Shipping Investor Sentiment and International Stock Return Predictability

Nikos C. Papapostolou* , Panos K. Pouliasis, Nikos K. Nomikos, and Ioannis Kyriakou

August 06, 2015

Abstract. Stock return predictability by investor sentiment has been subject to constant updating, but reaching a decisive conclusion seems rather challenging as academic research relies heavily on US data. We provide fresh evidence on stock return predictability in an international setting and show that shipping investor sentiment is a common leading indicator for financial markets. We establish out-of-sample predictability and demonstrate that investor sentiment is also economically significant in providing utility gains to a mean-variance investor. Finally, we find evidence that the predictive power of sentiment works best when negative forecasts are also taken into account.

JEL classification: G12, G02, C53

Keywords: Shipping Investor sentiment, stock return predictability, out-of-sample forecast performance

---

* Papapostolou, n.papapostolou@city.ac.uk, +44 (0) 20 7040 8620, Cass Business School, Faculty of Finance, 106 Bunhill Row, EC1Y 8TZ, London, United Kingdom; Pouliasis, p_pouliasis@city.ac.uk, +44 (0) 20 7040 8807, Cass Business School, Faculty of Finance, 106 Bunhill Row, EC1Y 8TZ, London, United Kingdom; Nomikos (corresponding author), n.nomikos@city.ac.uk, +44 (0) 20 7040 0104, Cass Business School, Faculty of Finance, 106 Bunhill Row, EC1Y 8TZ, London, United Kingdom; Kyriakou, ioannis.kyriakou@city.ac.uk, +44 (0) 20 7040 8738, Cass Business School, Faculty of Actuarial Science and Insurance, 106 Bunhill Row, EC1Y 8TZ, London, United Kingdom. We thank Steven Kou, Christian Ewald, and participants at the 2015 symposium on Quantitative Finance and Risk Analysis for their comments and suggestions. The authors also thank Kenneth French, Amit Goyal, Jeffrey Wurgler and Guofu Zhou for providing data on their websites.

1. Introduction

Investor sentiment as a predictor of the cross-section of stock returns has been identified in the literature by several studies. Brown and Cliff (2004) find that sentiment levels and changes are correlated with market returns, although the predictive power for stock returns is small. Lemmon and Portniaguina (2006) study the relationship between investor sentiment and small-stock premium and find that sentiment forecasts the returns of small and low institutional ownership stocks. Baker and Wurgler (2006, 2007) find sentiment to have larger effects on stocks whose valuation is highly subjective and difficult to arbitrage. Stambaugh, Yu, and Yuan (2012, 2014) explore the role of sentiment in a broad set of anomalies in cross-sectional returns and find that its predictive power is higher during high-sentiment periods. Huang *et al.* (2015) suggest that investor sentiment is important not only for cross-sectional returns, but also at the aggregate market level. However, evidence on the effects of sentiment exclusively focuses on cross-section results for the US stock market with Baker, Wurgler, and Yuan (2012) being the only study to explore investor sentiment, that is based on market proxies, in an international setting. Their study shows that annual investor sentiment is a contrarian predictor of cross-sectional international stock market returns in an in-sample framework. As such, there is no guarantee that the results are protected against in-sample overfitting and can be used to produce accurate forecasts of stock returns. Therefore, whether investor sentiment can predict international stock market returns remains an open question.

Our paper fills an important gap in the finance literature by providing a comprehensive picture of stock return predictability using measures of international investor sentiment that are based on market actions. Instead of focusing solely on US stock market data, we study stock return predictability in an international setting which allows broadening the evidence on the specific research question. The use of international stock return data and a common investor sentiment index provides also a natural out-of-sample test for earlier US based findings. Further, we complement earlier studies by Baker and Wurgler (2006, 2007), Baker, Wurgler, and Yuan (2012) and Huang *et al.* (2015) by looking into the predictive ability of investor sentiment for stock returns from a different angle, that of shipping investor sentiment.

Given the lack of a common measure of investor sentiment for stock markets internationally, earlier studies have attempted to use accurate proxies of it and consumer confidence indices are found to be highly adequate measures. However, these indices have flaws as direct measures of investor sentiment, since their construction is based on surveys and consumers' actions can differ substantially to that of their responses. We overcome this limitation using investor sentiment indices for the three major shipping markets (container, drybulk and tanker[2]), that are based on actual market sentiment proxies and the principles set out in Papapostolou *et al.* (2014), to study stock return predictability by investor sentiment in an international setting.

Why use shipping sentiment to predict international stock market returns? The importance of maritime transportation to the world economy has been highlighted as early as the 18th century by Adam Smith (1776) who illustrates the economic benefits offered by sea transportation. Today, seaborne trade is the backbone of the global economy but is hardly present in the finance literature. Approximately 80% of global trade by volume and over 70% of global trade by value are transported by sea and these shares are even higher in the case of developing countries (United Nations Conference on Trade and Development (UNCTAD), 2014). Our interest in the shipping

[2] See Supplementary Appendix for a description of the three markets, vessels' sizes and type specifications.

industry also stems from the extensive reference to the Baltic Dry Index[3] (BDI) as a leading economic indicator (and the opposing views) in financial press and blogs[4], and more recently, of the overall drybulk shipping market in the finance literature (Kilian, 2009; Apergis and Payne, 2013; Alizadeh and Muradoglu, 2014; Kalouptsidi, 2014; Papapostolou *et al.*, 2014; Greenwood and Hanson, 2015). The reliance on BDI as a leading indicator of the world economy is further highlighted by its inclusion in the construction of a number of economic series, including the Goldman Sachs Global Leading Indicator (GLI).

Shipping is undoubtedly a truly global industry and can be considered as representative of the general health of the world economy. Yet, making reference exclusively on BDI is not sufficient for two reasons. First, shipping is not only about the drybulk market of the industry and the transportation of raw materials. There are other shipping markets and commodities that are equally important; for example, crude oil, which is transported by tanker vessels and is one of the most vital natural resources of industrialized nations; or finished goods, which are transported by container vessels and are closely related to consumer end-demand. Neither of these markets is captured by BDI. Second, BDI reflects the balance between supply and demand and is not solely demand-driven[5]. Therefore, the index can practically fall in an environment of expanding raw

[3] The Baltic Dry Index (BDI) tracks the cost of shipping raw materials, such as coal, iron ore, steel, cement and grain, around the world.

[4] For example: FT Alphaville: Why does the BDI matter? (30/01/2008); Wall Street Journal: Shipping-cost Index drops (24/08/2009); Financial Times: Don't panic, the Baltic dry is a rubbish indicator (07/07/2010); The Source, Wall Street Journal: Baltic dry index watchers can relax (16/11/2010); Financial Times: The shipping news: BDI does not mean buy, buy, buy (04/10/2013).

[5] As noted by Jeremy Penn, CEO of The Baltic Exchange (Baltic drying up as a gauge, Wall Street Journal, 03/03/2010) "There are two elements to the BDI: demand and supply. When the supply of shipping is fairly stable, demand represents a good pointer to activity in primary industry. BDI is a good indicator of drybulk rates in the market; but we have never made great claims for it to be more than that".

materials demand if the supply of vessels grows faster. For the above reasons, BDI has to be treated with caution when referring to it as a leading indicator[6].

Our results answer a number of empirical questions: Can shipping sentiment be considered as a common leading indicator for financial assets? Is shipping sentiment economically significant in providing utility gains to a mean-variance investor? Should the investment community concentrate only on the drybulk market of the shipping industry, as suggested recently in the financial press and finance literature? We confirm empirically the significance of sentiment as a contrarian predictor of stock market excess returns and show that shipping sentiment constitutes a global predictor of financial assets in an in-sample and out-of-sample framework. The results are consistent under both time-series and pooled regressions, while the latter estimations mitigate the data-mining problem that may plague US stock market data (Ang and Bekaert, 2007). The monthly R^2 statistics are also higher than those reported in the literature (Baker and Wurgler, 2006; Baker, Wurgler, and Yuan, 2012; Huang *et al.*, 2015), implying that shipping sentiment can predict monthly stock market excess returns remarkably well. Tanker sentiment appears to be the strongest predictor and this may be attributed to the impact of oil on the economies and stock markets of industrialized countries that are heavily dependent on the commodity (Jones and Kaul, 1996; Driesprong, Jacobsen, and Maat, 2008), much of which is seaborne. However, we find weak predictability of stock market returns by oil price changes which may be caused by the subprime and financial crises included in the sample. When these periods are excluded, changes in oil prices can predict stock market returns but tanker sentiment still outperforms.

