ECONOMIC NEWSLETTER
No. 24-02 | July 2024

CONSTRUCTING HIGH-FREQUENCY AND THEMATIC ECONOMIC SENTIMENT INDICATORS FROM ONLINE NEWS ARTICLES

By
Alan Chester T. Arcin, Carmelita G. Esclanda-Lo, Chelsea Anne S. Ong, and Rossvern S. Reyes

Mr. Arcin is Senior Research Specialist at the Department of Economic Research. Ms. Esclanda is Bank Officer IV, Ms. Ong is Senior Research Specialist, and Mr. Reyes is Acting Bank Officer VI at the Department of Economic Statistics. This work was presented in the 3rd Bank for International Settlements – Irving Fisher Committee on Central Bank Statistics (BIS-IFC) Workshop on Data Science in Central Banking last 17-19 October 2023. The authors are grateful to Ms. Cherrie Mapa and Ms. Genna Paola Centeno for their invaluable contributions to this study.

Abstract

This study develops a high-frequency indicator based on online business and financial news articles to infer market sentiment, aptly named the News Sentiment Index (NSI). Specifically, two NSIs are constructed using dictionary and machine learning methods. Furthermore, this study utilizes topic modeling to explore sentiments by theme. Results indicate that the NSIs track key economic events and show strong positive correlation with other economic indicators. Sentiments of specific themes are also found to be positively and strongly correlated with economic indicators. This study finds that the use of NSIs and topic modeling can potentially complement existing survey-based measures to timely gauge market sentiment.

1. Introduction

Sentiment of economic agents can have a considerable impact on real economic activity. For instance, consumer sentiment can influence consumption and saving decisions, while business sentiment can affect investment, hiring, and pricing decisions. As such, measurement of market sentiment is an important component of macroeconomic surveillance.

Typically, sentiment or people's attitudes or opinions toward a particular topic is measured through surveys. The Bangko Sentral ng Pilipinas (BSP) deploys the Business Expectations Survey (BES) and Consumer Expectations Survey (CES) every quarter to gauge the economic sentiment of firms and households, respectively. However, conducting surveys can be tedious, often time-consuming, and expensive. The lags in the release of survey results imply that information that could be important for economic surveillance and policymaking may not be available on a near-real-time basis.

More recently, the accessibility of possible alternative high-frequency sources of information, such as online newspaper articles, and the growing field of sentiment analysis techniques opened opportunities for innovating sentiment measurement. This study explores the application of sentiment analysis techniques on news data to determine market sentiment and generate news-based sentiment indices (NSIs). Two indices are developed using: (1) dictionary method and (2) machine learning method.

This study adds to the increasing list of central banks [Federal Reserve Bank of San Francisco (Shapiro et al., 2020), Bank Negara Malaysia (Chong et al., 2022), Bank of Korea (Seo, 2022), Reserve Bank of Australia (Nguyen & Cava, 2020), and Narowdy Bank Polski (Marszal, 2022)] that have published studies involving sentiment analysis of online news data.

Overall, this research shows that the NSIs match the expected sentiment of key economic events. These indices display the potential to supplement existing surveys and provide a timely measurement of market sentiment. This study also delves into topic modeling to identify key themes that could be useful in predicting key economic indicators.

2. Data collection and pre-processing

Data were collected from multiple news sources in the web.[1] The sections of interest are limited to economy, business, finance, and related fields (e.g., stock market, banking). The collected news data were pre-processed by removing extra spaces and punctuation marks. Further steps include conversion of characters to lowercase and conversion of text into machine-readable format.[2]

2.1 Sentiment Analysis Techniques

Broadly, sentiment analysis is a technique used to determine the tone of text, whether positive, negative, or neutral. As discussed above, this study utilizes two approaches in determining sentiment: (a) dictionary-based method and (b) machine learning-based method.

