, 2017*(01).
Sinha, N. R. (2016). Underreaction to news in the US stock market. *Quarterly Journal of Finance, 6*(2), 165000501–165000546.
Sun, X. and Xu, W. C. (2014). Fast implementation of Delongs algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Processing Letters, 21*, 1389–1393.
Temin, P. (1976). *Did monetary forces cause the great depression?* Norton.
Tromp, E., Pechenizkiy, M., & Gaber, M. M. (2017). Expressive modeling for trusted big data analytics: Techniques and applications in sentiment analysis. *Big Data Analytics, 2017*(2), 5. https://doi.org/10.1186/s41044-016-0018-9
Tsai, M. F., & Wang, C. J. (2017). On the risk prediction and analysis of soft information in finance reports. *European Journal of Operational Research, 257*(1), 243–250. https://doi.org/10.1016/j.ejor.2016.06.069
Tuckett, D. (2017). The future of macroeconomics: why observation of the behaviour of human actors and how they combine within the economy, is an important next step. Conference paper, *Institute of New Economic Thinking* (INET), October, 2017.
Volkens, A. (2001). Quantifying the election programmes: Coding procedures and controls. In I. Budge, H.-D. Klingemann, A. Volkens, J. Bara, & E. Tanenbaum (Eds.), *Mapping policy preferences: Estimates for parties, electors, and governments 1945–1998* (pp. 93–109). Oxford University Press.
Wang, T., Zhao, S., Zhu, G., & Zeng, H. (2021). A machine learning-based early warning system for systemic banking crises. *Applied Economics, 53*(26), 2974–2992. https://doi.org/10.1016/j.ejor.2016.06.069
Wicker, E. (1980) A reconsideration of the causes of the banking panic of 1930. *Journal of Economic History, 40*.