[6] Especially in periods of oversupply of vessels as it has been the case recently. For example, during the period January 2010 to April 2014 the year-on-year average growth of the drybulk fleet (supply) was 11.61%, and the year-on-year average growth of the OECD industrial production (demand: assuming that industrial production is a proxy of drybulk shipping demand) was 2.68% (source: Clarkson Shipping Intelligence Network, http://www.clarksons.net/sin2010/). As a consequence, employing the BDI – which was recording new lows due to the oversupply of vessels – as a leading indicator of economic activity would generate the incorrect signal.

The out-of-sample R^2 statistics for tanker sentiment are sizeable and consistently significant at conventional levels, with an average value of 3.99%. However, we find evidence that tanker predictability power disappears when the forecast evaluation period starts within 16 months following the end of the subprime and financial crises. This loss of predictability can be attributed to the following facts: (1) this is a period during which stock markets worldwide have experienced a speedy and steady advance to higher levels; (2) tanker sentiment predictability power works best when negative forecasts are also taken into consideration. Furthermore, the certainty equivalent return (CER) gain and Sharpe ratio measures show that the forecasts based on lagged tanker sentiment are economically significant and can provide utility gains to a mean-variance investor.

To validate the unique predictive power of shipping sentiment, we carry out a set of robustness checks. First, we employ the monthly investor sentiment indices of Baker and Wurgler (2006) and Huang *et al.* (2015) to test whether shipping sentiment maintains its unique significant predictive power. Second, we compare shipping sentiment with physical market measures as predictors of excess stock returns. Third, we check if our results still hold when the recent subprime and financial crises are excluded from the full sample. The robustness checks reiterate the superior power of tanker sentiment compared to the container and drybulk sentiment and our initial inferences are not affected, supporting our claim that attention should be paid not only on the drybulk market and BDI, but also on the tanker market and the shipping industry at the aggregate level.

The remaining of the paper is structured as follows. In Section 2, we construct the shipping sentiment indices. In Section 3, we study the relationship between excess returns and shipping sentiment for a set of stock market indices and Section 4 presents the out-of-sample testing on return predictability. Section 5 investigates the economic significance of stock market forecasts based on the shipping sentiment indices. Section 6 goes over the robustness checks. Section 7 concludes the paper.

2. Shipping Sentiment indices

To some extent, a major challenge when empirically studying the importance of investor sentiment is that it is not directly observable. In their pioneering work, Baker and Wurgler (2006) construct an investor sentiment index for the US stock markets; whereas Baker, Wurgler, and Yuan (2012) extend the study to the construction of annual investor sentiment indices for six major stock markets. Das and Chen (2007) develop a methodology for extracting small investor sentiment from stock message boards and García (2013) constructs a measure of sentiment based on financial news from the New York Times. Moreover, Papapostolou *et al.* (2014) and Huang *et al.* (2015) exploit the six sentiment proxies of Baker and Wurgler (2006) to obtain a new investor sentiment index for shipping and an aligned investor sentiment index for the US stock markets respectively.

Measuring investor sentiment is subjective and there is no consensus on what the appropriate proxies should be (Schmeling, 2009). Extant literature (Barberis, Shleifer, and Vishny, 1998; Brown and Cliff, 2004; Baker and Wurgler, 2007; Barber, Odean, and Zhu, 2009; Antoniou, Doukas, and Subrahmanyam, 2013) suggests that there is a broad range of variables measuring sentiment. Our selection of sentiment proxies and the method to measure shipping sentiment follows Papapostolou *et al.* (2014). We combine five proxies that may reflect the sentiment of participants in the shipping markets, in addition to a component of non-sentiment related idiosyncratic variation. The selection process is based on the notion that individuals with positive (negative) sentiment make optimistic (pessimistic) judgments in relation to their investment decisions (Wright and Bower, 1992). Consequently, optimism or pessimism about the overall state of the shipping market may affect the decision of investors regarding the sale and purchase of second-hand or the order of newbuilding vessels.

We classify the proxies into three main categories. The first category is market expectations and includes the net contracting (NC) and money committed (MC) proxies. The second category is valuation where the price-to-earnings (PE) and second-hand-to-newbuilding vessel price (SNB) ratios are employed. The third category refers to liquidity and is captured by the turnover ratio (TURN). A detailed description of the sentiment proxies is provided in the Supplementary Appendix.

We calculate the proxies on a monthly basis for the following markets and sectors: (1) container market: panamax, sub-panamax and handymax sectors; (2) drybulk market: capesize, panamax, handymax and handysize sectors; and (3) tanker market: VLCC, suezmax and aframax sectors. Data is collected from Clarkson Shipping Intelligence Network over the period February 1996 to April 2014. [7] All proxies are de-trended using the one-sided Hodrick–Prescott filter[8] and their cyclical component is used in the analysis. As the proxies may embody a component that reflects underlying macroeconomic fundamentals, we remove the non-sentiment part by orthogonalizing the proxies to three macroeconomic variables: the G7 monthly industrial production growth and two recession-period dummies for the G7 and Major 5 Asia countries[9]. The macroeconomic variables are selected by taking into account the global nature of shipping markets, although we recognize that additional macroeconomic factors may still drive our proxies. Finally,

[7] Data availability for the container market begins in September 1996.

[8] In the analysis, we use a smoothing parameter of 14,000. Different values of the smoothing parameter were also tested without significant effect on the final estimated trend and cyclical components of the series. The PE and SNB ratios for a number of sectors were also de-trended as they appeared to be marginally stationary (unit root tests indicated stationarity only at the far end of the 10% significance level).

[9] Data provided by the Organisation for Economic Co-operation and Development (OECD).

following Baker, Wurgler, and Yuan (2012), we construct total, market and sector-specific sentiment indices using the first principal component method.

2.1. Total, Market and Sector-Specific Sentiment Indices

For each market and sector we construct a first-stage index comprising fifteen loadings given by the current, one-month lagged and two-month lagged orthogonalized proxies (denoted by $\perp$). This way, a lead-lag relationship between the proxies is allowed as a number of proxies may reflect a shift in sentiment earlier than other proxies (Brown and Cliff, 2004; Baker and Wurgler, 2006). To decide which proxies shall be included in the total shipping sentiment (SS) index, we estimate the correlation between the first-stage index and the current and lagged proxies. The proxies with the highest correlation qualify as the final sentiment proxies and the first principal component of the selected proxies provides the total sentiment index

$$SS_{j,q,t}^{total^{\perp}} = \alpha NC_{j,q,t-l}^{\perp} + \gamma MC_{j,q,t-l}^{\perp} - \delta PE_{j,q,t-l}^{\perp} + \varepsilon SNB_{j,q,t-l}^{\perp} + \omega TURN_{j,q,t-l}^{\perp}, \quad (1)$$

where j denotes the shipping market (container, drybulk, tanker), q the sector of the market (panamax, sub-panamax, handymax, …, aframax), t the current month and l the monthly lag. Table 1 presents the loadings and lags of the sentiment proxies. Looking at Table 1, we observe that the sentiment proxies enter the index with the expected sign. In terms of time order in reflecting sentiment, MC lags all other proxies while it carries the same time subscript across sectors and markets. Furthermore, NC and SNB appear to be the leading proxies, while carrying the same time subscript across sectors and markets (with the exception of the drybulk capesize sector). The variance explained by the first principal component in each sector q of market j ranges from 40% to 52% and the results are in line with Baker, Wurgler, and Yuan (2012) and

Papapostolou *et al.* (2014). Therefore, we conclude that most of the proxies' common variation is captured by one factor.