The dictionary-based method involves a pre-defined list of keywords and a set of rules.[3] This study employed a combination of some of the most used lexicons in finance and well-known general lexicons.[4] In particular, these off-the-shelf lexicons are (a) General Inquirer (GI); (b) Loughran-McDonald Sentiment Dictionary (LM); (c) FED Financial Sentiment Dictionary (FD): (d) Hu and Liu Opinion Lexicon (HL); and (e) Valence Aware Dictionary for Sentiment Reasoning (VADER). Using the above dictionaries, the sentiment score of an article is determined by subtracting the number of negative words from the number of positive words per article and dividing by the total number of words in the article.

Meanwhile, the machine learning method predicts the sentiment of text using a pre-trained model. Two machine learning models were used namely: Bidirectional Encoder Representations from Transformers or BERT (Devlin, 2018) and Financial BERT or FinBERT (Huang, A. H., Wang,

---

[1] The authors have coordinated with several media outlets to seek permission to web scrape their news data. The following media outlets have agreed to have their news articles retrieved online: Business World, Manila Bulletin, Inquirer.NET, Manila Standard, and Business Mirror.
[2] Conversion of text to machine readable format involves two steps. First, text is split into smaller units referred to as tokens (i.e., tokenization). For example, paragraphs are split into sentences and sentences into words. Second, the resulting tokens are converted into a representation of numerical data (i.e., vectorization).
[3] For this study, the negation rule was applied, where words switch from positive to negative and vice-versa when preceded by a negation word, e.g., “not” or “never”.
[4] A lexicon or dictionary is a collection of words where words are labeled with either positive, negative, or neutral tone.

H., and Yang, Y., 2022). In this method, the sentiment score of an article is derived by subtracting the number of negative sentences from the positive sentences and dividing by the number of sentences in the article.

To assess the performance of the different lexicons and machine learning models, a set of 3000 labeled sentences by the Department of Economic Research (DER) were used as a test dataset. The selected sentences were also taken from the same news data sources and contained at least one word from any of the five lexicons. This labeled dataset also enabled the creation of a Philippine-specific lexicon using a genetic algorithm (hereafter referred to as PH lexicon) that retrieves a combination of keywords that maximizes the number of correctly classified sentences from the test dataset.[5]

The reliability of sentiment classification using five off-the-shelf lexicons along with the constructed PH lexicon and the two pre-trained machine learning models were appraised vis-à-vis the manually labeled dataset based on accuracy and macro-F1 scores.[6] The lexicons are then ranked and the lexicon with the highest accuracy and macro F1 scores is selected to generate the NSI - dictionary method (NSI-D). The same is also applied to the machine learning methods where the model with the highest accuracy and macro-F1 is selected to generate the NSI-machine learning method (NSI-ML). Table 1 shows the accuracy and macro F1 scores of the various lexicons and machine learning models. As may be expected, the lexicon created using the genetic algorithm (PH lexicon) outperformed other lexicons, while FinBERT surpassed BERT in terms of accuracy and macro F1 score. Thus, both the PH lexicon and FinBERT model were chosen to generate the NSI-D and NSI-ML, respectively.

Table 1. Lexicons and Machine Learning Models: Accuracy and Macro F1 scores

Lexicons	Accuracy	Macro F1
PH Lexicon	0.66	0.64
LM Lexicon	0.52	0.46
FD Lexicon	0.51	0.51
VADER	0.51	0.42
HL Lexicon	0.50	0.46
GI Lexicon	0.45	0.42
Machine Learning Models	Accuracy	Macro F1
FinBERT	0.66	0.63
BERT	0.58	0.45

Source: Authors' estimates
Note: These are computed relative to the DER annotated dataset.

[5] Genetic algorithms are used for optimization problems, inspired by the evolutionary process.
[6] Accuracy measures the fraction of the model's correct predictions over total predictions. Meanwhile, macro F1 score combines both precision (a measure of the proportion of true positives that were correct) and recall (a measure of the proportion of true positives that were actually predicted).

2.2 Construction of the News Sentiment Indices

Following the selection of the best-performing lexicon and machine learning model, sentiment scoring of the entire news dataset covering the period from January 2018 to December 2022 was undertaken.

