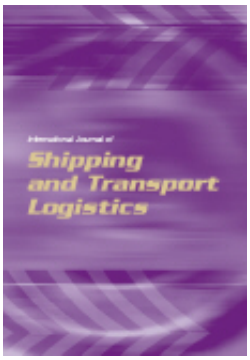
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Shipping and Transport Logistics

ISSN online: 1756-6525 - ISSN print: 1756-6517
https://www.inderscience.com/ijstl

Construction of Korean Shipping Industry Risk Sentiment Index using news articles

Sunghwa Park, Hyunsok Kim, Janghan Kwon, Taeil Kim

DOI: 10.1504/IJSTL.2023.10050213

Article History:

Received:	25 August 2021
Accepted:	06 July 2022
Published online:	15 January 2024

Copyright © 2023 Inderscience Enterprises Ltd.

Construction of Korean Shipping Industry Risk Sentiment Index using news articles

Sunghwa Park

Department of International Trade,
Gyeongsang National University,
501, 201dong, Jinju-daero,
Jinju-si, Gyeongsangnam-do, 52828, South Korea
Email: sh.park@gnu.ac.kr

Hyunsok Kim*

College of Economics and International Trade,
Pusan National University,
2, Busandaehak-ro 63beon-gil,
Geumjeong-gu, Busan, South Korea
Email: hyunsok.kim@pusan.ac.kr
*Corresponding author

Janghan Kwon

Ocean Economy Strategy Department,
Korea Maritime Institute,
26, Haeynag-ro 301beon-gil,
Yeongdo-gu, Busan, 49111, South Korea
Email: jkwon@kmi.re.kr

Taeil Kim

Maritime Policy Research Department,
Korea Maritime Institute,
26, Haeynag-ro 301beon-gil,
Yeongdo-gu, Busan, 49111, South Korea
Email: ktizorro@kmi.re.kr

Abstract: We develop the Shipping Industry Risk Sentiment Index (SRSI) to measure risk sentiment in Korea's shipping industry. The index is based on the frequency of news articles in Korean newspapers that contain terms related to the six C's of credit suggested by Grammenos (2013). The dynamics of the index show that it spikes near tight periods of elevated risk sentiment including global financial crisis and the court receivership of Korea's major shipping companies. We investigate predictive power of the index by using an autoregressive distributed lag (ARDL) model and impulse response functions from local projections. Empirical evidence shows that SRSI reflects risk sentiment in shipping industry and contains useful information to understand market situations.

Copyright © The Author(s) 2023. Published by Inderscience Publishers Ltd. This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Keywords: SRSI; news big data analysis; shipping market sentiment; local projections; ARDL.

Reference to this paper should be made as follows: Park, S., Kim, H., Kwon, J. and Kim, T. (2023) 'Construction of Korean Shipping Industry Risk Sentiment Index using news articles', *Int. J. Shipping and Transport Logistics*, Vol. 17, No. 4, pp.469–486.

Biographical notes: Sunghwa Park is a Professor in the Department of International Trade at Gyeongsang National University. He holds a PhD in Economics and has conducted research on maritime economics, shipping finance, risk management, and overall econometric analysis of the shipping market and shipping policy. He served as Director of the Maritime Finance Research Division at the Korea Maritime Institute, Busan, South Korea.

Hyunsok Kim is a Professor of Economics at the Pusan National University. His research interest is on the economics, time series econometrics and shipping economics. He has been working in applied research projects of econometric modelling. His work has been published in several journals on economics and finance.

Janghan Kwon is a senior researcher of the Ocean Economy and Statistics Research Department at Korea Maritime Institute, Busan, Korea. He holds a PhD in Economics and studied economic forecasting of Korean ocean and shipping industry, shipping companies' risk management and big data analysis.

Taeil Kim is a Research Fellow of the Maritime Policy Research Department at Korea Maritime Institute, Busan, Korea. He holds a PhD in Economics and conducted many projects on maritime economics, shipping finance, risk management and overall econometric analysis in the shipping market and shipping policy.

1 Introduction

The Korean shipping industry has faced several crises over the years. Severe fluctuations occurred in the 1980s due to the economic depression caused by the oil shock, and significant difficulties characterised the shipping market during the International Monetary Fund (IMF) bailout period in the 1990s, as well as in the aftermath of the global financial crisis of 2008. Also, Major Korean shipping companies became vulnerable to credit risk. For instance, Hanjin Shipping went bankrupt in 2017, and Pan Ocean and Hyundai Merchant Marine went through restructuring (Choi et al., 2018).

The shipping industry is capital-intensive and largely depends on borrowed capital as a means of financing to secure ships. Most revenues are from the freight rates paid by shippers, and the profit structure of the shipping industry is peculiar. In addition, the financial structure of shipping companies is weaker than that of firms operating in the manufacturing industry.

Grammenos (2013) proposed that the risk assessment of shipping companies should reflect the unique characteristics of the shipping industry. The business activities of shipping companies can be divided into revenues from maritime transport (or shipping), which companies operating their ships, and revenues from sales and purchase of ships

(S&P, henceforth). The demand for maritime transport largely depends on the demand for global trade and the trade volume, thus reflecting the procyclical feature of the shipping industry. Increasing and decreasing assets through the S&P of ships have a significant impact on the financial structure of shipping companies depending on when and how ships are acquired. The shipping industry as a whole is influenced by the shipping business cycle and is significantly affected by both the risks associated with individual companies and systemic risks common to all companies. Therefore, diagnosing the default of shipping companies at an early stage and developing appropriate risk management tools are challenging tasks.

The shipping industry plays a vital role in the Korean economy. Several studies used the shipping industry crisis index and the signals approach to determine the soundness of the shipping industry (Chen et al., 2014; Lee and Kim, 2016; Ministry of Oceans and Fisheries, 2015). Moreover, the government has implemented policies to reinforce the monitoring system for individual companies and the whole market and prevent the increase of risks due to shipping business fluctuations. For instance, the Plan for Improving Competitiveness of the Shipping Industry was announced in October 2016. However, this early warning system based on the signals approach has received severe criticism. In particular, some argue that the variables selection and the threshold for the risk alert may be arbitrarily chosen. While the results of previous studies based on developing an early warning index may help predict a future crisis in the industry and provide useful information regarding the overall market conditions, severe limitations exist in these approaches. In particular, the determinants of shipping business fluctuations and the risks associated with individual companies have been insufficiently addressed in previous research. Thus, an alternative approach is needed to evaluate the risk of shipping companies.

