

# Newspapers' Biases and Unbiased Economic Index

*Keywords: Business confidence index, Political orientation, Sentiment analysis, Deep learning, Dynamic factor model*

## Extended Abstract

All kinds of data from the web (e.g., news flashes and web search logs) to our personal information (e.g., credit card transactions and social media activities) are aggregated and exploited to optimize our decisions. In finance and economics, such data have been used as “alternative data” to predict financial and economic variables, including stock prices [3], market sentiment [4], gross domestic product [1], etc. For instance, newspaper articles can be used for predicting a business sentiment index (BSI for short). BSI gauges current/future mood in business, which helps economists, policymakers, and manufacturers make appropriate plans.

Traditional BSIs have been calculated based on a survey of business confidence, which is costly and time-consuming. Instead, Seki and Ikuta [6] used the Nikkei (Japan Economics Newspaper) as input to a deep learning model to nowcast a BSI, which was reported to be strongly and positively correlated (up to Pearson’s correlation coefficient  $r = 0.937$ ) with an existing survey-based index, specifically, Economy Watchers DI (EWDI for short)<sup>1</sup>. Additionally, they tested a different newspaper as input and reported a lower correlation with EWDI ( $r = 0.831$ ) and conjectured that the difference came from the fact that the Nikkei focuses on finance and economics, whereas the other is a general newspaper.

While the newspaper’s specialty would be playing an important role, it may not be the only factor that caused the difference. For instance, their political ideologies or leaning (conservative or right-leaning vs. liberal or left-leaning) may influence how news articles are composed even for the same event, such as an economic policy, especially when the policy is controversial. Because the computed BSI is based on news text, the tone and choice of words and expressions would directly affect the resulting index. Conversely, one may be able to quantify and observe, for example, the political position of an input newspaper through its BSI.

We investigated the very idea that newspapers’ biases could be observed through their BSIs. To this end, we conducted an experiment using two years’ worth of four national daily newspapers in Japan: Yomiuri, Nikkei, Mainichi, and Asahi, which are known to have different political orientations. Yomiuri and Nikkei are generally considered to be conservative and right-leaning, whereas Mainichi and Asahi are generally considered to be liberal and left-leaning. Additionally, Nikkei and Mainichi are often regarded as close to the center than the other two [5]. The following summarizes the methodology and the result of the experiment.

First, we computed a monthly news-based BSI for each newspaper by a Transformer-based model [6]. Figure 1 compares the resulting BSIs, where one can observe that Yomiuri’s BSI is the highest, followed by Nikkei, Asahi, and Mainichi, although the pattern has slightly changed after April 2020. In other words, they roughly align with their political orientations.

Then, we analyzed the four BSIs by a dynamic factor model (DFM), which is a time series analysis model typically used to identify a set of unobserved common factors underlying multiple time series. Specifically, this study adopted the DFM proposed by Berendrecht [2] and attempted to find a common dynamic factor as well as specific dynamic factors characterizing individual time series. A common dynamic factor can be interpreted as the latent business

---

<sup>1</sup>Economy Watchers DI is released monthly by the Cabinet Office of Japan.

sentiment underlying different newspapers, and the specific dynamic factors as their tones or biases on economic and financial events not explained by the common business sentiment.

Figure 2 (left) shows that the common dynamic factor in a thick line appears to capture the general business sentiment of the multiple BSIs of the different newspapers. On the other hand, Figure 2 (right) shows that the trends of the unstandardized specific dynamic factors are relatively stable even during the COVID-19 shock from January 2020 onward except for Nikkei. The upward trend of Nikkei after April 2020 was also seen in Figure 1 but is more clearly seen after the decomposition by DFM. The result demonstrates that DFM was useful to separate the general trend of business sentiment and individual newspapers' biases. The result also suggests that those biases are not necessarily static and may notably change in a relatively short period when a significant economic shock arises.

Lastly, we examined the common dynamic factor as an unbiased business sentiment and compared it with other economic indices (Composite Indices and Tankan) by cross-correlation. Figure 3 (left) shows that both the common dynamic factor (CDF) and Nikkei's BSI precede the coincident CI by a month but the correlation is much stronger for the CDF ( $r = 0.942$ ) than Nikkei's BSI ( $0.817$ ) and EWDI ( $r = 0.745$ ). Similarly, Figure 3 (right) shows that CDF was found to have the strongest correlation with Tankan ( $r = 0.910$ ) than Nikkei's BSI ( $0.759$ ) and EWDI ( $0.668$ ). The result indicates that CDF extracted from the four news-based BSIs is a good alternative to CI and Tankan.

The major findings and implications of this study are as follows: (1) The magnitudes of the news-based BSIs roughly correspond to their political positions; (2) The specific dynamic factors indicate the biases of the individual newspapers more accurately than raw BSI values, and Nikkei's bias has been continuously increasing after the COVID-19 shock started at the beginning of 2020, implying its position as an influential economic newspaper to encourage the economy; and (3) The common dynamic factor can be seen as an unbiased BSI better representing other economic indices (i.e., CI and Tankan) than Nikkei's BSI or EWDI. In conclusion, this study empirically showed that newspapers' political/economic biases can be monitored through news-based BSIs and that an unbiased BSI can be derived as a more accurate indicator of the other economic indices than EWDI or a raw BSI.