[INSERT TABLE 1 HERE]

A feature of the shipping industry is that companies operate vessels in more than one sector within a market and, as a result, sentiment may flow from one sector to another. We circumvent this issue and separate the overall market sentiment from the sector-specific sentiment. First, we construct the market sentiment index $SS_{j,t}^{market^{\perp}}$ based on the first principal component of the total sentiment indices ($SS_{q,j,t}^{total^{\perp}}$) for each market and the relevant sectors[10]:

$$SS_{container,t}^{market^{\perp}} = 0.579SS_{pana,t}^{total^{\perp}} + 0.534SS_{subpana,t}^{total^{\perp}} + 0.616SS_{hmax,t}^{total^{\perp}}, \quad (2)$$

$$SS_{drybulk,t}^{market^{\perp}} = 0.490SS_{cape,t}^{total^{\perp}} + 0.499SS_{pana,t}^{total^{\perp}} + 0.511SS_{hmax,t}^{total^{\perp}} + 0.500SS_{hsize,t}^{total^{\perp}}, \quad (3)$$

$$SS_{tanker,t}^{market^{\perp}} = 0.563SS_{vlcc,t}^{total^{\perp}} + 0.565SS_{suez,t}^{total^{\perp}} + 0.604SS_{afra,t}^{total^{\perp}}. \quad (4)$$

The sector-specific sentiment indices ($SS_{j,q,t}^{sector^{\perp}}$) are then obtained from the residuals of the regression of $SS_{j,q,t}^{total^{\perp}}$ for sector q on $SS_{j,t}^{market^{\perp}}$ of market j. The high correlation coefficients between the total and market sentiment indices, ranging from 0.88 to 0.97, imply that little

---

[10] For ease of notation, in Eqs. 2-4 we omit subscript j for each $SS_{j,q,t}^{total^{\perp}}$ originally defined in Eq. 1 as this is obvious from $SS_{j,t}^{market^{\perp}}$.

information is lost by $SS_{j,t}^{market^{\perp}}$. Further, the much lower correlation coefficients between the sector-specific indices imply that sentiment within sectors is captured more suitably by $SS_{j,q,t}^{sector^{\perp}}$ rather than $SS_{j,q,t}^{total^{\perp}}$. The full correlation matrix can be found in the Supplementary Appendix.

The market and sector-specific sentiment indices are plotted in Figure 1. It is shown that market sentiment is smooth, thus captures market-wide changes such as the recent financial and shipping crises. On the other hand, sector-specific sentiment indices move in a more erratic way and reflect the idiosyncratic features of each sector. As we are interested in the predictive power of shipping sentiment for stock returns on a global scale, we focus on the three shipping market sentiment indices[11] rather than the specific-sector indices (henceforth, for ease of notation, we use $SS_{j,t}^{\perp}$ when referring to the shipping market sentiment indices).

[INSERT FIGURE 1 HERE]

3. Predictability of International Stock Returns

We examine whether sentiment is statistically significant as a common predictor of international stock market excess returns measured in US dollar terms. We investigate the relationship between shipping sentiment and stock market excess returns of the G7 (Canada, France, Germany, Italy, Japan, United Kingdom and United States) countries and BRIC (Brazil, Russian Federation, India and China) countries that are heavily reliant on commodities and the shipping industry. All stock

[11] To validate the accuracy of the three market sentiment indices, a comprehensive empirical analysis has also been performed. The analysis builds on Papapostolou *et al.* (2014) and extends the study into the container and tanker markets. Further tests, excluding the NC and MC proxies, were also included in the analysis and the inferences remained unchanged.

market index data and the one-month US Treasury bill rate for the sample period February 1996 to April 2014 are obtained from Thomson Reuters Datastream.

We are interested in predicting stock market excess returns with shipping sentiment, but there is a possibility of causality running in the opposite direction, i.e., stock market excess returns affecting shipping sentiment. We check for time-series dependencies between the three shipping sentiment indices and stock market excess returns by performing Granger causality tests. The tests are conducted using a vector autoregressive framework and Table 2 reports the results.

[INSERT TABLE 2 HERE]

Overall, the tests in Table 2 cannot reject the null hypothesis that monthly stock market excess returns for country i (Brazil,…,United States) do not Granger-cause shipping sentiment of market j. However, the null hypothesis that shipping sentiment of market j does not Granger-cause monthly stock market excess returns for country i can be rejected. Therefore, the results provide evidence of one-way Granger causality from shipping sentiment to stock market excess returns[12].

3.1. Shipping Sentiment and Stock Market Returns

Several studies provide empirical evidence on a variety of predictors of stock market returns (Harvey, 1995; Ang and Bekaert, 2007; Welch and Goyal, 2008, among others). However, we are interested in whether shipping investor sentiment can predict international stock market excess returns. To this end, we run regressions of the type

[12] Exception to the overall results is the Canadian excess stock returns and the drybulk sentiment index, where we detect a two-way Granger causality; however, sentiment should not be regarded as a gift and as such may also be affected by market factors (Schmeling 2009). Furthermore, we identify no Granger causality between the Italian excess stock returns and the container and drybulk sentiment indices.

$$R_{i,t} = \beta_{i,0} + \beta_{i,j} SS^{\perp}_{j,t-1} + u_{i,t}\,, \tag{5}$$

i.e., the monthly return on a broad stock market index of country i and month t in excess of the one-month US Treasury bill rate ($R_{i,t}$) is regressed on the shipping sentiment index for market j that prevailed at month $t-1$. We also distinguish unique sentiment predictability effects from the Fama and French (2012) factors using the following multivariate regression model as in Baker and Wurgler (2006) and Stambaugh, Yu, and Yuan (2012),

$$R_{i,t} = \beta_{i,0} + \beta_{i,j} SS^{\perp}_{j,t-1} + \beta^{b}_{i,r} RMRF_{r,t-1} + \beta^{s}_{i,r} SMB_{r,t-1} + \beta^{h}_{i,r} HML_{r,t-1} + \beta^{w}_{i,r} WML_{r,t-1} + u_{i,t}\,. \tag{6}$$

The Fama and French (2012) control factors[13] for developed markets – namely, Global, Asia-Pacific, European, Japanese and North American factors – are constructed using six value-weight portfolios formed on size and B/M. $RMRF_{r,t}$ is region's r (Global, Asia-Pacific, Europe, Japan and North America) value-weighted return on the market portfolio over the one-month US Treasury bill rate, $SMB_{r,t}$ is the equal-weight average of the returns on the three small stock portfolios over the average of the returns on the three big stock portfolios for region r, $HML_{r,t}$ is the equal-weight average of the returns for the two high B/M portfolios minus the average of the returns for the two low B/M portfolio for region r and $WML_{r,t}$ is the equal-weight average of the returns for the two winner portfolios minus the average of the returns for the two loser portfolios

[13] The Fama and French factors for developed countries can be downloaded from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

for region r. We control for global factors, estimate the regression model of Eq. 6 and present the results in Table 3. The market excess returns are also categorized into Asia-Pacific, European, Japanese and North American regions, and Eq. 6 is re-estimated by accounting for the relevant Fama and French (2012) factors. The results are provided in the Supplementary Appendix.

Ordinary least squares (OLS) estimates of $\beta_{i,j}$ (denoted by $\hat{\beta}_{i,j}$) in the predictive regression Eq. 5 (Eq. 6) with the associated Newey-West *t*-statistics are reported in the second, sixth and tenth (fourth, eighth and twelfth) columns of Table 3. The constructed sentiment indices used in the predictive regressions are stationary and highly persistent and, as such, may impute biased coefficient and standard error estimates (Stambaugh, 1999; Ferson, Sarkissian, and Simin, 2003). To adjust for these biases and increase the robustness of our empirical results, we compute empirical *p*-values using the modified version of the wild bootstrap procedure (Gonçalves and Kilian, 2004; Cavaliere, Rahbek, and Taylor, 2010) in Rapach, Strauss, and Zhou (2013). We also account for the fact that investor sentiment is generally perceived as a contrarian predictor of stock returns and, to make our tests more powerful, we calculate *p*-values for the one-sided alternative hypothesis $\beta_{i,j} < 0$.

[INSERT TABLE 3 HERE]

Across all countries, the estimates of $\beta_{i,j}$ in Eq. 5 are negative and consistent with the existing literature[14]. The results support the contrarian feature of sentiment in relation to stock market excess returns. In particular, it is shown that when shipping sentiment is high, stock market

[14] The only exception to the rule is the statistical insignificance of the container and drybulk sentiment indices for the Italian stock market excess returns.

excess returns are lower over the coming month. The container and drybulk sentiment indices appear to be significant predictors of excess returns, but their predictive power is significantly lower compared to the tanker sentiment. All $\hat{\beta}_{i,tanker}$ estimates are higher than the corresponding $\hat{\beta}_{i,container}$ and $\hat{\beta}_{i,drybulk}$ estimates and this is also confirmed by the respective higher R^2 statistics (with the exception of the Chinese stock market excess returns). For example, a one-unit increase in tanker sentiment (which implies an increase in the standard deviation by one, as the indices are standardized) is associated with a 0.785% lower monthly excess return on the US stock market. The corresponding percentage reduction implied by the container and drybulk sentiment indices is 0.437% and 0.394% respectively.