The NSIs are calculated from the average sentiment scores of the articles in each month, adjusted by the number of articles per news source. The indices are also standardized to a historical average of zero.

Figure 1 presents the NSI-D and NSI-ML. Estimates higher than zero indicate positive sentiment, while estimates below zero suggest negative market sentiment.

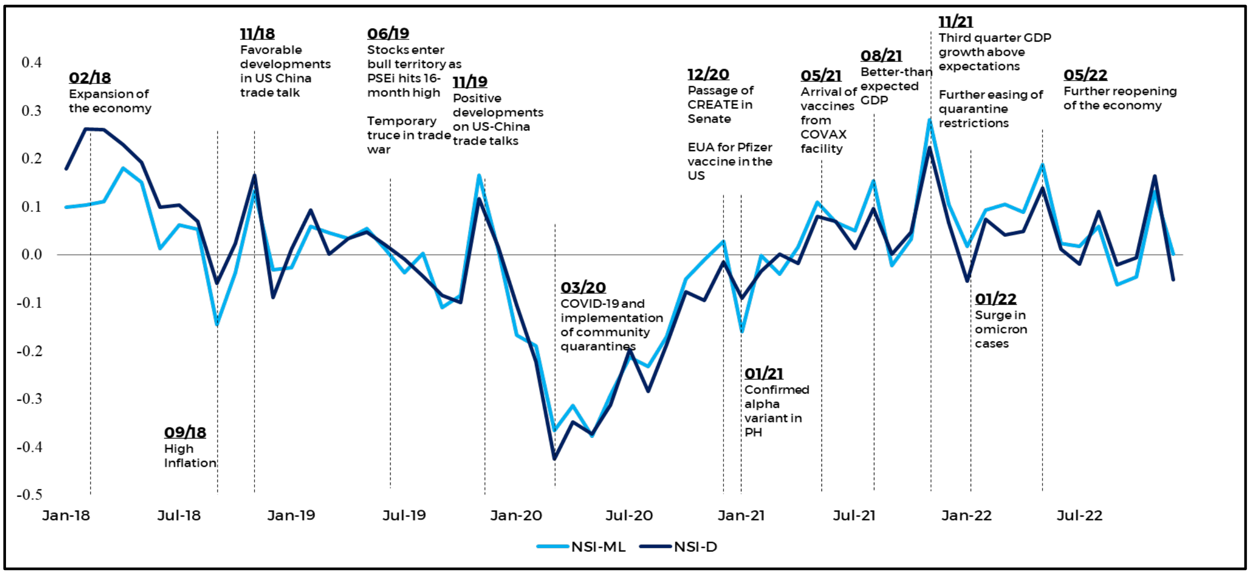
3. Analysis of results

This section presents an analysis of the NSI-D and NSI-ML estimates. Figure 1 shows that the indices co-move together, suggesting broadly similar estimates even if these NSIs are derived using two different methodologies.

More importantly, Figure 1 suggests that the NSI estimates match the expected market sentiment of key events that may have significant economic repercussions. For example, sentiment dipped and turned significantly negative, following the onset of the COVID-19 pandemic and the implementation of community quarantines. It marginally recovered in December 2020 (the first time the index registered a positive reading since the pandemic), coinciding with the passage of the Corporate Recovery and Tax Incentives for Enterprises (CREATE) bill in the Senate and the issuance of the emergency use authorization of the Pfizer vaccine in the US. Sentiment estimates were also above zero in May 2021 when the vaccines from the World Health Organization-led COVAX facility arrived in the Philippines. Further, the sentiment estimates captured the effects of economic surprises, such as when the actual gross domestic product (GDP) grew faster than expected. Other key events that match the peaks (positive sentiment) and troughs (negative sentiment) of the NSIs have been identified in Figure 1 for the readers' reference.