## References

- [1] Andreini, P., Hasenzagl, T., Reichlin, L., Senftleben-König, C. and Strohsal, T.: Nowcasting German GDP: Foreign factors, financial markets, and model averaging, *International Journal of Forecasting* (2021).
- [2] Berendrecht, W. and van Geer, F. C.: A dynamic factor modeling framework for analyzing multiple groundwater head series simultaneously, *Journal of Hydrology*, Vol. 536, pp. 50–60 (2016).
- [3] Derakhshan, A. and Beigy, H.: Sentiment analysis on stock social media for stock price movement prediction, *Engineering Applications of Artificial Intelligence*, Vol. 85, pp. 569–578 (2019).
- [4] Ge, Y., Qiu, J., Liu, Z., Gu, W. and Xu, L.: Beyond negative and positive: Exploring the effects of emotions in social media during the stock market crash, *Information Processing & Management*, Vol. 57, No. 4, p. 102218 (2020).
- [5] Kaneko, T., Asano, T. and Miwa, H.: Estimating Ideal Points of Newspapers from Editorial Texts, *The International Journal of Press/Politics*, Vol. 26, No. 3, pp. 719–742 (2021).
- [6] Seki, K., Ikuta, Y. and Matsubayashi, Y.: News-based Business Sentiment and its Properties as an Economic Index, *Information Processing & Management*, Vol. 59, No. 2, p. 102795 (2022).

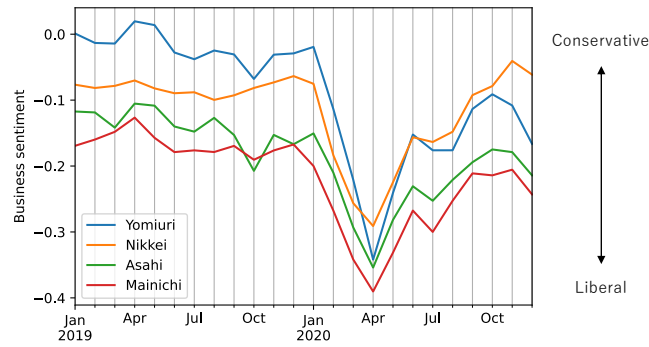


Figure 1: Comparison of BSIs computed from the four national newspapers. The magnitudes of the BSIs roughly align with the conceived political spectrum of the newspapers.

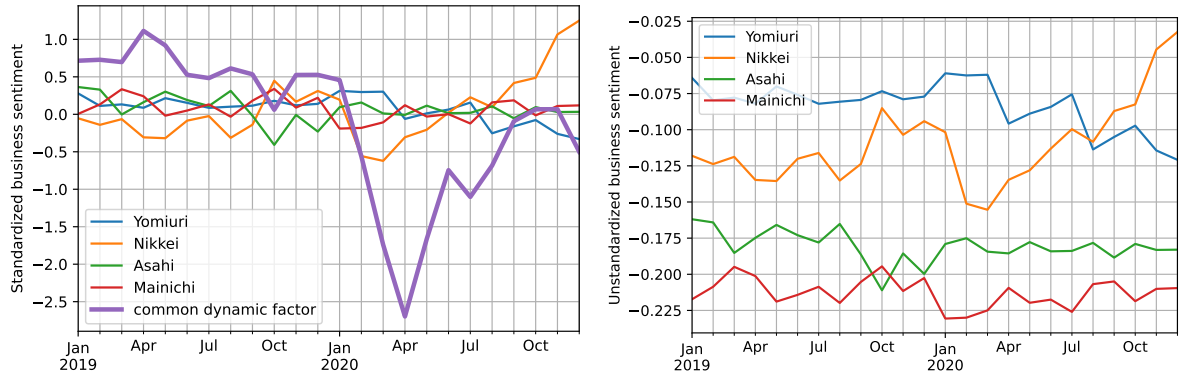


Figure 2: Specific and common dynamic factors for standardized series (left) and unstandardized specific dynamic factors (right). The latter shows the biases (e.g., political/economic tones) of the individual newspapers, where Nikkei's bias has noticeably changed during/after the COVID-19 shock in 2020.

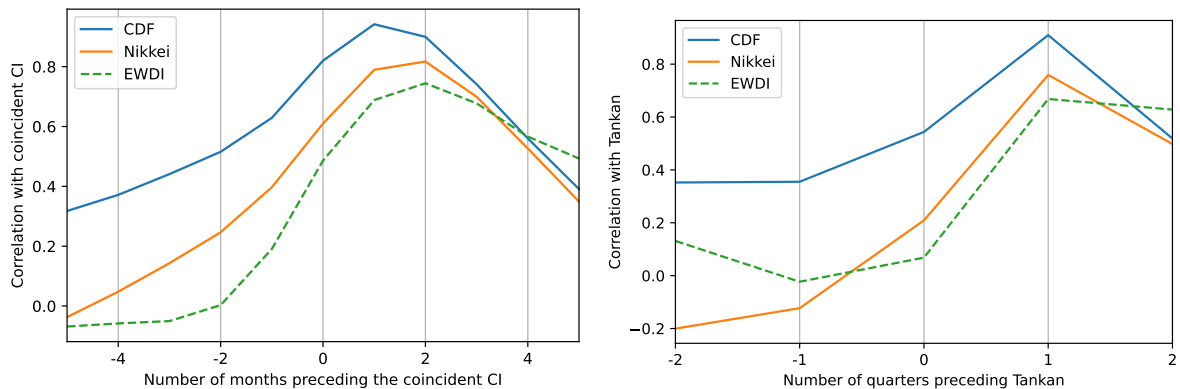


Figure 3: Cross-correlation of the common dynamic factor (CDF), Nikkei-based BSI, and EWDI with other economic indices, i.e., coincident CI (left) and Tankan (right). CDF can be seen as unbiased business sentiment underlying the different newspapers and more highly correlated with CI and Tankan than Nikkei's BSI and EWDI.