The estimated coefficients on shipping sentiment diminish when we control for $RMRF_{r,t}$, $SMB_{r,t}$, $HML_{r,t}$ and $WML_{r,t}$ (Eq. 6). We observe a markedly reduction in terms of statistical significance and predictive power for the container and drybulk sentiment indices. The container sentiment index appears to be a significant predictor of excess returns only for the Chinese stock market, whereas the drybulk sentiment index only for the Chinese and Russian stock markets. In contrast, the tanker sentiment index carries the highest predictive power and the estimated coefficients are statistically significant in six out of eleven countries.

The monthly R^2 statistics of Eq. 5 in the third, seventh and eleventh columns of Table 3 are higher than those reported in the literature (Baker and Wurgler, 2006; Baker, Wurgler, and Yuan, 2012; Huang *et al.*, 2015) suggesting that shipping sentiment can predict monthly stock market excess returns remarkably well. Asterisks attached to the R^2 statistics of Eq. 6 in the fifth, ninth and thirteenth columns indicate significance at the 10% level or better based on the wild bootstrapped *p*-values for testing the null hypothesis of no excess return predictability; the null hypothesis is rejected in all cases.

The pooled version of Eq. 5 and Eq. 6 is also estimated, as in Ang and Bekaert (2007) and Hjalmarsson (2010), by a generalized method of moments (GMM) procedure that accounts for possible heteroskedasticity and correlation among the market excess returns. The pooled models impose the restrictions that $\beta_{i,j} = \bar{\beta}_j,\ \beta_{i,r}^b = \bar{\beta}_r^b,\ \beta_{i,r}^s = \bar{\beta}_r^s,\ \beta_{i,r}^h = \bar{\beta}_r^h$ and $\beta_{i,r}^w = \bar{\beta}_r^w$. The pooled estimates are statistically significant at conventional levels based on the wild bootstrapped *p*-values. The only exceptions are the container and drybulk sentiment indices when controlling for the global factors. Furthermore, the tanker sentiment remains the strongest predictor. Finally, following Rapach, Strauss, and Zhou (2013), we use the multi-predictor augmented regression method (mARM) of Amihud, Hurvish, and Wang (2009) to check whether the wild bootstrap adequately adjusts for any biased coefficient and standard error estimates. mARM is explicitly designed to account for the Stambaugh (1999) bias which leads to over-rejection of the null hypothesis of no predictability. We conclude that the wild bootstrap is better for making inferences as we find *p*-values that are higher than the mARM *p*-values. In the case of Eq. 6, the wild bootstrapped *p*-values produce also fewer rejections of the null hypothesis of no return predictability (the *p*-values of the wild bootstrap and mARM are available by the authors upon request).

Overall, our results confirm the significance of shipping sentiment as a contrarian common predictor of international stock returns and highlight the superior predictive power of tanker sentiment compared to the container and drybulk sentiment. As such, we also provide evidence that emphasis should be placed on the shipping industry at the aggregate level and not only on the drybulk market of the industry and the BDI.

4. Out-of-Sample Testing

In-sample analysis provides efficient parameter estimates by utilizing the full sample, but it may cause a look-ahead bias. Welch and Goyal (2008) argue that out-of-sample testing must be used to assess genuine return predictability and to avoid in-sample over-fitting issues. In addition, out-of-sample tests are less affected by short sample periods that may cause distortions (Busetti and Marcucci, 2012). Welch and Goyal (2008) also show that excess return forecasts from predictive regressions on individual economic variables typically fail to outperform the historical average benchmark forecast in out-of-sample tests. We thus test whether the unrestricted model that uses lagged shipping sentiment index as a predictor can outperform the historical average benchmark model (restricted model) of excess stock returns. The unrestricted model in our tests is,

$$R_{i,t+1} = \beta_{i,0} + \beta_{i,j} SS_{j,t}^{\perp} + u_{i,t+1} \tag{7}$$

whereas in the restricted model we impose $\beta_{i,j} = 0$.

First, we define an in-sample period D to estimate the initial parameters for the unrestricted and restricted models. Next, we update the sentiment proxies to construct the shipping sentiment indices for months $t = D+1, D+2, \dots, T$ and obtain excess returns forecasts $\hat{R}_{i,t+1}$ based on the unrestricted model of Eq. 7, and excess returns forecasts $\bar{R}_{i,t+1}$ based on the restricted model. Hansen and Timmermann (2012) show that out-of-sample tests of predictive ability have better size properties when the forecast evaluation period is a relatively large proportion of the full sample. Therefore, the out-of-sample forecasts are computed recursively, with April 1996 to March 2003 and April 2003 to April 2014, respectively, as our initial estimation period (39% of full sample) and forecast evaluation period (61% of full sample). This way we have sufficient observations to accurately estimate the initial parameters and a long enough out-of-sample period for the forecast evaluation.

We evaluate the out-of-sample performance in terms of prediction errors based on Campbell and Thompson (2008) out-of-sample R^2 statistic

$$R_{OS}^2 = 1 - \sum_{t=D}^{T-1} \left(R_{i,t+1} - \hat{R}_{i,t+1}\right)^2 \Big/ \sum_{t=D}^{T-1} \left(R_{i,t+1} - \bar{R}_{i,t+1}\right)^2 ; \tag{8}$$

and the Clark and West (2007) adjusted mean-squared prediction error (MSPE-adjusted) statistic. R_{OS}^2 lies in the range $(-\infty, 1)$ and measures the proportional reduction in MSPE for the unrestricted forecasting model relative to the historical average benchmark, i.e., when $R_{OS}^2 > 0$ the unrestricted predictive regression forecast outperforms the historical average benchmark forecast in terms of MSPE. The MSPE-adjusted statistic tests the null hypothesis of equal MSPE $(H_0: R_{OS}^2 = 0)$ against the alternative that the unrestricted forecasting model has a lower MSPE than the historical average benchmark model $(H_\alpha: R_{OS}^2 > 0)$. As in Clark and McCracken (2012) and Rapach, Strauss, and Zhou (2013), the critical values for the MSPE-adjusted statistics are computed using the wild bootstrap procedure. We also calculate $R_{OS,\mathrm{CTT}}^2$ statistics using the Campbell and Thompson (2008) truncation (CTT) approach that sets $\hat{R}_{i,t+1} = max\left(0, \hat{\beta}_{i,0} + \hat{\beta}_{i,j} SS_{j,t}^{\perp}\right)$, where $\hat{\beta}_{i,0}$ and $\hat{\beta}_{i,j}$ are the estimates of $\beta_{i,0}$ and $\beta_{i,j}$ in Eq. 7.

[INSERT TABLE 4 HERE]

The second, fourth and sixth columns of Table 4, Panel A, report the R^2_{OS} statistics. Overall, we observe that the statistics are positive and indicate that the unrestricted model, which uses the information contained in lagged shipping sentiment, outperforms the historical average benchmark forecast model. Equivalently, we can say that the unrestricted model has lower MSPE than the restricted model which ignores shipping sentiment. The individual country statistics are sizeable, ranging on average from -0.11% to 3.99%, and statistically significant based on the MSPE-adjusted statistic[15]. Consistently with our in-sample results in Section 3.1., the individual country and average values of the R^2_{OS} statistics indicate that tanker sentiment has the highest predictive power out of the three sentiment indices. Overall, we conclude that the shipping sentiment indices are economically significant as a monthly out-of-sample R^2 of 0.5% may be perceived as a signal of substantial economic value (Campbell and Thompson, 2008). Interestingly, when we apply the CTT approach, the average $R^2_{OS,\mathrm{CTT}}$ statistics for the drybulk and tanker sentiment are lower, whereas for the container sentiment are higher. However, in the case of tanker sentiment, we also observe an increase in individual $R^2_{OS,\mathrm{CTT}}$ statistics for Brazil, India, Japan and United Kingdom. Even so, both of the unconstrained R^2_{OS} and constrained $R^2_{OS,\mathrm{CTT}}$ statistics suggest stronger out-of-sample predictive ability of tanker sentiment compared to the other two sentiment indices. The reduction in the constrained statistics for the tanker sentiment (although still high and significant) indicates that the predictive ability of the index works better when negative forecasts are taken into account as well. We attribute this to the fact that sentiment may play a more important role during recessions as shown by García (2013), who argues that predictability of stock returns using

[15] Exceptions where the historical average benchmark model outperforms the unrestricted model are the following: (1) the forecasts for the Brazilian, Canadian, German, Indian, Italian and Japanese stock market excess returns that are based on the container sentiment; and (2) the forecasts for the French, German and Italian stock market excess returns that are based on the drybulk sentiment.

sentiment is concentrated in recessions. We also estimate statistics $R^2_{OS,P}$ for the pooled version of Eq. 7 that imposes the restriction $\beta_{i,j} = \bar{\beta}_j$. In terms of bias-efficiency trade-off, the pooling restriction may improve the forecasting performance of the unrestricted forecasting model (Hjalmarsson, 2010). Looking at columns three, five and seven of Table 4, we observe that the $R^2_{OS,P}$ statistics are generally higher than the R^2_{OS} statistics and provide further evidence that excess return forecasts based on lagged shipping sentiment outperform the historical average benchmark forecasts.

