Received: 16 July 2018 | Revised: 11 January 2019 | Accepted: 12 February 2019

DOI: 10.1002/for.2584

RESEARCH ARTICLE

WILEY

Forecasting economic indicators using a consumer sentiment index: Survey-based versus text-based data

Minchae Song[1] | Kyung-shik Shin[2]

[1] Big Data Analytics, Ewha Womans University, Seoul, Republic of Korea

[2] School of Business, Ewha Womans University, Seoul, Republic of Korea

Correspondence
Kyung-shik Shin, School of Business, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-Gu, Seoul, 120-750, Republic of Korea
Email: ksshin@ewha.ac.kr

Abstract

Given the confirmed effectiveness of the survey-based consumer sentiment index (CSI) as a leading indicator of real economic conditions, the CSI is actively used in making policy judgments and decisions in many countries. However, although the CSI offers qualitative information for presenting current conditions and predicting a household's future economic activity, the survey-based method has several limitations. In this context, we extracted sentiment information from online economic news articles and demonstrated that the Korean cases are a good illustration of applying a text mining technique when generating a CSI using sentiment analysis. By applying a simple sentiment analysis based on the lexicon approach, this paper confirmed that news articles can be an effective source for generating an economic indicator in Korea. Even though cross-national comparative research results are suited better than national-level data to generalize and verify the method used in this study, international comparisons are quite challenging to draw due to the necessary linguistic preprocessing. We hope to encourage further cross-national comparative research to apply the approach proposed in this study.

KEYWORDS

consumer sentiment index, economic indicator forecasting, sentiment analysis, text mining

1 | INTRODUCTION

Keynes realized that employment and production decisions are based on expected consumer demand and that aggregate demand follows firms' production and employment decisions (Benhabib, Wang, & Wen, 2015). In response to the widespread belief that consumers' opinions and expectations influence the direction of the economy, a growing number of studies have set out to analyze the relationship between consumer attitudes and economic variables (Bram & Ludvigson, 1997; Gelper, Lemmens, & Croux, 2007; Ludvigson, 2004). Several papers argued that sentiment shocks can drive aggregate business conditions (Angeletos & La'O, 2013; Benhabib et al., 2015; Benhabib & Wen, 2004; Farmer & Guo, 1994). An economic interpretation of sentiment shocks is simply that economic agents such as households, firms, and governments sometimes become optimistic or pessimistic about future consumption, future investment, or future inflationary pressures, and those attitudes appear in the form of sentiment shocks. Sentiment shocks are therefore identified from the dynamic interactions between observed expectations and realized macroeconomic time series (Arias, 2016; Milani, 2017). Other researchers have suggested that sentiment may reflect information about future states of the economy held by individuals but not (yet) observed in publicly released data. Barsky and Sims (2012) found out that this dynamic information forms the main link between sentiment and future activity. For these reasons, consumer sentiment is treated as an important piece of economic information (Ludvigson, 2004; Uhl, 2011).

Previous studies showed that a survey-based consumer sentiment index (CSI)[1] is an effective leading indicator of real economic conditions (Acemoglu & Scott, 1994; Howrey, 2001). The CSI provides qualitative information for monitoring current economic situation and advances warning of turning points in economic activity. In South Korea, the Bank of Korea (BOK) releases the CSI at the end of each month, which attempts to account for economic and sentimental aspects of consumer behavior by asking households to assess current and upcoming economic conditions at personal and national levels.

As Figure 1 shows, trends in South Korea's domestic private consumption growth rate are closely related (with a correlation coefficient of 0.75) to the CSI.[2] Private consumption, which accounts for 50% of total gross domestic production (GDP), the CSI serves as a very important economic leading indicator for evaluating and forecasting domestic economic conditions. Also, private consumption of (preliminary) GDP in South Korea is only published quarterly, so indices with higher publishing frequencies can be helpful in not only presenting trends but also predicting the future.

Although the CSI offers relevant insights into private consumption and GDP, a traditional approach to measure consumer sentiment based on the survey has several limits. One possible weakness is that it takes considerable time to research, collect, and aggregate the data. If certain urgent issues arise, timely information will not be announced until the end of each month. In addition, the survey only contains information derived from questionnaire items, which means it can be difficult to catch up to the direct effect of newly arising issues. Hence it is necessary to find a way to complement the survey-based method.

For this purpose, we construct an index designed to measure consumers' economic sentiment using sentiment analysis, which is one of the text mining techniques. Unlike survey-based measures, our index extracts sentiment information from online economic news articles. News media can be great potential sources for generating economic indicators (Ulbricht, Kholodilin, & Thomas,

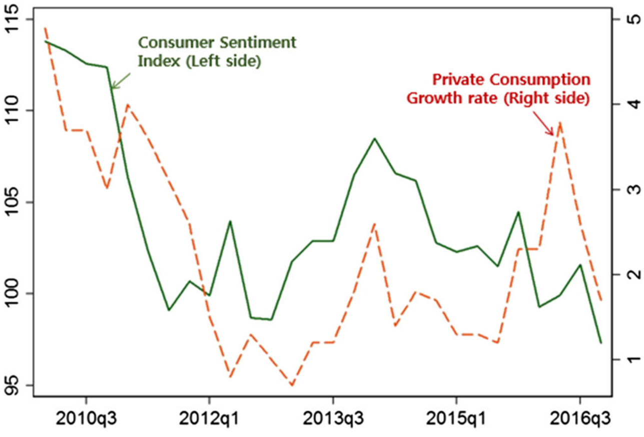

FIGURE 1 Trends of South Korea's private consumption growth rate and CSI [Colour figure can be viewed at wileyonlinelibrary.com]

2017). Firstly, news data can quickly capture the economic impact of specific issues and broadly cover certain issues from various angles. Secondly, online news is cheaper to obtain, announcing greater periodicities compared with survey-based measures. Thirdly, an impact of online news on individuals has steadily increased as there are several advantages of online news channels. These benefits include lower costs, easier multitasking, more news choices, in-depth and background information, 24/7 updates, customization, the ability to discuss the news with peers, the existence of different viewpoints, and the opportunity to “talk back” to the media (Chadha & Wells, 2016; Steensen, 2011). Lastly, news articles are relatively representative of household economic sentiments because they generally cover aggregates rather than individuals.

To be used as a meaningful economic indicator, it is important to check the effectiveness of the new sentiment index extracted from news articles. In line with existing scholarly assessments for confirming the relationship between CSI and macroeconomic indicators, we conduct a Granger causality test for a news consumer sentiment index (NCSI) and examine the sustained impact of sentiment shocks using impulse response function (IRF). Lastly, we estimate one step of out-of-sample prediction to investigate the new index's forecasting power. In almost all experiments, the NCSI strongly correlates with related contemporaneous key indicators, and news sentiment shocks predict well a household's future economic activities.

The contributions of this paper are that we attempt to construct a household's economic sentiment index using sentiment analysis to complement survey-based measures. We confirm through various statistical analyses that the newly generated index provides useful information for demand-side and supply-side of household

[1] CSI is also referred to as a consumer confidence index. In this study, we consistently use the CSI term.

[2] We conducted an augmented Dickey–Fuller (ADF) test and Phillips–Perron (PP) test to examine the presence of a unit root for the CSI and private consumption growth rate. The test results indicate that the two series have unit root. In addition, the NCSI and private consumption growth rate have a cointegrating relationship in the Johansen cointegration test.

economic activities in Korea. Although previous studies were less interested in the relationship between the CSI and the labor indicators, it is necessary to examine it because employment (unemployment) is one of the most important economic indicators to evaluate households' economic situations and has a strong correlation with a wide variety of measures. In particular, the impact of job security on economic sentiment increases when unemployment rates are high. On the demand side, we conduct our experiments by dividing goods into durables and nondurables and selecting certain services directly related to consumer economic sentiment. Because the purpose of consumption differs across goods and services, the impact of the sentiment index on those categories varies.

Cross-national comparative research results are better suited than national-level data to generalize and verify the method used in this study. However, international comparisons are quite challenging to draw due to the necessary linguistic preprocessing. Here, we demonstrate that Korean cases are a good illustration of using sentiment analysis when generating a CSI using text mining technique; we need to apply the same analysis to data from other countries to generalize and verify the proposed methodology.