Meanwhile, Table 2 presents the correlation of NSI-D and NSI-ML against selected economic indicators such as Philippine Stock Exchange Index (PSEi), Purchasing Managers' Index (PMI), and the confidence indices (CI) from the BSP BES and CES. As expected, the NSI-D and NSI-ML are positively correlated with the PSEi, PMI, BES CI, and CES CI. The correlation values for NSI-D for CES and NSI-ML for both BES and CES are not statistically significant.

Figure 1. News Sentiment Indices and Key Economic Events
January 2018 – December 2022



Source: Authors' estimates

Table 2. Correlation Coefficients of Models against Selected Economic Indicators

January 2018 – December 2022

Model	PSEi	PMI	BES	CES
NSI-D	0.61 (0.00)*	0.71 (0.00)*	0.49 (0.03)*	0.38 (0.10)
NSI-ML	0.53 (0.00)*	0.67 (0.00)*	0.43 (0.07)	0.28 (0.23)

*p-value < 0.05 – statistically significant relationship
Source: Authors' estimates

Note: Sentiment indices per topic were estimated using NSI-D approach. Values in parenthesis are corresponding p-values.

4. Topic modeling

This study also uncovers the themes that best characterize the collection of news articles. Topic modeling involves using an unsupervised machine learning technique to cluster text based on similarities. This study utilizes two of the most used methods for topic modeling: 1) Latent Dirichlet Allocation (LDA) and 2) Non-Negative Matrix Factorization (NMF). After evaluating the two models, NMF was chosen to retrieve the topics for the articles due to its speed and ease of use while yielding similar results to LDA.[7]

[7] LDA iteratively computes for distributions of word probabilities per document while NMF uses matrix factorization to retrieve topics from a collection of documents.

The topic modeling performed determined 12 as the optimal number of topics to represent the data. These topics or themes were energy, agriculture, services, corporate news, stock market, micro, small and medium enterprises (MSMEs), banking, trade, government finance, government projects, company earnings, and inflation. Figure 2 shows some of the word clouds associated with these topics.

Due to relative ease of implementation and higher correlation coefficients with economic indicators compared to NSI-ML (see Table 2), NSI-D was used to generate separate NSIs for each of the identified topics. Correlation coefficients were also computed for each of the topics with PSEi, PMI, BES and CES CI. Table 3 shows the topics whose sentiment scores posted the highest correlation with the monthly PSEi and PMI.

Figure 2. Word clouds of selected topics

Source: Authors' estimates

Table 3. Top Topics with NSIs that Exhibited Highest Correlation Coefficients with monthly PSEi and PMI

January 2018 – December 2022

Topic	PSEi	PMI
Company Earnings	0.70	0.54
Services	0.39	0.48
Trade	0.39	0.62
Banking	0.45	0.50
Infrastructure	0.41	0.34

Source: Authors' estimates
Note: Sentiment indices per topic were estimated using NSI-D approach. The calculated p-values are close to 0 (p-value < 0.05), hence statistically significant.

Company earnings (news articles relating to financial disclosures of firms) ranked highest among the topics that are correlated with the PSEi, followed by trade and services. Thus, it can be inferred that the NSI from news on company earnings can be used to gauge PSEi's movements and serve as a good indicator of market confidence. Meanwhile, the trend of market sentiment based on trade-related news (i.e., articles relating to imports and exports) appeared to coincide with the PMI series.

Table 4 shows the topics whose NSIs had the highest correlation coefficients with CIs of the BES and CES. NSI of news relating to company earnings ranked highest, followed by infrastructure and trade. This implies that market sentiment related to company earnings can also be an alternative measure of consumer and business outlook. Interestingly, these topics also displayed high positive correlation with the growth rate of the GDP.

Table 4. Top Topics with NSIs that Exhibited Highest Correlation Coefficients with quarterly BES CI, CES CI and GDP growth rate

January 2018 – December 2022

Topic	BES CI	CES CI	GDP
Company Earnings	0.77	0.79	0.81
Infrastructure	0.60	0.62	0.70
Trade	0.47	0.36	0.78

Source: Authors' estimates
Note: Sentiment indices per topic were estimated using NSI-D approach. The calculated p-values are close to 0 (p-value < 0.05), hence statistically significant.