[INSERT FIGURE 2 HERE]

To assess the consistency of out-of-sample gains, Welch and Goyal (2008) and Rapach, Strauss, and Zhou (2013) plot the cumulative differences of squared predicted errors of the historical average benchmark forecasts from the unrestricted model forecasts. One can determine whether the unrestricted model outperforms the historical average benchmark model by comparing the height of the plot at the beginning and end of the out-of-sample period; a higher value at the end of the out-of-sample period reflects a lower MSPE for the unrestricted model. The cumulative differences of squared forecast errors are plotted in Figure 2 and resemble the statistics presented in Table 4. The tanker sentiment as a predictor is superior to the container and drybulk sentiment and provides out-of-sample forecasting gains across all countries. Furthermore, some spikes in forecasting gains, in the case of the unconstrained R^2_{OS}, are concentrated around business cycle recessions (shaded areas in Figure 2 represent country-specific business cycle recessions according to the Economic Cycle Research Institute – www.businesscycle.com), and this is an indication of the importance of negative forecasts on the predictive ability of shipping sentiment. This phenomenon is in line with the results of Rapach, Strauss, and Zhou (2013) and Huang *et al*. (2015); Rapach, Strauss and Zhou (2010) and Henkel, Martin and Nardari (2011) who provide evidence that stock return predictability is mainly confined to economic recessions and base this finding on the fact that expected stock returns vary more during economic recessions than expansions.

Finally, we conduct additional tests to investigate the possibility of spurious out-of-sample forecasting performance associated with high R^2_{OS} statistics[16]. Hansen and Timmermann (2012) argue that the danger of spurious evidence of predictability induced by the search over the split point of the sample tends to be associated with short evaluation periods that start late in the sample. To ensure that our sample is not split in a way to favor our models, we calculate the R^2_{OS} and $R^2_{OS,\text{CTT}}$ statistics and the corresponding *p*-values for various forecast evaluation periods running from April 2003 through April 2013, with the end of the out-of-sample period fixed at April 2014. The figures of the statistics and *p*-values are provided in the Supplementary Appendix. In the case of container and drybulk sentiment, the *p*-values of R^2_{OS} and $R^2_{OS,\text{CTT}}$ are consistently above 10% with minor exceptions: the R^2_{OS} of container sentiment is positive and consistently significant at the 10% level up to the start of the recent subprime and financial crises only for China; the $R^2_{OS,\text{CTT}}$ is positive and consistently significant at the 10% level up to the end of the crises only for Canada and USA; and in the case of drybulk sentiment, only the R^2_{OS} statistics for Japan, China and Russia are positive and consistently significant at the 10% level up to the middle of the crises. Therefore, we cannot conclude that the lagged container and drybulk sentiment generates more precise forecasts of stock excess returns than the historical average benchmark model. Turning to the

[16] We note that in a separate contribution, Gargano and Timmermann (2012) report significant monthly R^2_{OS} statistics exceeding 4% when using the default return spread (difference between long-term corporate and government bonds return) to predict metals and raw industrials commodity spot price indices.

tanker sentiment, we can observe a significant difference: the R_{OS}^2 and $R_{OS,\mathrm{CTT}}^2$ statistics are on average positive and statistically significant for all countries (with the exception of Italy and Japan in the case of $R_{OS,\mathrm{CTT}}^2$) and up to the recent crises. This indicates that forecasts that start after the crises and are closer to the end of the full sample are unlikely to outperform the historical average forecasts. Interestingly, for forecast evaluation periods starting within 16 months following the end of the subprime and financial crises, we find *p*-values above the 10% threshold before falling below that level again; this is observed for Germany, France, Japan, UK and USA and is more evident in the case of the $R_{OS,\mathrm{CTT}}^2$ statistics. Finally, in the case of UK and USA, $R_{OS,\mathrm{CTT}}^2$ remains consistently below the 10% level throughout the various forecast evaluation periods, with only a slight increase above the threshold just after the end of the crises.

Overall, we conclude that our R_{OS}^2 and $R_{OS,\mathrm{CTT}}^2$ statistics could be higher and statistically significant if we had selected a forecast evaluation period starting later than April 2003, which forms evidence against spurious forecasting performance. Furthermore, it is clear that stock returns predictability based on tanker sentiment consistently outperforms the historical average benchmark and is statistically significant at the 10% level. However, the tanker predictability power disappears when the forecast evaluation periods start within 16 months following the end of the subprime and financial crises. The loss of predictability during this period can be attributed to stock markets worldwide experiencing a speedy and steady advance to higher levels, and the fact that tanker sentiment predictability power generally works best when negative forecasts are taken into consideration.

5. Economic Value of Shipping Sentiment and Asset Allocation Implications

To examine the economic value of the forecasts based on the lagged shipping sentiment indices, we compute the certainty equivalent return (CER) gain and Sharpe ratio for a mean-variance investor (Kandel and Stambaugh, 1996; Campbell and Thompson, 2008; Ferreira and Santa-Clara, 2011; Huang *et al.*, 2015). We assume that the investor takes positions across the stock index and the risk-free T-bills using the unrestricted regression forecasts generated by Eq. 7. At the end of period t, the investor assigns portfolio weight

$$w_{i,t} = \frac{1}{\gamma}\frac{\hat{R}_{i,t+1}}{\hat{\sigma}^2_{i,t+1}}, \tag{9}$$

to the stock index of country i for period $t+1$, where $\gamma = 1,3,5$ is the risk aversion coefficient and $\hat{\sigma}^2_{i,t+1}$ is the variance forecast. In addition, the investor allocates $1 - w_{i,t}$ of the portfolio to US Treasury bills, so that the realized portfolio return for period $t+1$ is

$$R^p_{i,t+1} = w_{i,t}R_{i,t+1} + RF_{t+1} \tag{10}$$

where RF_{t+1} is the gross risk-free return. Following Huang *et al.* (2015), we use a five-year moving window of past monthly returns to calculate the variance of the excess stock market return. We further constrain $w_{i,t}$ to lie between 0 and 1 to forbid short sales and allow no leverage, and assume transactions costs of 50 bps.

The CER of the portfolio for an investor who trades the stock index of country i is

$$\text{CER}_{p,i} = \hat{\mu}_{p,i} - 0.5\gamma\hat{\sigma}^2_{p,i} \tag{11}$$

where $\hat{\mu}_{p,i}$ and $\hat{\sigma}^2_{p,i}$ are the portfolio's sample mean and variance respectively over the forecasting evaluation period. The CER gain is the difference between the CER of the investor who employs the forecasts generated by Eq. 7 and the CER of an investor who instead uses the historical average forecasts. The annualized difference represents the annual portfolio management fee that an investor would be willing to incur for accessing the forecasts based on shipping sentiment. In addition, we calculate the portfolio's Sharpe ratio by dividing the portfolio mean excess return by the standard deviation of the portfolio excess return. To further assess the statistical significance of the CER gain and the Sharpe Ratio, we employ the stationary bootstrap[17].