The remainder of this paper is structured as follows. The next section describes related research, and Section 3 explains this study's methodology, focusing on the sentiment analysis of news articles. Section 4 deals with the experimental setup. Section 5 discusses the results, and Section 6 summarizes the work and makes suggestions for future research.

2 | RELATED WORKS

Sentiment analysis, which is also called opinion mining, uses computational methods to analyze opinions, sentiments, appraisals, attitudes, and emotions toward entities and their aspects that people express in written natural-language texts (Pang, Lee, & Vaithyanathan, 2002). In this context, opinion means the underlying positive or negative view implied in the text and aspect is the opinion target of the sentiment expression.

Many studies on sentiment analysis have been carried out on subjective texts, including blogs and product reviews (Daniel, Neves, & Horta, 2017; Kang, Yoo, & Han, 2012; Kontopoulos, Berberidis, Dergiades, & Bassiliades, 2013; Tang, Tan, & Cheng, 2009). In such texts, authors typically express their opinions freely and usually evaluate one or more aspects. On the other hand, relatively few researchers have paid attention to sentiment analysis for news articles (Balahur et al., 2010; Balahur & Steinberger, 2009; Balahur, Steinberger, van der Goot, Pouliquen, & Kabadjov, 2009; Van de Kauter, Breesch, & Hoste, 2015). In general, news articles aim to give the impression of objectivity, which means journalists often refrain from using clearly positive or negative words. They may express their opinions in other ways; for example, by embedding statements in more complex discourses or argumentative structures, they may quote other people who say what they think and feel. Therefore, automatically identifying sentiments is rather difficult compared to subjective texts, but lexically expressed opinions are still present in news data even if they occur less frequently than in product or service reviews (Balahur et al., 2010; Balahur & Steinberger, 2009). These features make opinion mining harder when applying an existing sentiment analysis techniques to news data. Balahur and Steinberger (2009) and Balahur et al. (2010) redefine tasks of sentiment analysis for news by identifying different aspects of sentiment in text and pinpointing what exactly they expect a sentiment analysis system to discover from news text based on the different aspects.

As unstructured textual information in social media, search queries, news, blogs, and forums has become an interesting topic of research, there has also emerged in the economic research area, for example, sentiment analysis for stock market prediction (Nguyen, Shirai, & Velcin, 2015; Oliveira, Cortez, & Areal, 2017; Perlin, Caldeira, Santos, & Pontuschka, 2017; Schumaker, Zhang, Huang, & Chen, 2012; Yu, Wu, Chang, & Chu, 2013). In addition, news and social media have been used to measure the economic sentiment index in several economic research studies (Daas & Puts, 2014; Shapiro, Kanjoya, & Wilson, 2017; Uhl, 2011). Uhl (2011) constructed a news sentiment index using a large corpus of economic news articles from LexisNexis. To conduct sentiment analysis at the document level, Uhl used a Naive Bayes (NB) classifier to distinguish between positive and negative sentiment from a predefined database of positive and negative words and phrases. A classification result of either positive or negative was generated for each article, and the daily sentiment data were then aggregated to quarterly values. The study focused on explaining consumer behavior by combining key economic variables such as personal income and savings, consumer prices, interest rates, stock prices, and exchange rates. Shapiro et al. (2017) also used the news media from LexisNexis and constructed two news sentiment series based on a lexical dictionary and a predictive model. They estimated a monthly sentiment index from the regressing sentiment scores by news article on the fixed effects of month and newspaper type. On the other hand, unlike the above two studies, Daas and Puts (2014) generated a new economic sentiment index using social media and conducted a sentiment classification at the sentence level. They measured the overall sentiment of a message by applying the lexicon-based approach, conducted various statistical analyses, and discussed the relationship between social media sentiment and consumer confidence in depth. The most interesting results of their study are that changes in the sentiment of social media routinely preceded changes in CSI and that the lag was of the order of 7 days. For this reason, they suggested that a weekly indicator with a frequency higher than monthly was an appropriate option for social media sentiment-based data.

3 | RESEARCH METHODOLOGY

3.1 | Sentiment analysis based on the lexicon approach

There are two main approaches to automatically extract sentiment from a text (Martín-Valdivia, Martínez-Cámara, Perea-Ortega, & Ureña-López, 2013). The lexicon-based approach, which is an unsupervised learning method, involves calculating sentiment based on the semantic orientation of words or phrases in the text (Godbole, Srinivasaiah, & Skiena, 2007; Tan & Wu, 2011; Turney, 2002). It uses a list of sentiment words and phrases to determine the overall sentiment of a document or sentence. The supervised learning approach involves building classifiers from labeled instances of documents or sentences (Mullen & Collier, 2004; Pang et al., 2002; Pang & Lee, 2008; Tan & Zhang, 2008). The latter approach could also be referred to as a machine-learning approach.

Previous studies (Zhang & Liu, 2011a, 2011b) demonstrated that if a large amount of labeled training data for a particular domain is available, the supervised learning approach usually generates superior performance because it can automatically consider domain-dependent sentiment expressions. On the other hand, the lexicon-based method cannot easily detect domain-dependent sentiment expressions without an algorithm capable of identifying such expressions and automatically classifying their sentiment polarities (Denecke, 2008; Moreo, Romero, Castro, & Zurita, 2012). However, supervised learning also has weaknesses. The main one is that a classifier trained in one domain usually does not work in other domains because words and even language constructs used to express opinions in different domains significantly vary (Ben-David, Blitzer, Crammer, & Sokolova, 2007; Glorot, Bordes, & Bengio, 2011). Above all, it is difficult to find large amounts of labeled data in the real world. News data are also not labeled. Supervised learning cannot apply to nonlabeled data, so unsupervised learning is often the only available option. For this reason, we apply the lexicon-based approach by constructing a sentiment lexicon dictionary of words annotated with semantic orientations. English has sentiment lexicon dictionaries, such as *SentiWordNet*, which has been widely used in various studies. However, Korean does not have such a dictionary, so we create it for analyses of this paper.

Dictionaries for lexicon-based approaches can be constructed manually, or automatically using seed words to expand the list of words (Turney, 2002). We use the former method. According to Taboada, Brooke, Tofiloski, Voll, and Stede (2011), manually created dictionaries are superior for two reasons. Firstly, they can exclude words with ambiguous meanings or words that convey a sentiment on some occasions but not on most. The probability of errors caused by the inner context of sentiment polarities varies depending on whether the polarity of each word is classified appropriately for the specific domain. Secondly, judicious restraint is necessary when expanding the dictionary. Taboada et al. (2011) indicated that adding more words to a dictionary does not always help performance because new words may add noise. On the other hand, one of the criticisms raised against the manual-based method is that the dictionaries are unreliable since they are annotated by humans (Andreevskaia & Bergler, 2008). In addition, it requires a large amount of time and effort to construct the dictionaries (Liu, 2012; Liu, 2015).

3.2 | Assessment

This section describes the methodology we used to investigate the relationship between NCSI and other macroeconomic indicators.

3.2.1 | Impulse response function

We compute a measure of NCSI shocks at different time lags using IRF. When predicting future economic situations, the most recent and past values of sentiment index are relevant. Demand-side and supply-side markets are closely related, so it is important to prioritize multivariate vector autoregression (VAR). However, the relatively short sample period of our analysis made it necessary to further reduce the number of factors to avoid an overfitting problem. We therefore estimate a simple autoregressive (AR) moving average (MA) model.[3] We

[3]In this study, we transformed all demand-side variables to the *y*/*y* growth rate in the previous year and used year-on-month changes for labor indicators.

consider the periods from 0 to 12 months and set an order of the AR and MA terms as two.

3.2.2 | Out of sample forecasting

Using the simple ARMA estimation explained in the previous section, we examine whether NCSI contains any predictive information about future economic conditions. To confirm a predictive power of the NCSI compared to that of CSI, a base model estimates an equation for the CSI as an explanatory variable. Then, we replace the CSI with the NCSI in a revised model. In the setting of ARMA(p, q) modeling we choose lag lengths of the model, with the lowest Akaike information criterion (AIC) being the best.[4] In the subsequent analysis, we conduct one-period-ahead out-of-sample forecasting. This method adds 1 month at a time, reestimating the model and calculating a series of forecasts for the current (nowcast) or following months (Vosen & Schmidt, 2011).