Studies have also attempted to develop economic sentiment indexes by surveying economic units to predict the economy in each country. Moreover, many studies have determined and forecasted the economy by measuring uncertainty based on changes in economic variables. However, this method cannot capture economic units' concerns about the economy in advance. Accordingly, many recent attempts have been made to explain the economy by selecting keywords and indexing them through unstructured data, such as news articles.

The shipping market contains uncertainty as well, and the fact that this uncertainty affects freight rates and investments is discussed in the literature. However, despite repeated crises in the shipping industry, too few representative variables are available to detect risk in the shipping industry. Thus, to determine the market conditions of the Korean shipping industry, a representative variable is needed to measure market agents' judgments regarding credit risk in the shipping industry. The purpose of this study is therefore to develop a risk sentiment index shipping industry that measures overall credit risk in the Korean shipping industry and evaluate the usefulness of the index written.

This study is novel in that it constructs a risk sentiment index for the shipping industry in Korea through text mining and verifies the predictive power of this index. Additionally, the risks of the shipping market, the global business, are segmented and analysed based on specific information of Korea. Specifically, cross-correlation analysis between the shipping industry risk index and Korea's shipping industry survey index, autoregressive distributed lag (ARDL) Model analysis between Korea's seaborne transport freight revenue and the index, and impulse response analysis using local

projection between them are performed. Through this empirical analysis, this study scientifically verifies the relationship between the index constructed in this study and the index representing the Korean shipping industry, and provides a new approach for risk management in the Korean shipping industry.

The rest of paper is structured as follows. Section 2 presents literature review. Section 3 describes the data we use to construct the Shipping Industry Risk Sentiment Index (SRSI). Section 4 empirically evaluates the SRSI in several ways. Conclusions follow in Section 5.

2 Literature review

The judgment expressed by the agents about the state of the economy plays a significant role in determining business fluctuations (Bram and Ludvigson, 1997; Gelper et al., 2007).

A popular approach is to measure market sentiment as survey in the expectation of economic agents. In Korea, the typical index representing the economic sentiment of companies is the Business Survey Index issued by the Bank of Korea (http://ecos.bok.or.kr/), based on the judgments and forecasts of survey participants regarding the trend of economic performance and released every month. However, the Business Survey Index has fixed survey items and, therefore, it cannot help judgments in the case of a specific event. In addition, every month, the Korea Maritime Institute issues the Shipping Business Survey Index (SBSI), which determines the overall shipping business trends and forecasts by surveying shipping business participants about the current economic status and future prospects.[1] However, the SBSI is problematic to use in time series analysis due to the short data accumulation period.[2]

With this awareness, recent studies have attempted to explain the market trends through sentiment analysis using unstructured data such as big data (Aggarwal and Zhai, 2012; Balahur et al., 2010; Heston and Sinha, 2016; Li et al., 2016).

In the presence of substantial uncertainty in the market, the typical economic unit postpones decision-making and adopts a wait-and-see approach until the uncertainty is reduced. Cautiousness induces market participants to deal with market changes in a passive manner, and, as a result, businesses risk tends to degenerate into recession. Some theoretical models claim that uncertainty shapes the decision-making of economic units and is a crucial factor for explaining economic phenomena (Fudenberg and Maskin, 2009). Research on the macro shock effects under uncertainty has been limited because while the policy ineffectiveness proposition is well understood in theory, empirical evidence on its macroeconomic importance is limited (Aastveit et al., 2013). However, uncertainty has been emphasised as a critical determinant of economic recessions since the financial crisis, and research on uncertainty has flourished.

Bontempi et al. (2016) argue that two ways to measure uncertainty exist. First, the finance-based method was proposed, which turns the volatility of the stock market into a variable using a statistical technique (Baker et al., 2016; Bloom, 2009). Second, the forecast-based method was suggested, which measures uncertainty based on the movements of economic variables (Rich and Tracy, 2010; Scotti, 2016). Bloom (2009) argued that uncertainty affects employment and capital adjustment costs; thus, if uncertainty increases, firms temporarily discontinue employment and investment, and real variables rapidly decrease in the short run. However, both methods are subject to

substantial limitations when trying to anticipate the concerns of economic units about the market.

Recently, there is a rapidly growing study on text mining through big data – using newspaper archives, in particular – to measure a variety of outcomes. This approach relies on variables that are built and indexed focusing on certain words that appear in the news (Baker et al., 2016; Bontempi et al., 2016; Rossi and Sekhposyan, 2015; Tobback et al., 2018).

A typical news-based variable is the economic policy uncertainty (EPU) variable proposed by Baker et al. (2016). The EPU is built by dividing the number of articles (based on keywords set by the researcher) by the total number of economics articles published by each newspaper company during the same period and standardising the data by the standard deviation of each newspaper company. Finally, the monthly average is normalised by setting the average of the total period of analysis as 100. However, the EPU index is designed to convey the overall uncertainty of macroeconomics rather than identify the crisis of individual markets. Moreover, the results of empirical analyses show a strong negative relationship between corporate investment and uncertainty.

Researches to measure uncertainty and sentiment about the global economy are being actively conducted, but limited in sentiment measurement for the shipping industry.

Papapostolou et al. (2014) provides the foundations of measuring market confidence in the dry bulk shipping market by investigating the role of sentiment and its implications for real assets. They employ shipping sentiment proxies (liquidity and valuation) to construct sentiment indices for the dry bulk shipping market that captures market expectations. The empirical findings also show that market sentiment serves as an indicator for future cycle phases in all sectors. Papapostolou et al. (2016) extends the work to shipping investor sentiment and stock return predictability and shows that shipping investor sentiment is a key predictor for financial markets. Park et al. (2018) develops the maritime news index that can measure the risk of the maritime industry in Korea using Korean news articles. As a result of forecasting regression analysis, the index is found to be useful for predicting market conditions. Bai et al. (2021) collects keywords from shipping-related magazines and constructs a shipping sentiment index for dry bulk shipping through text analysis. They find that it is helpful in predicting the dry bulk freight rate through vector autoregression (VAR) model and least absolute shrinkage and selection operation (Lasso) regression.