[INSERT TABLE 5 HERE]

Table 5 shows that forecasts based on tanker shipping sentiment can generate economic gains for a mean-variance investor, whereas the gains are limited when the forecasts are based on the container and drybulk sentiment indices. Specifically, the container sentiment has an average

[17] As pointed out by Sullivan, Timmermann, and White (1999), data snooping occurs when a dataset is used more than once for selection and inference purposes. Data snooping can increase the probability of having satisfactory results purely to chance or the use of posterior information, rather than the superior ability of the shipping sentiment based forecasts. We implement a reality check using 10,000 bootstrap simulations and generate artificial time series for the realized returns of the constructed portfolios; thus, generating a series of empirical distributions of mean returns and Sharpe ratios. We define loss function (LF) differentials between the unrestricted and historical average benchmark models based on the CER gain, $LF_i = CER^U_{p,i} - CER^R_{p,i}$, and Sharpe ratio statistics, $LF_i = SR^U_{p,i} - SR^R_{p,i}$ (where superscripts U and R denote the portfolios based on the unrestricted and the restricted models respectively), and generate collection $\left\{LF_i^{(Z)}\right\}_{W=1}^{10,000}$ where $LF_i^{(Z)}$ is the estimated statistic from the Zth bootstrapped sample. Then we test the null hypothesis that the portfolio based on the historical average forecasts is not outperformed by the portfolio based on the forecasts based on shipping sentiment, i.e., $H_0: E(LF_{i,Z}) \leq 0$. The bootstrapped *p*-values are given by the proportions of negative (one-tail test) $LF_i^{(Z)}$ across $Z = 1, \ldots, 10{,}000$. A description of the stationary bootstrap algorithm can be found in Politis and Romano (1994) and Sullivan, Timmermann, and White (1999).

CER gain of 0.57% when the risk aversion coefficient is 1, and the average gain drops to 0.23% and 0.16% when the risk aversion coefficients are 3 and 5 respectively. The average CER gain for the drybulk sentiment based forecasts is 0.94%, 0.35% and 0.23% when the risk aversion is 1, 3 and 5 respectively. The tanker sentiment stands out again in terms of economic value and statistical significance, with an average CER gain of 3.67% when the risk aversion is 1, and smaller positive average CER gains of 1.40% and 0.97% when the risk aversion is 3 and 5 respectively. In terms of Sharpe ratios the inferences remain unchanged.

Overall, the results demonstrate that the tanker sentiment can offer economic value for a mean-variance investor. More specifically, an investor would be willing to incur an average annual portfolio management fee up to 3.67% to have access to the forecasts based on the tanker sentiment instead of using the historical average forecasts.

6. Additional Robustness Checks

6.1. Alternative Predictors of Stock Returns

We employ the Baker and Wurgler (2006) investor sentiment index (S^{BW}) and the Huang *et al.* (2015) aligned investor sentiment index (S^{PLS}) as alternative measures[18] of investor sentiment and test whether shipping sentiment still carries unique significant predictive power for stock returns. Both indices are based on US sentiment proxies and our analysis is focused on the excess returns of the US stock market accordingly. To this end, we estimate,

$$R_{US,t} = \beta_{US,0} + \beta_{US,j} SS^{\perp}_{j,t-1} + \beta^{k}_{US} S^{k}_{t-1} + u_{US,t}, \qquad k = BW, PLS, \tag{12}$$

and

[18] S^{BW} and S^{PLS} can be downloaded from http://apps.olin.wustl.edu/faculty/zhou/.

$$R_{US,t} = \beta_{US,0} + \beta_{US,j} SS_{j,t-1}^{\perp} + \beta_{US}^{k} S_{t-1}^{k} + \beta_{US}^{b} RMRF_{US,t-1} + \beta_{US}^{s} SMB_{US,t-1} + \beta_{US}^{h} HML_{US,t-1} + \beta_{US}^{w} WML_{US,t-1} + u_{US,t}\ , \tag{13}$$

where $R_{US,t}$ is the excess return on the US stock market index, S_{t-1}^{BW} the orthogonalized BW investor sentiment index and S_{t-1}^{PLS} the orthogonalized aligned investor sentiment index. Eqs. 12-13 are estimated as described in Section 3.1.

[INSERT TABLE 6 HERE]

The second, third, fifth and sixth columns of Table 6 report the estimates of $\beta_{US,j}$ and β_{US}^{k} (denoted by $\hat{\beta}_{US,j}$ and $\hat{\beta}_{US}^{k}$) of Eqs. 12-13. When we control for the S^{BW} index we observe that the container and drybulk sentiment indices are not statistically significant, whereas the estimates for the tanker sentiment and S^{BW} index are negative and statistically significant. When we estimate Eq. 13, where the Fama and French US factors are also accounted for, we find that the container and drybulk sentiment indices are still insignificant in predicting stock market returns while the S^{BW} index is negative and statistically significant. On the contrary, the tanker sentiment index is capable of predicting stock returns, whereas the S^{BW} index is statistically insignificant. The results remain qualitatively similar when we use the S^{PLS} index as a control variable. Overall, the robustness check shows that only the tanker sentiment retains its unique predictive power and that it is a better predictor of the US stock market excess returns than the container and drybulk sentiment indices.

Next, we check whether the shipping sentiment indices capture different elements of sentiment from the US investor sentiment indices. The correlation coefficients between the shipping sentiment indices and the US investor sentiment indices are modest and show evidence against a tight statistical link, whereas no time-series dependencies are detected between them according to the Granger causality tests[19]. To further ensure that the shipping sentiment indices do not include any component that stems from the US investor sentiment indices, we orthogonalize our sentiment proxies to the US investor sentiment indices and re-construct the shipping sentiment indices. We then re-estimate Eqs. 12-13 and the results[20] are consistent with our original analysis, although the container sentiment becomes statistically significant when we control for S^{PLS}. Therefore, taking into account the fact that the shipping sentiment indices are constructed by proxies outside of the US financial markets and the unique predictive power of tanker sentiment that is orthogonalized to US investor sentiment, we conclude that the shipping sentiment indices do not include any component of US investor sentiment and capture different elements of sentiment.

The US consumer sentiment index provided by Thomson Reuters/University of Michigan could also be utilized in the robustness check tests as an alternative behavioral predictor. Huang *et al*. (2015) test the Michigan Consumer Sentiment Index and find that it fails to forecast significantly aggregate stock returns in the US. Furthermore, the Michigan Consumer Sentiment Index measures the current and expected financial conditions of households as determined by consumer opinion; whereas the S^{BW} and S^{PLS} indices reveal investors' sentiment about the stock market which is captured by market sentiment proxies. As such, in terms of information revealed and sentiment proxies employed, our shipping sentiment indices share analogies with the S^{BW} and

[19] The full sample correlation coefficients for S_t^{BW} (S_t^{PLS}) and $SS_{j,t}^{\perp}$ range between 0.133 (0.084) to 0.332 (0.225). The correlation coefficients for S_t^{BW} (S_t^{PLS}) and $SS_{j,t}^{\perp}$ when the sample is split in two halves range between -0.221 (0.118) to 0.317 (0.331) for the first half and between 0.133 (-0.021) to 0.327 (0.299) for the second half. We also find no Granger causality between the indices and the results are consistent for 1, 6 and 12 lags.

[20] The results are available from the authors upon request.

S^{PLS} indices. Therefore, we avoid the flaws of survey-based sentiment indices by not utilizing the Michigan Consumer Sentiment Index, as what participants reply and actually do can differ substantially.

Furthermore, we perform out-of-sample tests (see Section 4) to compare the excess return forecasts computed by the two unrestricted models $R_{US,t+1} = \beta_{US,0} + \beta_{US,j} SS_{j,t}^{\perp} + u_{US,t+1}$ and $R_{US,t+1} = \beta_{US,0} + \beta_{US}^{k} S_t^{k} + u_{US,t+1}$. The R_{OS}^2 and MSPE-adjusted statistics (provided in the Supplementary Appendix) also suggest that tanker sentiment has stronger out-of-sample forecasting power than the S^{BW} and S^{PLS} indices. In addition, we check in an in- and out-of-sample framework whether shipping sentiment is a significant predictor for stock returns when we consider the 14 economic predictors[21] employed by Welch and Goyal (2008) as control variables. Our inferences with regard to the unique predictive power of shipping sentiment remain unchanged and the estimation results can be found in the Supplementary Appendix. Finally, to discount the possibility that the predictive ability of shipping sentiment is transmitted to the international stock markets through shipping stocks[22], we also test separately the impact of shipping sentiment on a portfolio of shipping and non-shipping stocks. We construct an equally-weighted US shipping stock index using all shipping stocks trading in the US equity market and an orthogonalized – to the shipping stocks – US stock market index. The results (available from the authors) remain qualitatively and quantitatively similar to the original inferences in the paper. Therefore, we can

[21] The economic variables can be downloaded from http://www.hec.unil.ch/agoyal/.

[22] We believe that is highly unlikely for the predictability to be transmitted to the international stock markets through shipping stocks, since there are only few shipping stocks trading on stock exchanges internationally and their low market capitalization plays a minor role on the general stock market indices and performance. For example, in April 2014 the market capitalization of 45 shipping stocks in the US equity markets was standing at $US 38.2 billion, representing a market share of 0.19% and 0.23% in terms of the NYSE and S&P500 total market capitalization respectively (we concentrate on the US stock market as this is, by far, the largest and most liquid market for shipping stocks).

safely conclude that shipping sentiment – even though industry related – can also be utilized to explain market-wide stock returns; probably due to its significance in international trade and economies as mentioned in the introduction.