4 | EXPERIMENTAL SETUP AND DATA DESCRIPTION

4.1 | NCSI construction

NCSI construction consists of three broad steps:

1. collecting news articles and preprocessing the texts;
2. applying sentiment analysis for each article; and
3. generating a monthly time series index by aggregating each sentiment lexicon.

4.1.1 | Collection

As data collection takes a considerable amount of time, we limit the target to 11 daily general newspapers and nine daily economic newspapers provided by a news portal site. Similar to CSI, which surveys consumer attitudes regarding present situations and expectations for future business conditions, we collect news articles with titles containing the keywords "*household business condition*," "*household income*," or "*household expenditure*." We take this criterion because search keywords are commonly used terms rather than jargon. If we set the search condition as a keyword appearing in the body of text, many news articles unrelated to judging household's economic sentiment would be included in the dataset. We find that narrowing the scope of the target by setting the search condition to the title of articles greatly alleviates this problem. As a result, we collect 781 articles consisting of 14,080 sentences from January 2014 to March 2017.

4.1.2 | Preprocessing

In contrast to numeric data, text data require preprocessing, which is transforming unstructured data into structured data. In general, preprocessing consists of tokenization, stop-word removal, and stemming. Before tokenization, we first clean up special characters, English, and numbers in the news articles. Then, we conduct tokenization, which involves splitting each article into sentences and then segmenting each sentence into words. In stop-word removal step, we delete common words in the articles that are not specific or discriminatory to different polarity classes. Finally, we apply the stemming by removing a word's prefixes and suffixes. After the preprocessing procedure, we obtain a primary dataset of candidates to be registered in the sentiment lexicon dictionary.

4.1.3 | Sentiment lexicon dictionary construction

Previous research found that most sentiment lexicons are often expressed with adjectives, adverbs, and verbs (Benamara, Cesarano, Picariello, Reforgiato, & Subrahmanian, 2007; Brody & Elhadad, 2010; Cesarano et al., 2006; Hatzivassiloglou & McKeown, 1997; Hatzivassiloglou & Wiebe, 2000; Hu & Liu, 2004; Subrahmanian & Reforgiato, 2008; Taboada, Anthony, & Voll, 2006). Inspired by the results of these studies, we conduct morpheme analysis to tag part-of-speech (POS) information of each word.[5] Then, we manually annotate a semantic orientation of each word.[6] Words registered in the sentiment lexicon dictionary are used to generate the NCSI. The composition of the sentiment lexicon dictionary is illustrated in Table 1.

[4]There are several criteria for choosing the optimal lag lengths in a time series: AIC; Schwartcz information criterion (SIC); Hannan–Quinn criterion (HQ). The discrimination function differs from one criterion to another. According to related research, AIC is more appropriate when observations are less than 60 (Liew, 2004).

[5]In this study, we performed tagging using the Hannanum morphological analyzer. To be specific, we used the Simplepos09 code in the KoNLP package of R.

[6]Three researchers who majored in economics and have worked in national economic research institutes for more than 5 years performed this work.

TABLE 1 Composition of sentiment lexicon dictionary

Class	Number of words	Proportion	
Negative	5,385	21.5%	(64.4%)
Positive	2,981	11.9%	(35.6%)
Neutral	16,677	66.6%	(—)
Total words	25,043	100.0%	(100.0%)

4.1.4 | NCSI generation

The aggregation of individual surveys responded to economic conditions on five classes (very positive, positive, neutral, negative, very negative) generates the CSI (see Equation (2)). Some research classifies it as five classes (Socher et al., 2013). In this paper, we classify three classes as positive, negative, and neutral and then, generate the NCSI by aggregating the number of positive, neutral, and negative words (1 for a positive word, 0 for a neutral word, and − 1 for a negative word).[7] The current approach to constructing a new sentiment index greatly reduces its variance, and its effect diminishes the usefulness of the NCSI as an economic indicator. This finding indicates that applying existing sentiment analysis methodologies in generating economic indicators has its limitations. Therefore, it is necessary to develop either a more in elaborate approach or an innovative use of an existing method.

$$\mathrm{NCSI} = \frac{\sum[\text{positive word} \times 1 + \text{neutral word} \times 0 + \text{negative word} \times (-1)]}{\sum(\text{total positive} + \text{neutral} + \text{negative words})} \tag{1}$$

In comparison with other economic indicators, we create it monthly, but it can be expanded to various frequencies (e.g., weekly or even daily) depending on the data point of aggregation. The number of news articles collected in each month and the number of words appeared in each article is different, but we do not perform normalization since the difference is adjusted during the calculation process of the index.

$$\mathrm{CSI} = \frac{\sum[\text{very positive} \times 1 + \text{positive} \times 0.5 + \text{neutral} \times 0 + \text{negative} \times (-0.5) + \text{very negative} \times (-1)]}{\sum \text{total responses}} \tag{2}$$

4.2 | Variable description

4.2.1 | Consumer sentiment index

The samples for the CSI are statistically designed to be representative of domestic households.[8] The questionnaire of the CSI includes three separate headline figures: one for how people feel currently, one for how they feel the general economy is going, and one for how they expect things will be in 6 months. The CSI interprets reference values of less than 100 (the long-term average) as pessimistic and values greater than 100 are more optimistic than the long-term average. This means that, regardless of their current financial situation, when consumers feel confident about overall economic conditions and their personal financial futures, they are more likely to purchase greater amounts of goods and services and vice versa.

4.2.2 | Demand-side variables

Consumer goods, including a wide range of retail products and services, reflect the demand side. While overall demand for nondurable goods such as food and necessaries is not likely to fluctuate wildly, the level of durable-goods spending on optional purchases, such as automobiles and electronics, varies greatly depending on economic conditions. Consumers' purchasing decisions regarding durables, nondurables, and services are influenced by different motives; thus one may expect that consumer sentiment does not equally affect a decision of purchasing each consumption (Ludvigson, 2004; Throop, 1992). For confirming this, we select durables and nondurables indices as well as a retail sales index. In addition, a service industry production index includes various service industries, so we choose certain service items which are directly related to consumer economic sentiment. These

[7]Although not presented in this paper, we found out that the trend of NCSI was very similar including or excluding neutral words (a correlation coefficient of 0.99). Therefore, we included neutral words in our NCSI in order to generate a sentiment index in a manner similar to how the existing CSI was constructed, and then conducted the statistical analysis in Section 5.

[8]In the middle of each month, a survey of 2,200 city households from across the nation is conducted.

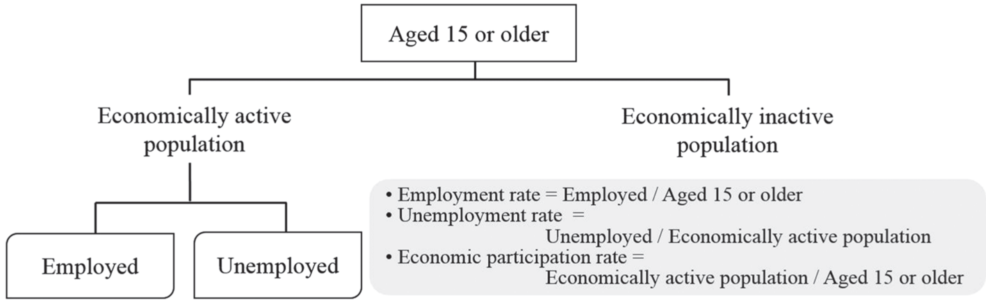

FIGURE 2 Composition of labor indicators

indices are announced at the end of each month by the National Statistical Office.

4.2.3 | Supply-side variables

Employment and unemployment levels largely affect the consumption and consumer sentiment. When employment rates increase, consumers begin to spend more and consumer confidence tends to increase. If unemployment rates rise, the opposite occurs. Hence economic sentiment and employment/unemployment rates are inextricably linked.