The advantages of an index that uses the news and big data are expandability and economic feasibility. Big data analysis allows changes in the proposed keywords, when necessary, to develop a new market condition index. Therefore, the resulting index is highly flexible and requires relatively little economic cost and time.

Moreover, big data analysis based on the news can build long-term time series regardless of time and space. The survey data of market participants, such as those used to build the SBSI, cannot generate information if the survey is not conducted and no data are collected. However, the proposed Shipping News Business Index is based on the news, and big data analysis can generate data for any point in which news archives exist, thus making it possible to construct long time series.

Therefore, the most commonly used indicator for assessing risk in the shipping industry the Shipping Business Index. However, to determine the market conditions of the Korean shipping industry, appropriate indicators are needed to measure the judgments of market agents on the credit risk of the shipping industry.

As mentioned above, this study measures the credit risk of Korean shipping companies and the uncertainty of the shipping industry as a whole using the news and big data. Like Park et al. (2018), the EPU index method by Baker et al. (2016) is adopted for building the SRSI, and keywords are selected using the frequency of news articles that include 26 keywords in five categories based on the 'six C's' of credit proposed by Grammenos (2013). To test the utility of the index, cross-correlation analysis and forecasting regression analysis are performed. Unlike Park et al. (2018), this study consists of the SRSI, Baltic Dry Index (BDI), and World Trade Volume Index (WTVI) as predictors of the seaborne transport freight revenue, and analyse including the COVID-19 period. In addition, the dynamic relationship between the SRSI and seaborne transport freight revenue is analysed by applying the local projections analysis of Jordà (2005).

Our approach suggests that an index using news big data analysis can be a useful proxy for something that hard to observe, which could be especially valuable in industries with fewer data sources as in shipping industry.

3 Constructing the SRSI

News about the credit risk of the shipping industry was extracted from BIGKinds (https://www.bigkinds.or.kr).[3] BIGKinds has the advantage of being able to accurately extract related articles by setting several common keywords at once or by setting keywords sequentially among articles from all newspapers in Korea. The index is created using news articles from January 2000 to June 2021 obtained from 19 daily and financial newspapers represented by BIGKinds. Table 1 shows the media companies used in the study, the period of analysis, circulation in 2019, and the ranking of circulations. The circulation of the media companies used covers 64.03% of total circulation in 2019.

The keywords used to measure the credit risk of the shipping industry are chosen by using the 'six C's of credit' proposed by Grammenos (2013) as the theoretical basis. The business activities of shipping companies can be broadly classified into two types: profit generation by operating the ships, and profit generation by S&P ships. Thus, the credit risk of shipping companies is mainly caused by the fluctuations of the ship value, which is the main collateral of the shipping companies. The six C's of credit compromises character/capacity, capital, company, conditions, and collateral.

Character/capacity evaluates the long-term plans and strategies of the company. Investment, finance, chartering, risk management, and human resources are emphasised as factors to consider in evaluating character/capacity. This factor values the achievement of management efficiency based on low costs and high profits compared to competitors. Capital reflects the capital-intensive character of the shipping industry. The credit risk of a company is evaluated by analysing the adequacy of the shares by shipowners and managers, financial structure, and financial statements. A company is evaluated based on its input and output structure. Inputs relate to the employment type, contracts with charterers, and choices of charterers, while outputs relate to operating expenses such as insurance and management. Market share is also a major target of consideration. Conditions deal with the overall world economy, shipping market, as well as financial and political situations. The interest rate in the financial market and the government's policy direction regarding trade are also considered. The world economy is evaluated in terms of global industrial production and maritime trade volume, and the evaluation items of the shipping market include supply-demand analysis and prospects as well as

regulations. Finally, collateral relates to the fleet structure and condition as the major assets of shipping companies as well as the lifespan of ships, and it also includes securities aside from ships.

Table 1 Media companies used for building the SRSI

Media company	*Circulation*	*Ranking*	*Period included*
Chosun Ilbo	1,212,208	1	2018m1–2021m6
Dong-A Ilbo	925,919	2	2000m1–2021m6
JoongAng Ilbo	861,984	3	2000m1–2021m6
Maeil Business Newspaper	706,760	4	2000m1–2021m6
Korea Economic Daily	526,908	6	2000m1–2021m6
Hankyoreh	205,748	7	2000m1–2021m6
Munhwa Ilbo	195,068	8	2000m1–2021m6
Hankook Ilbo	219,672	9	2000m1–2021m6
Kyunghyang Shinmun	190,677	10	2000m1–2021m6
Seoul Shinmun	160,348	11	2000m1–2021m6
Money Today	82,322	18	2001m1–2021m6
Seoul Economic Daily	102,442	20	2000m1–2021m6
Kukmin Ilbo	74,685	21	2000m1–2021m6
Naeil Shinmoon	56,279	24	2003m10–2021m6
Herald Business	54,934	28	2000m1–2021m6
Aisa Business Daily	35,030	38	2008m1–2021m6
Financial News	37,439	39	2000m1–2021m6
Aju Business Daily	23,022	53	2007m11–2021m6
Segye Times	2,030	161	2000m1–2021m6

Source: Korea Audit Bureau of Certification

The major difficulty in creating variables related to the news is discriminating between good and bad news (Kang et al., 2015). A representative study that measures uncertainty based on news articles, Baker et al. (2016) also conducts audit study based on human readings for 6 months to prepare EPU indices. Since the Korean language presents many adversative relationships in the sentence structure, determining whether the news is good or bad by a keyword search is particularly hard. In addition, there are problems in that synonyms of keywords, distinction between verbs and nouns, and shipping can be used as company names or similar words such as dry bulk and tanker.