6.2. Shipping Sentiment and the Physical market

As a second robustness check, we test whether the drybulk or the tanker physical markets are superior predictors of stock market excess returns compared to the drybulk and tanker sentiment indices. In our analysis, the Baltic Dry Index (BDI) and the West Texas Intermediate (WTI) oil price are employed as proxies for the drybulk and the tanker physical markets respectively. To perform this type of robustness check, we run the following regressions:

$$R_{i,t} = \beta_{i,0} + \beta_i^k R_{t-1}^k + u_{i,t}, \qquad k = BDI, WTI, \tag{14}$$

where R_{t-1}^{BDI} is the monthly return on BDI at month $t-1$ and R_{t-1}^{WTI} is the monthly price change of WTI (for more details about the estimation see Section 3.1.).

[INSERT TABLE 7 HERE]

The second and fourth columns of Table 7 report the estimates of β_i^{BDI} and β_i^{WTI} (denoted by $\hat{\beta}_i^{BDI}$ and $\hat{\beta}_i^{WTI}$) of Eq. 14. In terms of the drybulk physical market and BDI, $\hat{\beta}_i^{BDI}$ is statistically significant in explaining stock market excess returns in nine out of eleven countries. Further, the BDI is a superior predictor – either in terms of variance explained or statistical significance – only for the Canadian and Italian stock market excess returns[23]. As such, the robustness check results confirm the superiority of drybulk sentiment. Next, we find no predictive ability for WTI as $\hat{\beta}_i^{WTI}$ is statistically insignificant in explaining stock market excess returns in all cases. Therefore, the question of whether tanker sentiment carries superior predictive power compared to the WTI measure is straightforward to answer. The results are not in line with Driesprong, Jacobsen, and Maat (2008) who argue that changes in oil prices predict stock market returns worldwide. However, one needs to take into account the different sample periods that the empirical results are based on. The weak predictability of stock market returns by oil price changes may be attributed to the subprime and financial crises included in the sample. When these periods are excluded, as we show in the next Section, changes in oil prices can predict stock market returns. As a result, we may assume that the established negative relationship between oil price changes and stock returns did not hold during the recent crises, as plummeting oil prices were followed by abysmal stock returns. We also perform out-of-sample tests (see Section 6) to compare the excess return forecasts computed by the following two unrestricted models $R_{i,t+1} = \beta_{i,0} + \beta_{i,j} SS_{j,t}^{\perp} + u_{i,t+1}$, $R_{i,t+1} = \beta_{i,0} + \beta_i^k R_t^k + u_{i,t+1}$. The R_{OS}^2 and MSPE-adjusted statistics (provided in the Supplementary Appendix) suggest that the sentiment indices display stronger out-of-sample forecasting power relative to the drybulk and tanker physical markets proxies.

Finally, we check whether the tanker sentiment is a proxy of oil price returns. The correlation coefficient between tanker sentiment and oil price returns is modest[24] and the Granger causality tests show that there is no feedback from oil price changes to tanker sentiment, whereas the null hypothesis that tanker sentiment does not Granger-cause oil price returns can be rejected.

[23] The cyclical component of BDI, given by the one-sided HP filter, was also tested as a possible proxy for the drybulk physical market and the results remained qualitatively similar.

[24] The full sample correlation coefficient is -0.214; and -0.013 and -0.252 when the sample is split in two halves.

The tanker sentiment index is already orthogonalized to macroeconomic factors; however, to ensure that it does not include any component that stems from monthly oil price returns and possibly other macroeconomic determinants[25], we orthogonalize the tanker sentiment proxies to oil price returns and re-construct the index. The results (available from the authors upon request) on the new orthogonalized tanker sentiment index as a predictor of stock excess returns corroborate the ones in Tables 3, 7 and 8. Therefore, we can be confident that the tanker sentiment index does not include any component of oil price returns and is not a proxy of oil prices.

6.3. Robustness of Results to Subprime and Financial Crises

A natural question that may arise is whether our results still hold when the recent subprime and financial crises are excluded from the full sample period. We follow the estimation procedure described in Section 3.1. and replicate the tests for Eqs. 5 and 14 by examining the sample period that excludes the subprime and financial crises[26].

[INSERT TABLE 8 HERE]

The second, fourth and sixth columns of Table 8 report the estimates of $\beta_{i,j}$ of Eq. 5, and the eighth and tenth columns the estimates of β_i^{BDI} and β_i^{WTI} of Eq. 14. The reduced sample results record a significant reduction of predictive power for the container sentiment index and may raise

[25] Much of the oil price surge between 2003 and mid-2008 has been associated with an unexpectedly booming world economy, notable in emerging Asia (Kilian and Hicks, 2013; Kilian and Murphy, 2014).

[26] The exclusion period is December 2007 to June 2009. We choose December 2007 as the starting point as the first market interventions by central banks occurred during this month. December 12th, 2007: the Central banks of US, the European Union, Canada and Switzerland announce a plan to provide at least US$90 billion in short-term financing to banks. December 18th, 2007: The European Central Bank injects US$500 billion into the financial system and The Bank of England auctions off US$20 billion in three-month loans (source: Bloomberg). Further, we end the period in June 2009, which is the month that the US economy exited recession.

questions with regard to its consistency as predictor of stock market excess returns. For instance, $\hat{\beta}_{i,container}$ coefficients are statistically significant in explaining stock market excess returns in two out of eleven countries. The results for the drybulk sentiment index and BDI are also weakened, as the $\hat{\beta}_{i,drybulk}$ and $\hat{\beta}_{i,bdi}$ coefficients are statistically significant predictors of stock market excess returns in five and two out of eleven countries respectively. In contrast, the results for WTI are strengthened and in line with Driesprong, Jacobsen, and Maat (2008). Finally, our initial inferences in relation to tanker sentiment do not change, as the results remain qualitatively and quantitatively similar compared to the results of the full sample period. Hence, the consistent superior predictive power of tanker sentiment is once more highlighted.

7. Conclusions

The predictability of stock returns by investor sentiment has been subject to constant updating over time. Despite the mixed evidence, there seems to be little chance of reaching a decisive conclusion using US data alone. Surprisingly, up until now there has been only one study (Baker, Wurgler, and Yuan, 2012) that uses investor sentiment indices based on market actions and data other than that of the US stock market. Yet, this study is concerned with in-sample predictability of stock returns and, as such, there is no guarantee that in-sample return predictability can produce more accurate forecasts compared to the historical average benchmark forecasts. To overcome the unavailability of common measures of investor sentiment for stock markets internationally, we use shipping sentiment as a common predictor of international stock returns.

Our analysis confirms empirically the significance of tanker sentiment as a contrarian global predictor of financial assets in an in- and out-of-sample framework. The out-of-sample R^2 statistics are sizeable and we find that tanker sentiment is not only statistically but also economically significant in providing utility gains to a mean-variance investor. The predictability power of tanker sentiment is stronger when negative forecasts are taken into account, and this is evident by the loss of power observed for forecasting evaluation periods that start within 16 months following the end of the subprime and financial crises. Finally, contrary to the common belief in the financial press and finance literature that the sole interest should be on the drybulk market and the BDI, we show evidence that one should concentrate on the shipping industry at the aggregate level, with more attention paid on the tanker rather than the drybulk market.

References

Alizadeh, A.H., and G.Y. Muradoglu. “Stock market efficiency and international shipping market information.” *Journal of International Financial Markets*, *Institutions and Money,* 33 (2014), 445-461.

Amihud, Y.; C.M. Huvrich; and Y. Wang. “Multiple-predictor regressions: hypothesis tests.” *Review of Financial Studies,* 22 (2009), 413-434.

Ang, A., and G. Bekaert. “Return predictability: is it there?” *Review of Financial Studies,* 20 (2007), 651-707.

Antoniou, C.; J.A. Doukas; and A. Subrahmanyam. “Cognitive dissonance, sentiment and momentum.” *Journal of Financial and Quantitative Analysis,* 48 (2013), 245-275.

Apergis, N., and J.E. Payne. “New evidence on the information and predictive content of the Baltic Dry Index.” *International Journal of Financial Studies,* 1 (2013), 62-80.