In South Korea, the National Statistical Office announces an employment trends report at the end of each month. The main items are economic participation, employment, and unemployment rates (see Figure 2). The employment rate is the proportion of employed in the population aged 15 and over. The unemployment rate is the proportion of unemployed in the economically active population (consisting of employed and unemployed). The rest of the population aged 15 and over is an economically inactive population. The classification of those without a job as economically active or economically inactive depends on whether they are actively seeking work.

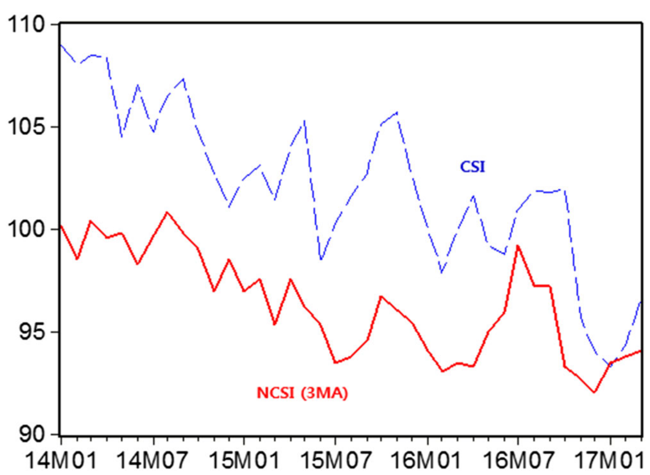

FIGURE 3 Trends of NCSI and CSI [Colour figure can be viewed at wileyonlinelibrary.com]

5 | EMPIRICAL RESULTS

5.1 | NCSI versus CSI

The public's main sources of information on the economy are news media, so economic news affects the sentiment of economic agents such as households, firms, and governments. Doms and Morin (2004) and Starr (2012) investigated whether news has an effect on consumer sentiment and showed that it plays a role in fluctuations in consumer sentiment. In this paper, we compare the trends of NCSI and CSI to look at the relationship between the two indices. Specifically, we use the 3-month average (3MA) for the NCSI (see Figure 3). We notice that two series are strongly correlated (with a correlation coefficient of 0.63). Next, we seek to relate pairs of time series through assessments of their cross-correlations, and the correlation coefficient is the highest for a lag of 1–2 months (see Figure 4). In addition, the NCSI Granger-causes the CSI, and the two have a cointegrating relationship in the Johansen cointegration test.[9]

5.2 | Granger causal relationship

The extent to which the economic sentiment index offers relevant, timely insights regarding economic conditions can be translated into a Granger causality framework. A (set of) time series is said to Granger-cause another time series if the former has incremental predictive power for predicting the latter (Wilms, Gelper, & Croux, 2016). In this section, we discuss the Granger causality test results to confirm that our newly created index has causality and predictive power for key economic indicators.

[9]We conducted an ADF test and PP test to examine the presence of a unit root for the CSI and the NCSI. The test results indicate that the two series have unit root.

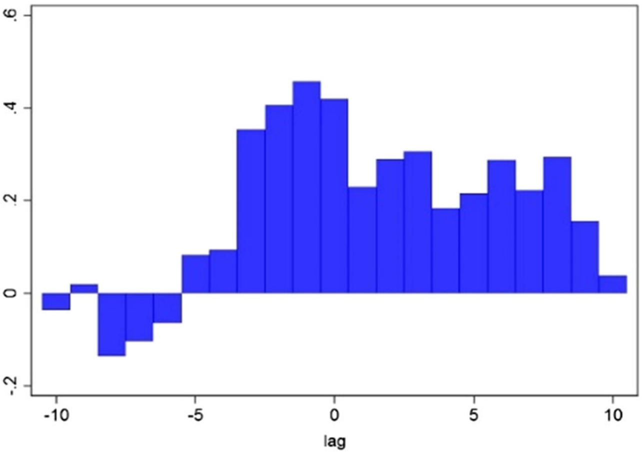

FIGURE 4 Cross-correlation of NCSI and CSI [Colour figure can be viewed at wileyonlinelibrary.com]

Table 2 summarizes the results of the Granger causality tests for the demand-side variables. The NCSI and CSI show no causality with respect to the retail sales index and the service industry production index. Conversely, the NCSI shows a significant causal relationship with the indicators that segmented the characteristics of consumption. These results are consistent with the discussion in Section 4.2, in which we mentioned that the impact of economic sentiment on goods and services differs greatly because the purposes for consumption vary. However, there is no Granger causality for the CSI on the demand-side variables, except for educational services.

TABLE 4 Results of granger causality test for supply-side variables

Employment rate does not granger-cause NCSI	**3.063†**	**(0.061)**
Employment rate does not granger-cause unemployment rate	**7.065****	**(0.003)**
Economic participation rate does not granger-cause unemployment rate	**7.138****	**(0.003)**

Note. Symbols denote significance level at †10% and. **1%.

Next, we investigate the Granger causal relationship between sentiment index and labor indicators (see Table 3). The NCSI shows a significant effect on the rates of economic participation and unemployment. However, as with the demand side, the CSI does not have a Granger cause on those variables except for the unemployment rate.

One notable thing is that neither NCSI nor CSI is significant for the employment rate. Meanwhile, the employment rate has a predictive power on the NCSI and the unemployment rate. In addition, the economic participation rate only has a causal influence on the unemployment rate (see Table 4). Figure 5 summarizes the Granger causal relationship between NCSI and labor indicators. We think that this relationship is closely related to the definition of labor indicators.

Suppose that the population over 15 years of age has no significant change and that only increases/decreases in employed/unemployed are possible. By definition,

TABLE 2 Result of granger causality test for demand-side variables

Null hypothesis: NCSI/CSI does not granger-cause *Y*		**NCSI**		**CSI**	
		***F*-stat.**	**(Prob.)**	***F*-stat.**	**(Prob.)**
Demand-side variables (*Y*)	Retail sales index	1.532	(0.232)	0.271	(0.764)
	Durables index	**3.385***	**(0.046)**	0.098	(0.906)
	Nondurables index	**2.577†**	**(0.092)**	0.438	(0.649)
	Service production index	0.578	(0.567)	0.375	(0.691)
	Accommodation and foods service index	**3.001†**	**(0.064)**	0.518	(0.601)
	Education service index	**4.575***	**(0.018)**	**5.974****	**(0.006)**

Note. Symbols denote significance level at **1%, *5%, and. †10%. The bold indicates that the F-statistic is statistically significant.

TABLE 3 Result of granger causality test for supply-side variables

Null hypothesis: NCSI/CSI does not granger-cause *Y*		**NCSI**		**CSI**	
		***F*-stat.**	**(Prob.)**	***F*-stat.**	**(Prob.)**
Supply-side variables (*Y*)	Economic participation rate	**3.849***	**(0.032)**	0.294	(0.747)
	Employment rate	1.264	(0.296)	0.292	(0.749)
	Unemployment rate	**6.132****	**(0.006)**	**3.557***	**(0.040)**

Note. Asterisks denote the significance level at **1% and. *5%. The bold indicates that the F-statistic is statistically significant.

FIGURE 5 Granger causal relationship between the NCSI and the labor indicators

these indicators determine employment status (new employment or retirement or dismissal) and classify people without a job as unemployed (i.e., belonging to the economically active population) or economically inactive population depending on whether they are actively seeking work. Therefore, there are two major cases in which the unemployment rate and the economic participation rate change. In the first case, the economically inactive population becomes the economically active population if people intend to get a job. This is possible even if there is no change in employed. A firm's hiring decision is an external factor that cannot be decided by households in the short term; an individual's intention to get a job determines changes in labor indicators, which could be reflected in the NCSI. Therefore, the NCSI can be the significant leading indicator for the unemployment rate and economic participation rate. In the second case, the unemployment rate and the economic activity participation rate change when the unemployed or economically inactive populations are employed and vice versa. Therefore, the employment rate has a causal influence on both unemployment and economic participation rates. In addition, when the employment rate increases (decreases), individuals see the economic optimistically (pessimistically), which has a Granger cause for the NCSI.