Being aware of this issue, we survey the keywords that a professional journalist likely uses when the events related to six C's of credit occur. Then the keywords used to build the SRSI are proofread by a professional journalist and are not likely to be affected by the adversative structure of sentences. We also conduct audit study of 100 randomly selected articles to determine whether the contents indicated an increase or decrease in the credit risk of the shipping industry.

This study selects 26 keywords that can be used when the credit risk of each field increases based on the five fields of the six C's in credit rating. Table 2 shows the contents of the six C's and the keywords used for building the SRSI. When selecting

news articles through the selected keywords, this study used 'shipping' as a common keyword, which should be considered along with all the selected keywords to only extract articles related to the shipping industry. We collect articles including common keyword and keywords for each category, and generate the number of articles for each category. For instance, when articles related to capital category are collected, we use articles that contain the following terms: 'shipping' and 'creditor' or 'court receivership' or 'bankruptcy' or 'default' or 'sell-off' or 'uncertainty'.

Table 2 Keywords used for building the SRSI based on the six C's of credit

Category	*Main content*	*Keyword*	*No. of articles*
Common keyword	Selection of news articles related to the shipping industry	Shipping	79,119
Character/capacity	Goals and long-term strategies of companies	Risk, slowdown, minus, deterioration, failure, all-time low, depression, deficit, slump, loss	30,897
Capital	Financial structure	Creditor, court receivership, bankruptcy, default, sell-off, uncertainty	21,161
Company	Input and output structure	Dismissal/discharge, cutback, restructuring	11,927
Conditions	Financial market, world economy, shipping market	Dark clouds, gloom, decline, bad news, oversupply	14,133
Collateral	Fleets and securities	Scrapping/demolition, drought	1,001

Source: Grammenos (2013), Kim and Park (2017)

A total of 79,119 articles are collected, and the number of articles for each category is 30,897 for character/capacity, 21,161 for capital, 11,927 for company, 14,133 for conditions, and 1,001 for collateral.

Baker et al. (2016) proposes the EPU index to measure policy uncertainty in macroeconomics. The EPU index is based on the frequency of news articles that report policy uncertainty. The SRSI is also created based on the method introduced by Baker et al. (2016). The method can be described as follows.

First, the number of monthly economics articles in each newspaper is collected by focusing on the keywords based on the six C's of credit introduced in the previous section. To generate the credit risk to the total number of articles, $Y_{i,t}$ the number of articles in newspaper i at time t reflecting credit risk, $X_{i,t}$ is divided by the total number of economics-related articles in newspaper i at time t, $\text{Total}_{i,t}$.[4]

$$Y_{i,t} = \frac{X_{i,t}}{\text{Total}_{i,t}}, \tag{1}$$

To standardise the $Y_{i,t}$ we employ variance σ_i^2 of $Y_{i,t}$ from each newspaper and generate normalised following equations (2) information;

$$Z_{i,t} = \frac{Y_{i,t}}{\sigma_i}, \tag{2}$$

From the normalised information, we derive following equation (3), monthly averaged μ_t;

$$\mu_t = \frac{1}{n}\sum_{i=1}^{n} Z_{i,t} \tag{3}$$

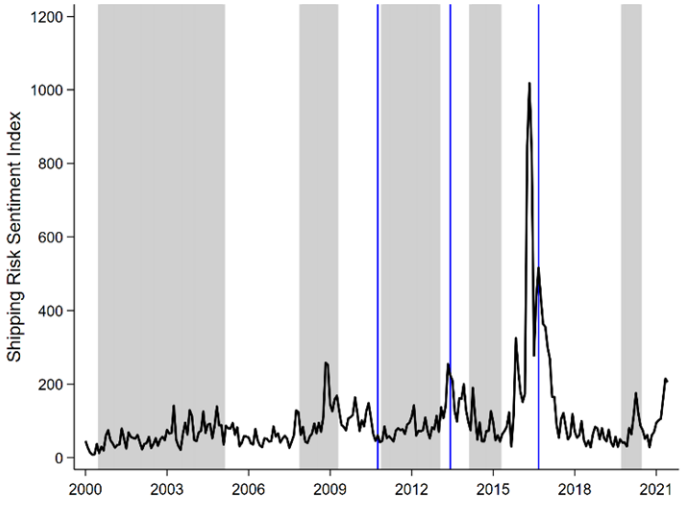
To generate the SRSI (hereafter $SRSI_t$) total mean is converted to 100

$$SRSI_t = \mu_t \times \frac{100}{M}, \tag{4}$$

where M is the total mean for the sample period of μ_t. Thus, when SRSI, $SRSI_t$ is greater than 100, it is implied the credit risk has increased. On the contrary if it is lower than 100, then it represents stable credit risk situation.

Figure 1 shows the SRSI from January 2000 to June 2021. The SRSI reflects the frequency of articles in major Korean media companies. The index is constructed using 79,119 news articles. The black line in Figure 1 represents the SRSI and the grey shaded areas indicate recession periods according to OECD business cycle chronology. The blue vertical lines indicate when major shipping companies in Korea filed for court receivership.[5]

Figure 1 Shipping Industry Risk Sentiment Index (see online version for colours)



Notes: The SRSI is built using the number of articles on the credit risk of shipping companies among the economics articles of 19 media companies, including daily and financial newspapers in Korea, from January 2000 to June 2021. The grey shaded regions show recessions as dated by OECD. The blue vertical lines indicate when major shipping companies in Korea filed for court receivership.

The SRSI was stable under 100, but it showed an increasing pattern during the business contraction period, indicating that the index behaves countercyclically. Aside from this period, the SRSI also significantly increased before the major shipping companies in Korea filed for court receivership. The index lagged during the global financial crisis of 2008 as, in these periods, the crisis in the shipping industry was caused by macroeconomic factors. However, the SRSI was leading when major shipping companies

filed for court receivership. In particular, the SRSI skyrocketed in 2017 when Hanjin Shipping, the number 1 shipping company in Korea, went bankrupt. Therefore, the SRSI can provide useful information for diagnosing and predicting the credit risk of the Korean shipping industry.