It may also be noted that the above topic-level NSIs were more correlated with the selected economic indicators than the general NSIs. This suggests that topic modeling may bring out specific facets of news that are more significant in explaining co-movement with selected economic indicators, particularly the BES, CES and GDP growth. Therefore, these topic NSIs may be used as potential complementary indices of the previously mentioned key economic measures.

5. Conclusion

News-based indices offer a faster and more cost-effective approach to high-frequency sentiment measurement by leveraging widely available news articles and innovations in sentiment analysis techniques.

Evaluation of the indices shows that the constructed NSIs match expected sentiment for critical economic events. This suggests that NSIs have the potential to complement the BSP's existing BES and CES and provide a timely measurement of market sentiment.

Meanwhile, topic modeling helped identify key themes in the corpus of news data utilized in this study. Notably, some topic-level NSIs demonstrated higher correlations with key economic indicators than broad NSIs. This implies that topic modeling can help uncover specific news themes that may have more predictive content for key economic variables, particularly the BES, CES and GDP growth rate. While correlation does not imply causality, the high correlations between the identified topics and key economic indicators are encouraging and suggest that further research on this quantitative method is worth pursuing.

References

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3, 993-1022.

Chong, E., Ho, C. C., Ong, Z. F., & Ong, H.H, (2022). Using news sentiment for economic forecasting: a Malaysian case study. IFC Bulletins chapters, in: Bank for International Settlements (ed.), Machine learning in central banking, volume 57, Bank for International Settlements.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805. Retrieved from https://arxiv.org/abs/1810.04805

Huang, Allen and Wang, Hui and Yang, Yi, FinBERT - A Large Language Model for Extracting Information from Financial Text (July 28, 2020). Contemporary Accounting Research, Forthcoming, Available at SSRN: https://ssrn.com/abstract=3910214 or http://dx.doi.org/10.2139/ssrn.3910214

Holland, John (1992). Adaptation in Natural and Artificial Systems. Cambridge, MA: MIT Press. ISBN 978-0262581110.

Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

Kanzari, D., Nakhli, M. S., Gaies, B., & Sahut, J. (2023). Predicting macro-financial instability – How relevant is sentiment? Evidence from long short-term memory networks. Research in International Business and Finance, 65, 101912. https://doi.org/10.1016/j.ribaf.2023.101912

Lee, D. D., & Seung, H. S. (2000). Algorithms for non-negative matrix factorization. In Advances in Neural Information Processing Systems (pp. 556-562).

Li, X., Ding, Y., Chen, X., & Lin, X. (2019). FinBERT: A Pretrained Language Model for Financial Communications. arXiv preprint arXiv:1908.10063. Retrieved from https://arxiv.org/abs/1908.10063

Marszal, Anna (2022). What news can really tell us? Evidence from a news-based sentiment index for financial markets analysis, NBP Working Papers 349, Narodowy Bank Polski.

Nguyen, K., & Cava, G. L. (2020). Start Spreading the News: News

Sentiment and Economic Activity in Australia. https://www.rba.gov.au

Seo, Beomseok (2022). Machine-Learning-Based News Sentiment Index (NSI) of Korea (한국의 기계학습 기반 뉴스심리지수) (2022). Bank of Korea WP 2022-15, Available at SSRN: https://ssrn.com/abstract=4230611 or http://dx.doi.org/10.2139/ssrn.4230611

Shapiro, Adam Hale, Moritz Sudhof, and Daniel Wilson (2020). "Measuring News Sentiment," Federal Reserve Bank of San Francisco Working Paper 2017-01. Available at https://doi.org/10.24148/wp2017-01

Stevens, K., Kegelmeyer, P. W., Andrzejewski, D., & Buttler, D. (2012). Exploring Topic Coherence over Many Models and Many Topics. In Empirical Methods in Natural Language Processing (pp. 952–961). Available at: https://www.aclweb.org/anthology/D12-1087.pdf