5.3 | Time structure of consumer sentiment shock

Here, we discuss IRF analysis results, which examined the lagged structure of the impact of sentiment shocks on economic indicators.[10] We consider innovations of one standard deviation rather than one-unit shocks because the variables have different scales (Lütkepohl, 2013). We only report the results the indicators of the demand side and the supply side showed as statistically significant, because the impulse responses are zero if one of the variables does not Granger-cause the other variables.

As shown in Figure 6, NCSI has a stronger impact on durable goods than on spending on nondurable goods. Apart from these results, durables, nondurables, accommodation, and food services respond strongly to sentiment index shocks (3–6). Meanwhile, the response of education service is less than 1. This implies that the effect of sentiment index on educational spending is relatively inelastic compared to other goods and services.

Next, Figure 7 shows the impulse responses of supply-side variables. NCSI shocks have long-lasting effects on economic participation and unemployment rates compared to the demand-side variables. This means that the NCSI has longer-lasting effects on the supply side, and adjustments in the labor market take place over a long period.

5.4 | Forecasting

This section discusses the forecasting power of NCSI for major economic indicators. The coefficient of the CSI for the retail sales index proves to be insignificant, whereas the NCSI proves significant (see Table 5).

The coefficient of NCSI is positive at 0.122, which is intuitively valid. The more optimistically consumers perceive economic conditions, the more consumption increases. Although it is not significant, the coefficient of the CSI is also positive and low at 0.072. In addition, the R^2 (higher), root mean square error (RMSE; lower), and mean absolute error (MAE; lower)—the criteria for determining the improvement of predictive power—show better results in the model using NCSI as an explanatory variable. These results correspond to the results of the above Granger causality test. On the other hand, neither sentiment index has a significant effect on the service industry production index, because the service industry includes various types of industries that are separated from households' economic sentiment.

[10]Before applying ordinary least squares to Equation 3, we conduct an ADF test and PP test to examine the presence of a unit root for the variables used for estimation. The test results indicate that demand- and supply-side variables are stationary, so they do not need any adjustments. Meanwhile, the CSI and the NCSI series need to be differentiated by an order of 1 to become stationary. We conducted these adjustments of experimental variables equally for the ARMA estimation in Section 5.4.

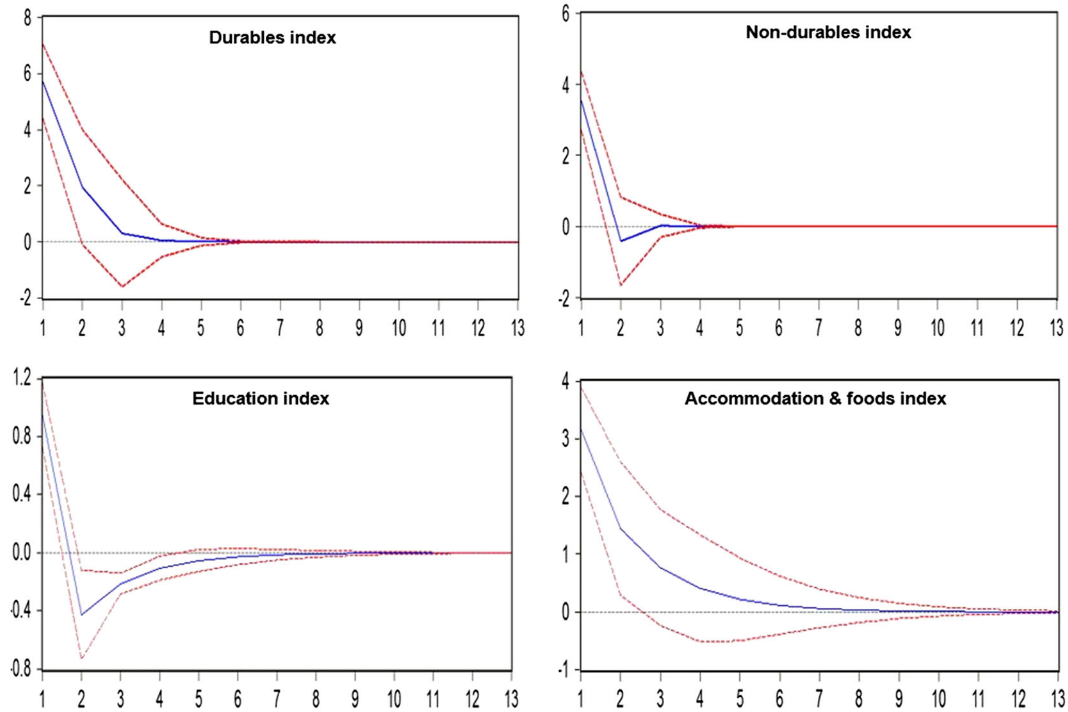

FIGURE 6 Impulse response function of NCSI shock to demand side variables [Colour figure can be viewed at wileyonlinelibrary.com]

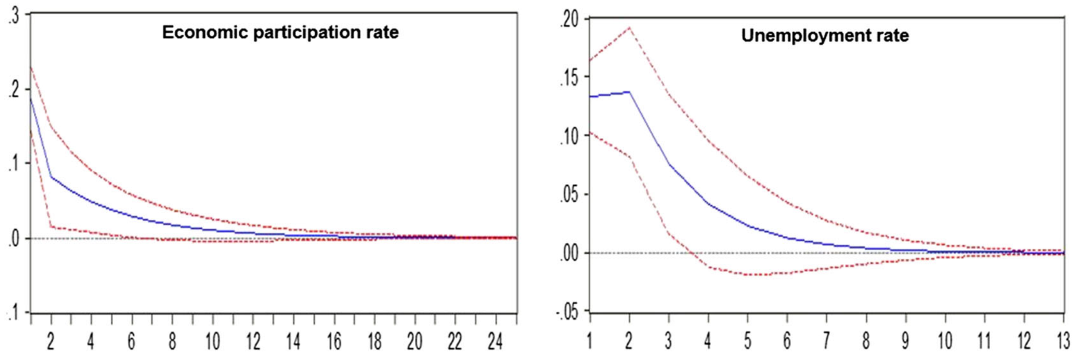

FIGURE 7 Impulse response function of NCSI shock to supply side variables [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 5 Estimation results for demand-side variables

	Retail sales index: MA(2)				Service production index: MA(2)			
	Coeff.	(Prob.)	Coeff.	(Prob.)	Coeff.	(Prob.)	Coeff.	(Prob.)
NCSI	—	—	**0.122**†	**(0.096)**	—	—	−0.009	(0.675)
CSI	0.072	(0.645)	—	—	0.031	(0.592)	—	—
R^2	0.208		0.260		0.084		0.080	
RMSE	2.256		2.179		0.726		0.754	
MAE	1.699		1.624		0.545		0.576	

Note. In the equation, AR term, MA term and constant term are included, but those estimation results are omitted in the table. Symbol denotes significance level at †10%.

Table 6 shows estimates of consumer goods divided into durable and nondurable goods. Similar to the results of the retail sales index, those of the model using NCSI as a predictor show an improved fitness and predictability over the model using CSI as an explanatory variable.

TABLE 6 Estimation results for durable and nondurable goods

	Durables index: ARMA(1, 1)				Nondurables index: ARMA(1, 1)			
	Coeff.	**(Prob.)**	**Coeff.**	**(Prob.)**	**Coeff.**	**(Prob.)**	**Coeff.**	**(Prob.)**
NCSI	—	—	**−0.305**†	**(0.055)**	—	—	**0.287***	**(0.034)**
CSI	0.367	(0.351)	—	—	−0.068	(0.798)	—	—
R^2	0.113		0.143		0.042		0.159	
RMSE	5.593		5.382		3.668		3.342	
MAE	4.653		4.314		2.682		2.481	

Note. In the equation, AR term, MA, term and constant term are included, but those estimation results are omitted in the table. Symbols denote significance level at †10% and *5%.