4 Evaluating the SRSI

The impact of the shipping sentiment index constructed by Bai et al. (2021) has a positive effect on dry bulk freight rates such as BDI, BCI and BPI. In other words, since the risk sentiment index for the shipping industry constructed in this study has an inverse relationship with the shipping economy, if risk sentiment increases, Korea's seaborne transport freight revenue and SBSI will deteriorate.

In this section, we examine the explanatory power of the SRSI. To this end, correlation test with the SBSI, regression analysis with Korea's seaborne transport freight revenue, and impulse response analysis with Korea's seaborne transport freight revenue are performed.

4.1 Correlation test with the business survey index

As uncertainty about credit risk increases in the shipping industry, companies reduce their investments; business conditions, profitability, and financial conditions deteriorate; and sales decrease. Thus, this study analyses the cross-correlations between each item in the SBSI, which describes the industry, profitability, and financial conditions, and the SRSI.

Korea Maritime Institute has conducted a business survey of the shipping companies in Korea since 2013:M8. The participants in the survey are corporate companies that are members of the Korea Shipowners Association. The number of corporations is 144 and the average number of respondents is approximately 80. The SBSI is compiled to grasp the business conditions for the current month and outlooks for the next month by surveying companies' sentiment. A total of four items, including business conditions, profitability, financial conditions, sales are surveyed.

A cross-correlation analysis measures correlations by considering the time lags of two variables. Cross-correlations can be presented for various time lags through a cross-correlation diagram and analysis, and the lead-lag relationship between two variables can also be explained.

$$corr\left(X_t, Y_{t-s}\right) = \frac{\operatorname{cov}\left(X_t, Y_{t-s}\right)}{\sqrt{\operatorname{var}\left(X_t\right)\operatorname{var}\left(Y_{t-s}\right)}}, \tag{5}$$

In equation (5), if s is greater than 0, the equation gives the correlation coefficient of the historical value of variable Y and the current value of variable X. If s is 0, the equation gives the simultaneous correlation coefficient of the two variables. If s is less than 0, if s is greater than 0 ($0 < s < t$), the equation gives the correlation coefficient of the current value of variable X and the future value of variable Y.

The SRSI is considered to have a high correlation with the economic sentiment and outlooks of the Korean shipping industry. Hence, this study tested its cross-correlation with the SBSI, which represents the sentiment on shipping business conditions.

Table 3 shows the results of the cross-correlation test between the two indexes (by the time lag). A negative correlation exists between the two indexes regardless of the time lag. The SRSI has a statistically significant correlation with the SBSI three months before. In particular, a statistically significant correlation exists with business conditions and sales in both sentiment for current month and outlooks for following month.

Table 3 Cross-correlation analysis between the SRSI and the SBSI

Time lag (h)	*Sentiment for current month*			
	Business conditions	*Profitability*	*Financial conditions*	*Sales*
–3	–0.23**	–0.04	–0.11	–0.35***
–2	–0.26**	–0.09	–0.11	–0.37***
–1	–0.24**	–0.07	–0.06	–0.35***
0	–0.2*	–0.03	–0.05	–0.32***
Time lag (h)	*Outlooks for following month*			
	Business conditions	*Profitability*	*Financial conditions*	*Sales*
–3	–0.35***	–0.18*	–0.16	–0.36***
–2	–0.37***	–0.22**	–0.21**	–0.37***
–1	–0.32***	–0.2*	–0.16	–0.33***
0	–0.26**	–0.17*	–0.13	–0.3***

Notes: This table reports the cross-correlation coefficients between the SRSI at point t and the SBSI at point $t + h$. The estimation period is from 2013:M8 to 2021:M6.
*p < .10, **p < .05, ***p < .01.

4.2 Regression analysis

In this subsection, we apply an ARDL model to examine whether SRSI has predictive power on Korean seaborne transport freight revenue which is a proxy for real economic activity of Korean shipping industry. The seaborne transport freight revenue is attained from the Bank of Korea's extended balance of payment services classification (EBOPS). Specifically, letting Y_t denote monthly measure of seaborne transport freight revenue, we estimates the following an ARDL model:

$$\Delta Y_{t+h} = \alpha + \beta_1 SRSI_t + \beta_2 \Delta bdi_t + \beta_3 \Delta wtvi_t + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-1} + \varepsilon_{t+h}, \quad (6)$$

where ΔY_{t+h} is cumulative percentage change of seaborne transport freight revenue from $t - 1$ to h and $h \geq 0$ is the forecast horizon;[6] $SRSI_t$, Δbdi_t, and $\Delta wtvi_t$ represent the SRSI, log-difference of the BDI and the WTVI of Centraal Planbureau (CPB), respectively. WTVI reflects the evolution in the global merchandise trade, which covers the international trade of 81 countries, accounting for 99% of trade worldwide (Ebregt, 2020). We transform the BDI and the WTVI to first difference of log as the results for augmented Dicky-Fuller (ADF) test reveals that the underlying time series are non-stationary. Our specification investigates the marginal information content of the SRSI conditional on the BDI and the WTVI, two key indicators which are known to have predictive power regarding the shipping market (Stopford, 2008; Park et al., 2021).

The optimal number of lags is four suggested through Akaike information criterion (AIC). The data are monthly and range from January 2006 to June 2021. We obtain the data from diverse sources. Seaborne transport freight revenue (Y) is obtained from the Bank of Korea; the BDI is collected from the Clarksons (https://sin.clarksons.net) dataset and the WTVI is taken from CPB Netherlands Bureau for Economic Policy Analysis (https://www.cpb.nl/en/worldtrademonitor). Our sample period lies between 2006:M1 and 2021:M6 due to data availability.

Table 4 ADF test

Variable	*ADF*		
	Level	*Log*	*Log-difference*
SRSI	–5.18***	–5.83***	–20.62***
BDI	–2.21	–2.37	–12.32***
WTVI	–0.94	–1.51	–13.68***

Notes: *, **, *** indicate that null hypothesis is rejected at statistical significance level of 10%, 5%, and 1%, respectively.

Table 5 SRSI forecasting regression analysis results

	1-month forecast horizon			*4-month forecast horizon*		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
$SRSI_t$	–0.01*	–0.01*	–0.01**	–0.01*	–0.01*	–0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Δbdi_t		0.09***	0.07**		0.09**	0.08*
		(0.03)	(0.03)		(0.04)	(0.04)
$\Delta wtvi_t$			1.54***			0.61
			(0.45)			(0.61)
ΔY_{t-1}	0.44***	0.41***	0.36***	0.65***	0.63***	0.62***
	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)
ΔY_{t-2}	–0.35***	–0.33***	–0.33***	0.22**	0.20**	0.20**
	(0.07)	(0.07)	(0.07)	(0.09)	(0.09)	(0.09)
ΔY_{t-3}	0.48***	0.48***	0.43***	0.07	0.08	0.07
	(0.07)	(0.07)	(0.07)	(0.09)	(0.09)	(0.09)
ΔY_{t-4}	–0.15**	–0.15**	–0.14**	–0.17**	–0.17**	–0.17**
	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)
RMSE	10.20	9.966	9.665	12.73	12.60	12.60
R^2	0.30	0.34	0.38	0.65	0.65	0.66
Observations	179	179	179	175	175	175

Notes: This table reports the estimates of an ARDL regression. The dependent variables in columns 1–3 are the growth rates of seaborne transport freight revenue for the 1-month forecast horizon, and those in columns 4–6 are the growth rates of seaborne freight revenue for the 4-month forecast horizon. Constant terms are omitted for convenience. The analysis period is from 2006:M1 to 2021:M6. *p < .10, **p < .05, ***p < .01.

Table 5 reports estimates of an ARDL regression. The results for the forecast horizon of one month are presented in columns 1 through 3 while those for the four-month horizon are shown in columns 4 through 6.[7]

SRSI is statistically significant at the 1 and 4-month horizons, which proves that the index has significant explanatory power as a predictor variable of freight income.

In the one-month forecast horizon model, the coefficients on *SRSI* in column 1, which only includes *SRSI*, have lower statistical significance and R^2 than the coefficients reported in columns 2 and 3. This result indicates that the explanatory power of SRSI is lower than that of the model that controls for WTVI and BDI in the short run. World trade volume and freight rates are primarily determined by global macroeconomic conditions, outside of the Korean shipping industry, and the shipping industry perceives the changes in macroeconomic conditions with a time lag. The fact that the coefficients on *SRSI* in the model that controls for world trade volume and BDI are statistically significant indicates that *SRSI* has sufficient explanatory power for phenomena that cannot be explained by world trade volume and freight rates.

In the four-month forecast horizon model, *SRSI* shows statistically significant results regardless of the controls for the BDI and the WTVI. However, the coefficients on the WTVI is not statistically significant in the four-month forecast horizon model.

4.3 Impulse response analysis

We use Jordà (2005) local projections to investigate the full dynamic relation between seaborne transport freight revenue and SRSI. Local projections are widely used in time series analysis recently, as they have been found to be more robust to misspecification, and easily allow for inclusion of control variables (Ramey, 2016; Mian et al., 2017). The local projections impulse responses to risk sentiment shocks are derived by estimating a set of regressions, for $h = 0, \ldots, 24$:

$$y_{t+h-1} = \alpha^h + \gamma^h \Delta bdi_t + \theta^h \Delta wtvi_t + \sum_{j=1}^{4} \beta^h SRSI_{t-j} + \sum_{j=1}^{4} \delta^h y_{t-j} + \epsilon^h_{t+h-1}, \tag{7}$$

In equation (7), we can compute the impulse response functions by estimating the sequence of coefficients $\{\hat{\beta}^h\}$. By doing so, we can trace the effects of risk sentiment shocks to seaborne transport freight revenue dynamically.

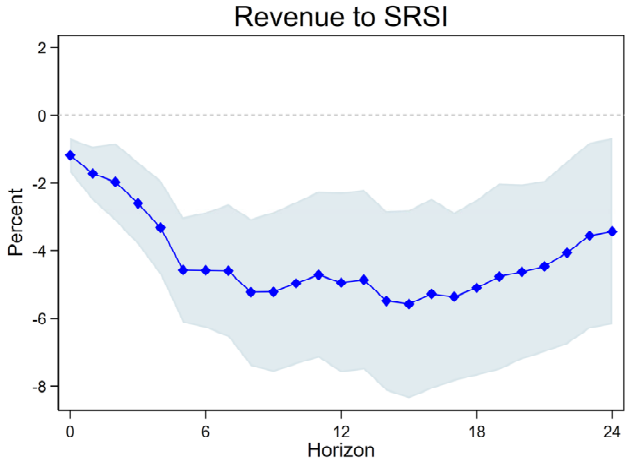
Figure 2 shows the impulse response of seaborne transport freight revenue to risk sentiment shocks, along with 90% confidence intervals based on Newey-West standard errors (Newey and West, 1987). As shown in Figure 2, there is a significant decrease in seaborne transport freight revenue after risk sentiment shocks. Following a surprise increase in *SRSI*, freight revenue declines by a little less than 2% on impact, continues to decline for 18 months. The negative revenue response shows prolonged and quite persistent pattern. Figure 2 shows maximum estimated drops of 5.57% in freight revenue, which is about 1.60% of the services credit of Korea in 2021.

4.4 Discussion

In this section, we apply various econometric methods to examine the performance of the SRSI. The results from the correlation test with the SBSI suggest that the SRSI is

negatively correlated with the SBSI. In particular, the SRSI tends to move in the opposite direction with overall market condition and sales of shipping companies, indicating that the news-based analysis can reflect the sentiment in the shipping market.

Figure 2 Impulse response to risk sentiment shocks (see online version for colours)



Notes: This figure displays the impulse response function of seaborne transport freight revenue to risk sentiment shocks. The horizontal axes measure the analysed horizon and the vertical axes measure the percentage change. The line with squares is the OLS point estimates, and the shaded regions are 90% confidence band based on Newey-West corrections for standard errors.