TABLE 7 Estimation results for supply-side variables

	Employment rate: MA(2)				Unemployment rate: ARMA(1, 1)			
	Coeff.	**(Prob.)**	**Coeff.**	**(Prob.)**	**Coeff.**	**(Prob.)**	**Coeff.**	**(Prob.)**
NCSI	—	—	**−0.006***	**(0.011)**	—	—	**−0.006****	**(0.008)**
CSI	0.014	(0.234)	—	—	0.004	(0.628)	—	—
R^2	0.564		0.591		0.562		0.646	
RMSE	0.212		0.205		0.139		0.126	
MAE	0.178		0.169		0.116		0.107	

Note. In the equation, AR term, MA term and constant term are included, but those estimation results are omitted in the table. Asterisks denote the significance level at **1% and. *5%.

Apart from these results, a regression coefficient sign is quite interesting. The coefficients of CSI and NCSI have opposite sign for durables and nondurables. In considering the features of durables and nondurables, the coefficient of NCSI appears more reasonable than that of CSI. By their nature, durable goods tend to be more expensive and can be used repeatedly, providing consumers with a flow of services over a number of years. Thus the decision to purchase durable goods requires some confidence in the economic outlook. Prior to the recession, households increased their purchases of durable goods and built up a stock of durables. With the onset of the recession, the worsening employment outlook, and decreases in income and wealth, households may have been more cautious when deciding whether to purchase additional durables. In other words, when it comes to durable goods, consumers require adequate time and money to make purchasing decisions. Meanwhile, nondurable goods have no time lags regarding purchasing decisions because of their nature, and individuals respond promptly to their financial conditions.

Lastly, Table 7 presents estimates of the employment and unemployment rates. The results of the model using NCSI as an explanatory variable improve in terms of fitness and predictability compared with the model using CSI as a predictor. The coefficient value of NCSI for unemployment and employment rates is very low due to the time structure of the impact of NCSI shocks. As seen in Section 5.3, shocks on durables and nondurables are largest in the first 1–2 months and most of the shocks disappear after 3 months. On the other hand, the relatively small impact on the supply side continues for more than 6 months. Therefore, estimates on the supply-side

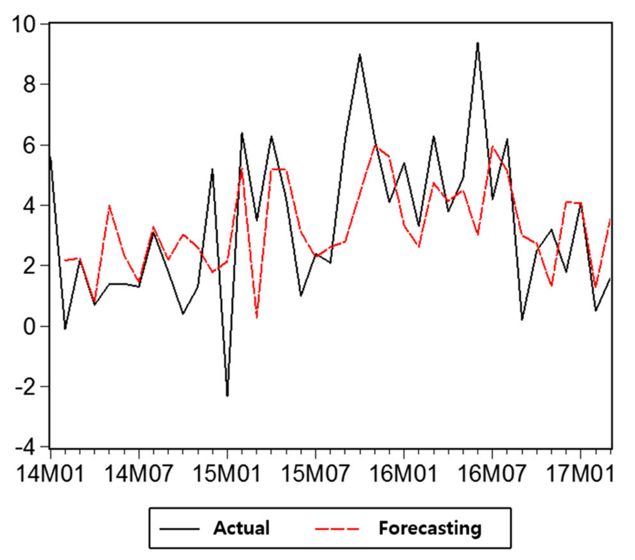

FIGURE 8 Retail sales index [Colour figure can be viewed at wileyonlinelibrary.com]

variables are relatively low compared to those on the demand-side variable.

Overall, the NCSI performs better than the CSI on the Granger causality and forecasting test, although the two indices are closely related. We think there are two reasons for this result. Firstly, as mentioned in Section 5.1, the NCSI could be a leading indicator for the CSI. If individual attitudes regarding present situations and expectations for future business conditions are largely influenced by news media and sentiment changes alter their economic decision (e.g., spending), news data have a higher predictive power for other economic indicators. Starr (2012) demonstrated that consumer sentiment is significantly affected by "economic news shocks" and news shocks significantly affect aggregate economic activity. Secondly, news articles can better explain household's economic conditions rather than individual subjective judgment, because experts analyze a financial phenomenon based on available facts.

Based on the above ARMA estimation, Figures 8–13 compare actual values with the results of the one-step-ahead-of-the-sample prediction. Forecasting values for the retail sales, durables, and nondurables indices show similar trends to the actual values. In addition, prediction

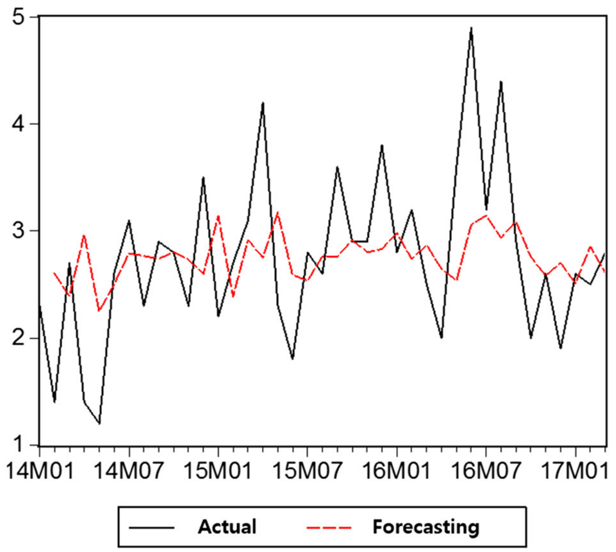

FIGURE 9 Service production index [Colour figure can be viewed at wileyonlinelibrary.com]

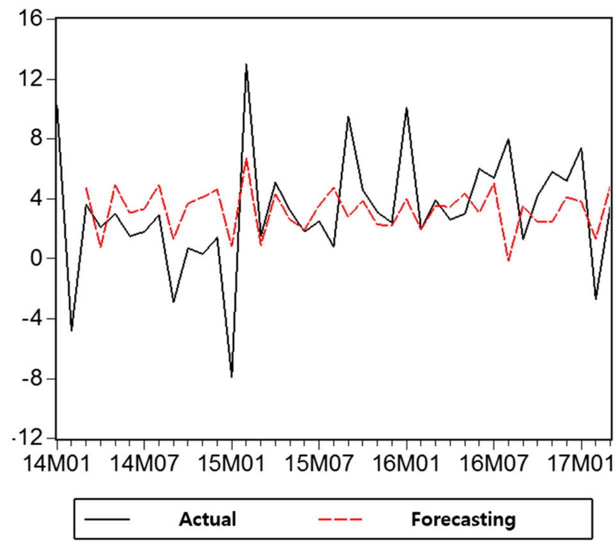

FIGURE 11 Nondurables index [Colour figure can be viewed at wileyonlinelibrary.com]

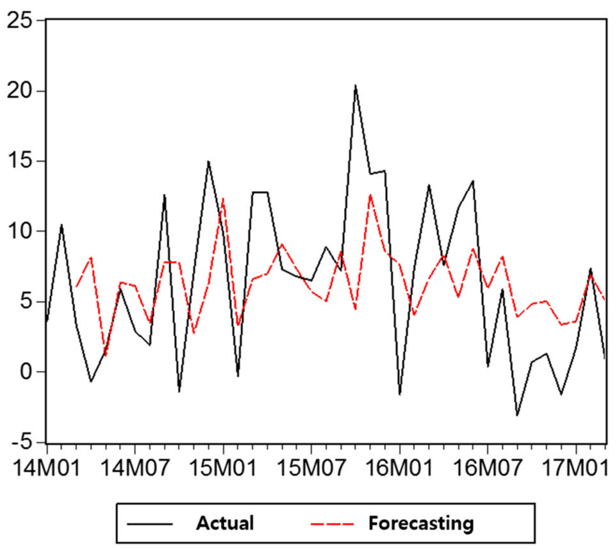

FIGURE 10 Durables index [Colour figure can be viewed at wileyonlinelibrary.com]