To investigate the predictive ability of the risk sentiment on the shipping market, we estimate the ARDL model. The results from the regression shows that the SRSI has significant predictive content at both 1 and 4-month forecast horizons. The dynamic relation between the shipping market and the risk sentiment is further examined using impulse response analyses using the local projections. The results from the impulse response analyses suggest that that an increase in risk sentiment predicts lower revenue growth, and the magnitude of sentiment shock are sizable, indicating that the SRSI could be a significant predictor for the Korean Shipping market. These findings are in line with Baker et al. (2016) and Bai et al. (2021) who find the index based on news-based analysis provides useful information about future economic conditions.

5 Concluding remarks

The SRSI measures the risk sentiment of the Korean shipping industry using the text analysis, thus providing useful information to detect and prevent risks in the Korean shipping industry. The proposed index is countercyclical, as it increases not only in business contraction periods in Korea but also during the crisis of major shipping companies.

The SRSI has a significant negative correlation with the SBSI and a statistically significant correlation with business situations and sales.

The results of the forecasting regression and impulse response analysis of maritime transport freight income show that the proposed index has a statistically significant explanatory power as a predictor variable for freight income, which proves that the index is useful for detecting the credit risk of the shipping industry in advance. These results are in line with the research results that the impact of the shipping sentiment index constructed by Bai et al. (2021) has a positive effect on dry bulk freight rates.

The contribution of this study lies in its new approach, namely news big data analysis, to measure the credit risk of the shipping industry. The proposed index is expected to help the government manage and supervise risk in the shipping market and carry out efficient decision-making.

Future studies should identify additional keywords that comprehensively connote the risk of the shipping industry in addition to the keywords proposed by this study and improve the index development method. In addition, although this study measures risk sentiment for the overall shipping industry in Korea, future research needs to analyse it by dividing it into segmented markets such as containers, dry bulk, and tankers.

Finally, to determine the degree of credit risk in Korea's shipping industry, this study estimates the SRSI only through news articles in Korea. However, almost no boundary exists between countries in the shipping business. Therefore, although the risk sentiment index of the shipping industry in this study is constructed through Korean newspaper articles, it can be used to understand the status of the global shipping market.

Nevertheless, future research should analyse the overall global shipping market along with shipping industry risk by creating a Global Shipping Market News Index based on global news article keywords obtained from global shipping journals and Google Trends. This issue should be further discussed by policy makers, authorities, shipping industry participants, and academia.

Acknowledgements

This research is a revised and expanded version of 'A Study on Default Prediction of Shipping Companies based on Financial Information' supported by the research grant RP2016-01 of Korea Maritime Institute and part of the first author's PhD thesis (Park, 2019). This paper is the sole responsibility of its authors. The views represented here do not necessarily reflect those of the Korea Maritime Institute. We are very grateful to the anonymous referees for their constructive comments, which improved both the content as well as the presentation of the paper.

The open access fee of this work is supported by China Merchants Energy Shipping.

References

Aastveit, K.A., Natvik, G.J.J. and Sola, S. (2013) *Economic Uncertainty and the Effectiveness of Monetary Policy*, Norges Bank Working Paper 17 [online] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2353008.

Aggarwal, C.C. and Zhai, C. (2012) 'An introduction to text mining', in *Mining Text Data*, pp.1–10, Springer, Boston, MA, USA.

Bai, X., Lam, J.S.L. and Jakher, A. (2021) 'Shipping sentiment and the dry bulk shipping freight market: new evidence from newspaper coverage', *Transportation Research Part E: Logistics and Transportation Review*, Vol. 155 [online] https://doi.org/10.1016/j.tre.2021.102490.

Baker, S.R., Bloom, N. and Davis, S.J. (2016) 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics*, Vol. 131, No. 4, pp.1593–1636 [online] https://doi.org/10.1093/qje/qjw024.

Balahur, A., Steinberger, R., Kabadjov, M., Zavarela, V., Van Der Goot, E., Halkia, M., Pouliquen, B. and Belyaeva, J. (2010) 'Sentiment analysis in the news', *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC'2010)*, Valletta, Malta, 19–21 May, pp.2216–2220.

Bank of Korea [online] http://ecos.bok.or.kr/ (accessed 30 June 2021).

BIGKinds [online] https://www.bigkinds.or.kr (accessed 30 June 2021).

Bloom, N. (2009) 'The impact of uncertainty shocks', *Econometrica*, Vol. 77, No. 3, pp.623–685 [online] https://doi.org/10.3982/ECTA6248.

Bontempi, M.E., Golinelli, R. and Squadrani, M. (2016) *A New Index of Uncertainty based on Internet Searches: A Friend or Foe of Other Indicators?*, Quaderni – Working Paper DSE No. 1062 [online] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2746346.

Bram, J. and Ludvigson, S. (1997) *Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race*, No. 9708, Federal Reserve Bank of New York.

Chen, J., Lu, L., Lu, J.J., Feng, Y. and Wu, Y. (2014) 'Framework development of shipping market early-warning system by using prosperity index: a case study of Shanghai dry bulk shipping market', *Transportation Research Board 93rd Annual Meeting, 2014*, Washington DC, 12–16 January 2014, Report/Paper No. 14-4860, 17pp.

Choi, J.S., Kim, K.. and Han, H.J. (2018) 'Development of early warning index in Korean shipping industry by using signal approach', *Maritime Policy and Management*, Vol. 45, No. 8, pp.1007–1020 [online] https://doi.org/10.1080/03088839.2018.1467058.

Clarksons [online] https://sin.clarksons.net (accessed on 30 June 2021).

CPB Netherlands Bureau for Economic Policy Analysis [online] https://www.cpb.nl/en/worldtrademonitor (accessed 30 June 2021).

Ebregt, J. (2020) *The CPB World Trade Monitor: Technical Description*, CPB Background Document, CPB Netherlands Bureau for Economic Policy Analysis [online] https://www.cpb.nl/sites/default/files/omnidownload/CPB-Background-Document-April2020-The-CPB-World-Trade-Monitor-technical-description-update_27.pdf.