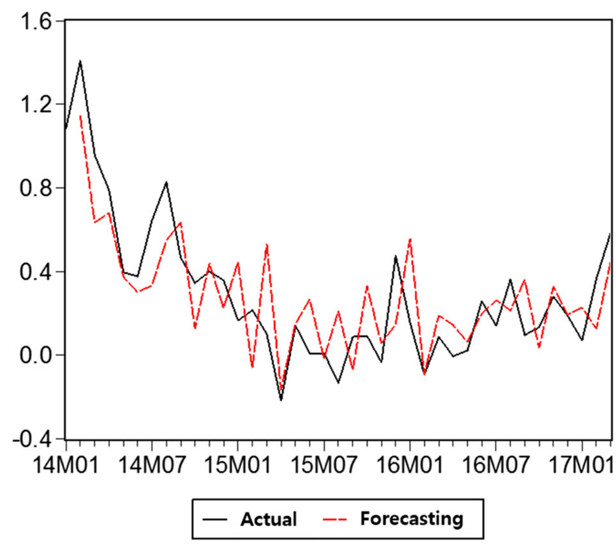

FIGURE 12 Employment rate [Colour figure can be viewed at wileyonlinelibrary.com]

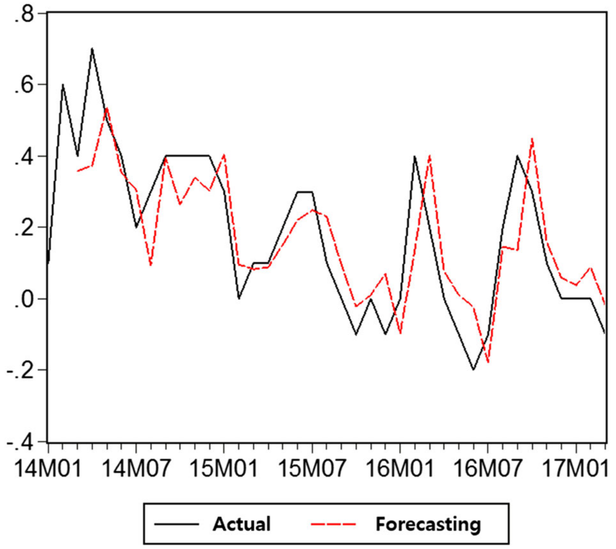

FIGURE 13 Unemployment rate [Colour figure can be viewed at wileyonlinelibrary.com]

values for labor indicators are very similar to the actual data; thus NCSI can provide useful information regarding trends and outlooks in the labor market. However, the service industry production index will likely need to consider other variables.

6 | CONCLUSIONS

Given the confirmed effectiveness of the survey-based CSI as a leading indicator of real economic conditions, the CSI is actively used in making policy judgments and decisions in many countries. However, although the CSI offers qualitative information for presenting current conditions and predicting a household's future economic activity, the survey-based method has several limitations. In this context, we constructed the NCSI using sentiment analysis of news articles to complement the CSI.

By applying a simple sentiment analysis based on the lexicon approach, we demonstrated that news articles can be an effective source for generating an economic indicator in Korea. To check the usefulness of NCSI, we conducted a Granger causality test, IRF analysis, and out-of-sample prediction. The results of the works are summarized as follows:

1. We found out that NCSI and CSI have a cointegrating relationship, and NCSI could be a leading indicator for the survey-based measure.
2. NCSI showed Granger causality and predictive power for demand-side and supply-side economic indicators. In addition, NCSI outperformed the CSI in most cases.
3. NCSI had a high predictive power for labor indicators because people's intentions of getting jobs can be reflected in the NCSI.

This paper does not attempt a full linguistic analysis that involves analyses of word senses or argument structures; this is a limitation of our research, and further work in that direction remains possible. The proposed method of generating a new sentiment index vastly decreases its variance; it has a negative effect on the performance of the NCSI as an economic indicator. Therefore, future studies should focus on designing a more elaborate approach or on the innovative use of an existing approach.

Textual data such as news articles and social networks (e.g., Twitter, Facebook, and blogs), are generated at high speeds and cover a wide range of topics. Such sources can quickly capture an economic impact of specific economic issues; hence they can be great potential sources for generating economic indicators. Although few previous studies have applied such unstructured data in economic analysis, various applications that use text mining techniques will see significant growth if their usefulness is confirmed. We hope to encourage further cross-national comparative research to apply the approach proposed in this study.

ACKNOWLEDGMENTS

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors declare no conflicts of interest.

REFERENCES

Acemoglu, D., & Scott, A. (1994). Consumer confidence and rational expectations: Are agents' beliefs consistent with the theory. *Economic Journal*, *104*(422), 1–19.

Andreevskaia, A., & Bergler, S. (2008). When specialists and generalists work together: Domain dependence in sentiment tagging. In *Proceedings of 46th Annual Meeting of Association for Computational Linguistics* (pp. 290–298). Stroudsburg, PA: ACL.

Angeletos, G. M., & La'O, J. (2013). Sentiments. *Econometrica*, *81*(2), 739–779.

Arias, A. (2016). *Sentiment shocks as drivers of business cycles (Working Paper 782)*. Santiago, Chile: Central Bank of Chile.

Balahur, A., & Steinberger, R. (2009). Rethinking sentiment analysis in the news: From theory to practice and back. In *Proceedings of the 1st Workshop on Opinion Mining and Sentiment Analysis* (pp. 1–12). Hanover, MA: Now.

Balahur, A., Steinberger, R., Kabadjov, M., Zavarella, V., van der Goot, E., Halkia, M., … Belyaeva, J. (2010). Sentiment analysis in the news. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, et al. (Eds.), *Proceedings of the 7th International Conference, Language Resources and Evaluation* (pp. 2216–2220). Paris, France: European Language Resources Association.

Balahur, A., Steinberger, R., van der Goot, E., Pouliquen, B., & Kabadjov, M. (2009). Opinion mining on newspaper quotations. In *Proceedings of the Workshop Intelligent Analysis and Processing of Web News Content* (pp. 523–526). Washington, DC: IEEE Computer Society.

Barsky, R. B., & Sims, E. R. (2012). Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review*, *102*(4), 1343–1377.

Benamara, F., Cesarano, C., Picariello, A., Reforgiato, D., & Subrahmanian, V. S. (2007). Sentiment analysis: Adjectives and adverbs are better than adjectives alone. In *Proceedings of International Conference on Web and Social Media* (pp. 203–206). Menlo, Park, CA: International Joint Conferences on Artificial Intelligence.

Ben-David, S., Blitzer, J., Crammer, K., & Sokolova, P. M. (2007). Analysis of representations for domain adaptation. In B. Schölkopf, J. C. Platt, & T. Hoffman (Eds.), *Advances in Neural Information Processing Systems* (pp. 137–144). La Jolla, CA: NIPS Foundation.

Benhabib, J., Wang, P., & Wen, Y. (2015). Sentiments and aggregate demand fluctuations. *Econometrica*, *83*(2), 549–585.

Benhabib, J., & Wen, Y. (2004). Indeterminacy, aggregate demand, and the real business cycle. *Journal of Monetary Economics*, *51*(3), 503–530.

Bram, J., & Ludvigson, S. C. (1997). *Does consumer confidence forecast household expenditure? A sentiment index horse race (Research Paper 9708)*. New York, NY: Federal Reserve Bank of New York.

Brody, S., & Elhadad, N. (2010). An unsupervised aspect–sentiment model for online reviews. In *Human Language Technologies: 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 804–812). Stroudsburg, CA: ACL.

Cesarano, C., Dorr, B., Picariello, A., Reforgiato, D., Sagoff, A., & Subrahmanian, V. S. (2006). Oasys: An opinion analysis system. In *Proceedings of the Association for the Advancement of Artificial Intelligence, 06 Spring Symposium: Computational Approaches to Analyzing Weblogs* (pp. 21–26). Palo Alto, CA: AAAI Press.

Chadha, K., & Wells, R. (2016). Journalistic responses to technological innovation in newsrooms. *Digital Journalism*, *4*(8), 1020–1035.

Daas, P. J., & Puts, M. J. (2014). Social media sentiment and consumer confidence (ECB Statistics Paper, No. 5). Frankfurt, Germany: European Central Bank.

Daniel, M., Neves, R. F., & Horta, N. (2017). Company event popularity for financial markets using Twitter and sentiment analysis. *Expert Systems with Applications*, *71*(1), 111–124.

Denecke, K. (2008). Using SentiWordNet for multilingual sentiment analysis. In *Proceedings of the IEEE 24th International Conference on Data Engineering workshop* (pp. 507–512). Washington, DC: IEEE Computer Society.

Doms, M., & Morin, N. (2004). *Consumer sentiment, the economy, and the news media* (FEDS Working Paper. No. 2004–51. Washington, DC: Federal Reserve Board of Governors.

Farmer, R., & Guo, J. T. (1994). Real business cycles and the animal spirits hypothesis. *Journal of Economic Theory*, *63*, 42–72.

Gelper, S., Lemmens, A., & Croux, C. (2007). Consumer sentiment and consumer spending: decomposing the Granger causal relationship in the time domain. *Applied Economics*, *39*, 1–11.

Glorot, X., Bordes, A., & Bengio, Y. (2011). Domain adaptation for largescale sentiment classification: A deep learning approach. In L. Getoor, & T. Scheffer (Eds.), *Proceedings of the 28th International Conference on Machine Learning* (pp. 513–520). Madison, WI: Omnipress.

Godbole, N., Srinivasaiah, M., & Skiena, S. (2007). Large-scale sentiment analysis for news and blogs. In *Proceedings of the International Conference on Weblogs and Social Media* (Vol. 7. No. 21, pp. 219–222). Boulder, Colorado: International Conference on Weblogs and Social Media(ICWSM).

Hatzivassiloglou, V., & McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics* (pp. 174–181). Stroudsburg, PA: ACL.

Hatzivassiloglou, V., & Wiebe, J. M. (2000). Effects of adjective orientation and gradability on sentence subjectivity. In *Proceedings of the 18th International Conference on Computational Linguistics, COLING* (Vol. 1, pp. 299–305). Stroudsburg, PA: ACL.

Howrey, E. (2001). The predictive power of the index of consumer sentiment. *Brookings Papers on Economic Activity*, *32*(1), 175–216.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 168–177). New York, NY: ACM.

Kang, H., Yoo, S. J., & Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications*, *39*(5), 6000–6010.

Kontopoulos, E., Berberidis, C., Dergiades, T., & Bassiliades, N. (2013). Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications*, *40*(10), 4065–4074.

Liew, K. S. (2004). What lag selection criteria should we employ? *Economics Bulletin*, *33*(3), 1–9.

Liu, B. (2012). Sentiment analysis and opinion mining. In *Synthesis lectures on human language technologies* (Vol. 5, No. 1) (pp. 1–167). San Rafael, CA: Morgan & Claypool.

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge, UK: Cambridge University Press.

Ludvigson, S. C. (2004). Consumer confidence and consumer spending. *Journal of Economic Perspectives*, *18*(2), 29–50.

Lütkepohl, H. (2013). *Introduction to multiple time series analysis*. Berlin, Germany: Springer.

Martín-Valdivia, M. T., Martínez-Cámara, E., Perea-Ortega, J. M., & Ureña-López, L. A. (2013). Sentiment polarity detection in Spanish reviews combining supervised and unsupervised approaches. *Expert Systems with Applications*, *40*, 3934–3942.

Milani, F. (2017). Sentiment and the US business cycle. *Journal of Economic Dynamics and Control*, *82*, 289–311.

Moreo, A., Romero, M., Castro, J. L., & Zurita, J. M. (2012). Lexicon-based comments-oriented news sentiment analyzer system. *Expert Systems with Applications*, *39*, 9166–9180.

Mullen, T., & Collier, N. (2004). Sentiment analysis using support vector machines with diverse information sources. In *Proceedings of the 2004 Conference on Empirical Methods on Natural Language Processing Association for Computational Linguistics* (pp. 412–418). Stroudsburg, PA: ACL.

Nguyen, T. H., Shirai, K., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, *42*, 9603–9611.

Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications*, *73*, 125–144.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1–2), 1–135.

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing* (Vol. 10, pp. 79–86). Stroudsburg, PA: ACL.

Perlin, M. S., Caldeira, J. F., Santos, A. A. P., & Pontuschka, M. (2017). Can we predict the financial markets based on Google's search queries. *Journal of Forecasting*, *36(4*, 454–467.

Schumaker, R. P., Zhang, Y., Huang, C. N., & Chen, H. (2012). Evaluating sentiment in financial news articles. *Decision Support Systems*, *53*(3), 458–464.

Shapiro, A. H., Kanjoya, M. S., & Wilson, D. (2017). *Measuring news sentiment (Working Paper 2017–01)*. San Francisco: CA Federal Reserve Bank.

Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C., Ng, A., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of Conference on Empirical Methods in Natural Language Processing* (pp. 1631–1642). Stroudsburg, PA: ACL.

Starr, M. A. (2012). Consumption, sentiment, and economic news. *Economic Inquiry*, *50*(4), 1097–1111.

Steensen, S. (2011). Online journalism and the promises of new technology. *Journalism Studies*, *12*(3), 311–327.

Subrahmanian, V. S., & Reforgiato, D. (2008). Ava: Adjective–verb–adverb combinations for sentiment analysis. *Intelligent Systems*, *23*(4), 43–50.

Taboada, M., Anthony, C., & Voll, K. (2006). Creating semantic orientation dictionaries. In *Proceedings of 5th International Conference on Language Resources and Evaluation* (pp. 427–432). Paris, France: European Language Resources Association.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, *37*, 267–307.

Tan, S., & Wu, Q. (2011). A random walk algorithm for automatic construction of domain-oriented sentiment lexicon. *Expert Systems with Applications*, *38*, 12094–12100.

Tan, S., & Zhang, J. (2008). An empirical study of sentiment analysis for Chinese documents. *Expert Systems with Applications*, *34*, 2622–2629.

Tang, H., Tan, S., & Cheng, X. (2009). A survey on sentiment detection of reviews. *Expert Systems with Applications*, *36*(7), 10760–10773.

Throop, A. W. (1992). Consumer sentiment: Its causes and effects. *Federal Reserve of San Francisco Review*, *1*, 35–39.

Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (pp. 417–424). Stroudsburg, PA: ACL.

Uhl, M. W. (2011). Explaining US consumer behavior with news sentiment. *ACM Transactions on Management Information Systems*, *2*(2), 9–18.

Ulbricht, D., Kholodilin, K. A., & Thomas, T. (2017). Do media data help to predict German industrial production? *Journal of Forecasting*, *36(5*, 483–496.

Van de Kauter, M., Breesch, D., & Hoste, V. (2015). Fine-grained analysis of explicit and implicit sentiment in financial news articles. *Expert Systems with Applications*, *42*, 4999–5010.

Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: Survey-based indicators vs. Google Trends. *Journal of Forecasting*, *30*, 565–578.

Wilms, I., Gelper, S., & Croux, C. (2016). The predictive power of the business and bank sentiment of firms: A high-dimensional Granger Causality approach. *European Journal of Operational Research*, *254*(1), 138–147.

Yu, L. C., Wu, J. L., Chang, P. C., & Chu, H. S. (2013). Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news. *Knowledge-Based Systems*, *41*, 89–97.

Zhang, L., & Liu, B. (2011a). Extracting resource terms for sentiment analysis. In *Proceedings of the 5th International Joint Conference on Natural Language Processing* (pp. 1171–1179). Chiang Mai, Thailand: Asian Federation of Natural Language Processing.

Zhang, L., & Liu, B. (2011b). Identifying noun product features that imply opinions. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics* (pp. 575–580). Stroudsburg, PA: ACL.

AUTHOR BIOGRAPHIES

Minchae Song received a Ph.D. in Big Data Analytics at Ewha Womans University and a M.A. degree in Economics at Ewha Womans University. Her research interests focus on the machine learning algorithms, natural language processing, sentiment analysis and text mining.

Kyung-shik Shin is a professor of School of Business at Ewha Womans University. He received an undergraduate degree in Management from Yonsei University, an MBA degree in Management from George Washington University in Washington D. C, and his Ph.D. in Management Engineering at KAIST (Korea Advanced Institute of Science and Technology). His interests include the artificial intelligence applications, big data analytics, business intelligences, data mining, and text mining.

How to cite this article: Song M, Shin K. Forecasting economic indicators using a consumer sentiment index: Survey-based versus text-based data. *Journal of Forecasting*. 2019;38:504–518. https://doi.org/10.1002/for.2584