Fudenberg, D. and Maskin, E. (2009) 'The folk theorem in repeated games with discounting or with incomplete information', in *A Long-Run Collaboration on Long-Run Games*, pp.209–230 [online] https://doi.org/10.1142/9789812818478_0011.

Gelper, S., Lemmens, A. and Croux, C. (2007) 'Consumer sentiment and consumer spending: decomposing the Granger causal relationship in the time domain', *Applied Economics*, Vol. 39, No. 1, pp.1–11 [online] https://doi.org/10.1080/00036840500427791.

Grammenos, C.T. (2013) 'Revisiting credit risk, analysis and policy in bank shipping finance', in *The Handbook of Maritime Economics and Business*, pp.807–840, Informa Law from Routledge, London, UK.

Heston, S.L. and Sinha, N.R. (2016) 'News versus sentiment: predicting stock returns from news stories finance and economics discussion series', 2016-048 [online] http://dx.doi.org/10.17016/FEDS.2016.048.

Jordà, Ò. (2005) 'Estimation and inference of impulse responses by local projections', *American Economic Review*, Vol. 95, No. 1, pp.161–182, DOI: 10.1257/0002828053828518.

Kang, S.A., Kim, Y.S. and Choi, S.H. (2015) 'Study on the social issue sentiment classification using text mining', *Journal of the Korean Data & Information Science Society*, Vol. 26, No. 5, pp.1167–1173, DIO: 10.7465/jkdi.2015.26.5.1167.

Kim, T.I. and Park, S.H. (2017) *A Study of Default Prediction for Shipping Companies using Financial Information*, Korea Maritime Institute, Busan, Korea.

Lee, J.K. and Kim, I.S. (2016) 'An analysis on crisis indices of shipping markets in Korea based on the signals approach', *Journal of International Trade & Commerce*, Vol. 12, pp.485–496, DOI: 10.16980/jitc.12.5.201610.485.

Li, J., Xu, Z. and Tang, L. (2016) 'Forecasting oil price trends with sentiment of online news articles', *Procedia Computer Science*, Vol. 91, pp.1081–1087 [online] https://doi.org/10.1016/j.procs.2016.07.157.

Mian, A., Sufi, A. and Verner, E. (2017) 'Household debt and business cycles worldwide', *The Quarterly Journal of Economics*, Vol. 132, No. 4, pp.1755–1817 [online] https://doi.org/10.1093/qje/qjx017.

Ministry of Oceans and Fisheries (2015) *Early Warning System Building Project for the Shipping Market*.

Newey, W.K. and West, K.D. (1987) 'Hypothesis testing with efficient method of moments estimation', *International Economic Review*, pp.777–787 [online] https://doi.org/10.2307/2526578.

Papapostolou, N.C., Nomikos, N.K., Pouliasis, P.K. and Kyriakou, I. (2014) 'Investor sentiment for real assets: the case of dry bulk shipping market', *Review of Finance*, Vol. 18, No. 4, pp.1507–1539 [online] https://doi.org/10.1093/rof/rft037.

Papapostolou, N.C., Pouliasis, P.K., Nomikos, N.K. and Kyriakou, I. (2016) 'Shipping investor sentiment and international stock return predictability', *Transportation Research Part E: Logistics and Transportation Review*, Vol. 96, pp.81–94 [online] https://doi.org/10.1016/j.tre.2016.10.006.

Park, S. (2019) *A Study on Business Fluctuation and Risk Assessment in the Shipping Market*, unpublished PhD thesis, Pusan National University, Busan, Korea.

Park, S., Kwon, J. and Kim, T. (2018) 'Developing the maritime news index using news big data analysis', *Ocean Policy Research*, Vol. 33, No. 1, pp.281–301.

Park, S., Kwon, J. and Kim, T. (2021) 'An analysis of the dynamic relationship between the global macroeconomy and shipping and shipbuilding industries', *Sustainability*, Vol. 13, No. 24 [online] https://doi.org/10.3390/su132413982.

Ramey, V.A. (2016) 'Macroeconomic shocks and their propagation', *Handbook of Macroeconomics*, Vol. 2, pp.71–162 [online] https://doi.org/10.1016/bs.hesmac.2016.03.003.

Rich, R. and Tracy, J. (2010) 'The relationships among expected inflation, disagreement, and uncertainty: evidence from matched point and density forecasts', *The Review of Economics and Statistics*, Vol. 92, No. 1, pp.200–207 [online] https://doi.org/10.1162/rest.2009.11167.

Rossi, B. and Sekhposyan, T. (2015) 'Macroeconomic uncertainty indices based on nowcast and forecast error distributions', *American Economic Review*, Vol. 105, No. 5, pp.650–55, DOI: 10.1257/aer.p20151124.

Scotti, C. (2016) 'Surprise and uncertainty indexes: real-time aggregation of real-activity macro-surprises', *Journal of Monetary Economics*, Vol. 82, pp.1–19 [online] https://doi.org/10.1016/j.jmoneco.2016.06.002.

Stopford, M. (2008) *Maritime Economics*, 3rd ed., Routledge, London.

Tobback, E., Naudts, H., Daelemans, W., de Fortuny, E.J. and Martens, D. (2018) 'Belgian economic policy uncertainty index: improvement through text mining', *International Journal of Forecasting*, Vol. 34, No. 2, pp.355–365 [online] https://doi.org/10.1016/j.ijforecast.2016.08.006.

Notes

1 Korea Maritime Institute press release, November 30, 2016.

2 The Shipping Industry Business Survey Index has been issued since August 2013.

3 Big Kinds is the news archive of the Korea Press Foundation.

4 The number of articles varies depending on the size of the newspaper company, so the units are standardised by dividing by the total number of economics articles.

5 Korea Line Corporation, under court receivership (October 2010); Pan Ocean, under court receivership (June 2013), and Hanjin Shipping, under court receivership (September 2016).

6 $\Delta Y_{t+h} = \ln\left(\frac{Y_{t+h}}{Y_{t-1}}\right) \times 100.$

7 We set the forecast horizon 1 and 4-month as the number of optimal lags suggested by AIC.