Baker, M., and J. Wurgler. “A catering theory of dividends.” *Journal of Finance,* 59 (2004), 1125-1165.

Baker, M., and J. Wurgler. “Investor sentiment and the cross-section of stock returns.” *Journal of Finance,* 61 (2006), 1645-1680.

Baker, M., and J. Wurgler. “Investor sentiment in the stock market.” *Journal of Economic Perspectives*, 21 (2007), 129-151.

Baker, M.; J. Wurgler; and Y. Yuan. “Global, local and contagious investor sentiment.” *Journal of Financial Economics,* 104 (2012), 272-287.

Barber, B.M.; T. Odean; and N. Zhu. “Do retail trades move markets?” *Review of Financial Studies*, 22 (2009), 151-186.

Barberis, N.; A. Shleifer; and R. Vishny. “A model of investor sentiment.” *Journal of Financial Economics*, 49 (1998), 307-343.

Busetti, F., and J. Marcucci. “Comparing forecast accuracy: a Monte Carlo investigation.” *International Journal of Forecasting*, 29 (2012), 13-27.

Brown, G.W., and M.T. Cliff. “Investor sentiment and the near-term stock market.” *Journal of Empirical Finance*, 11 (2004), 1-27.

Campbell, J.Y., and S.B. Thompson. “Predicting the equity premium out of sample: can anything beat the historical average?” *Review of Financial Studies*, 21 (2008), 1509-1531.

Cavaliere, G.; A. Rahbek; and R.A.M. Taylor. “Cointegration rank testing under conditional heteroskedasticity.” *Econometric Theory*, 26 (2010), 1719-1760.

Clark, T.E., and W. McCracken. “Reality checks and nested forecast model comparisons.” *Journal of Business and Economic Statistics*, 30 (2012), 85-110.

Clark, T.E., and K.D. West. “Approximately normal tests for equal predictive accuracy in nested models.” *Journal of Econometrics*, 138 (2007), 291-311.

Das, S.R., and M.Y. Chen. “Yahoo! for Amazon: Sentiment extraction from small talk on the web.” *Management Science*, 53 (2007), 1375-1388.

Driesprong, G.; B. Jacobsen; and B. Maat. “Striking oil: Another puzzle?” *Journal of Financial Economics*, 89 (2008), 307-327.

Fama, E.F., and K.R. French. “Common risk factors in the returns on stocks and bonds.” *Journal of Financial Economics*, 33 (1993), 3-56.

Fama, E.F., and K.R. French. “Size, value and momentum in international stock returns.” *Journal of Financial Economics*, 105 (2012), 457-472.

Ferreira, M., and P. Santa-Clara. “Forecasting stock market returns: The sum of the part is more than the whole.” *Journal of Financial Economics,* 100 (2011), 514-537.

Ferson, W.E.; S. Sarkissian; and T.T. Simin. “Spurious regressions in financial economics.” *Journal of Finance,* 58 (2003), 1393-1414.

García, D. “Sentiment during recessions.” *Journal of Finance,* 68 (2013), 1267-1300.

Gargano, A., and A. Timmermann. “Predictive dynamics in commodity prices.” Working paper, University of California, San Diego, USA and CREATES, Denmark, (2012).

Gonçalves, S., and L. Kilian. “Bootstrapping autoregressions with conditional heteroskedasticity of unknown form.” *Journal of Econometrics*, 123 (2004), 89-120.

Greenwood, R., and S.G. Hanson. “Waves in ship prices and investment.” *Quarterly Journal of Economics*, 130 (2015), 55-109.

Hansen, P.R., and A. Timmermann. “Choice of sample split in out-of-sample forecast evaluation. Working paper, European University Institute, Florence, Italy, (2012).

Harvey, C.R. “Predictable risk and returns in emerging markets.” *Review of Financial Studies*, 8 (1995), 773-812.

Henkel, S.J.; J.S. Martin; and F. Nardari. “Time-varying short-horizon predictability.” *Journal of Financial Economics*, 99 (2011), 560-580.

Hjalmarsson, E. “Predicting global stock returns.” *Journal of Financial and Quantitative Analysis,* 45 (2010), 49-80.

Huang, D.; F. Jiang; J. Tu; and G. Zhou. “Investor sentiment aligned: a powerful predictor of stock returns.” *Review of Financial Studies,* 28 (2015), 791-837.

Jones, C.M., and G. Kaul. “Oil and the stock markets.” *Journal of Finance,* 51 (1996), 463-491.

Kalouptsidi, M. “Time to build and shipping prices.” *American Economic Review,* 104 (2014), 564-608.

Kandel, S., and R.F. Stambaugh. “On the predictability of stock returns: An asset allocation perspective.” *Journal of Finance,* 51 (1996), 385-424.

Killian, L. “Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market.” American Economic Review, 99 (2009), 1053-1069.

Kilian, L. and B. Hicks. “Did unexpectedly strong economic growth cause the oil price shock of 2003-2008?” Journal of Forecasting, 32 (2013), 385-394.

Kilian, L. and D.P. Murphy. “The role of inventories and speculative trading in the global market for crude oil.” Journal of Applied Econometrics, 29 (2014), 454-478.

Lemmon, M., and E. Portniaguina. “Consumer confidence and asset prices: some empirical evidence.” *Review of Financial Studies,* 19 (2006), 1499-1529.

Neely, C.J.; D.E. Rapach; J. Tu; and G. Zhou. “Forecasting the equity risk premium: the role of technical indicators.” *Management Science*, 60 (2014), 1772-1791.

Papapostolou, N.C.; N.K. Nomikos; P.K. Pouliasis; and I. Kyriakou. “Investor sentiment for real assets: the case of drybulk shipping.” *Review of Finance,* 18 (2014), 1507-1539.

Rapach, D.E.; J.K. Strauss; and G. Zhou. “International stock return predictability: what is the role of the United States?” *Journal of Finance,* 68 (2013), 1633-1662.

Rapach, D.E.; J.K. Strauss; and G. Zhou. “Out-of-sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy.” *Review of Financial Studies,* 23 (2010), 821-862.

Schmeling, M. “Investor sentiment and stock returns: some international evidence.” *Journal of Empirical Finance,* 16 (2009), 394-408.

Smith, A. “An inquiry into the nature and causes of the wealth of nations.” London: W. Strahan and T. Cadell (1776).

Stambaugh, R.F. “Predictive regressions.” *Journal of Financial Economics,* 54 (1999), 375-421.

Stambaugh, R.F.; J. Yu; and Y. Yuan. “The short of it: investor sentiment and anomalies.” *Journal of Financial Economics*, 104 (2012), 288-302.

Stambaugh, R.F.; J. Yu; and Y. Yuan. “The long of it: odds that investor sentiment spuriously: predicts anomaly returns.” *Journal of Financial Economics,* 114 (2014), 613-619.

Sullivan, R.; A. Timmermann; and H. White. “Data-snooping, technical trading rule performance, and the bootstrap.” *Journal of Finance,* 54 (1999), 1647–1691.

United Nations Conference on Trade and Development (UNCTAD). “Review of maritime transport.” New York and Geneva: United Nations (2014).

Welch, I., and A. Goyal. “A comprehensive look at the empirical performance of equity premium prediction.” *Review of Financial Studies*, 21 (2008), 1455-1508.

Wright, F., and G. Bower. “Mood effects on subjective probability assessment.” *Organizational Behavior and Human Decision Process,* 52 (1992), 276-291.

Table 1. Total Sentiment Index Components – Loadings and Lags

$SS_{j,q,t}^{total^{\perp}} = \alpha NC_{j,q,t-l}^{\perp} + \gamma MC_{j,q,t-l}^{\perp} - \delta PE_{j,q,t-l}^{\perp} + \varepsilon SNB_{j,q,t-l}^{\perp} + \omega TURN_{j,q,t-l}^{\perp}$ (Eq. 1) is the first principal component of the five orthogonalized sentiment proxies in market *j* and sector *q* during month *t*. NC is net contracting, MC money committed, PE the price-to-earnings ratio, SNB the second-hand-to-newbuilding price ratio, and TURN the turnover ratio. The orthogonalized proxies labelled by $\perp$ are the residuals from the regression of each of the five raw sentiment proxies on the G7 industrial production growth and two-recession period dummies for the G7 and Major 5 Asia countries. The table reports the first principal component loadings α, γ, δ, ε and, ω of the five orthogonalized proxies and their respective time lags l.

q	$\alpha\ (l)$	$\gamma\ (l)$	$\delta\ (l)$	$\varepsilon\ (l)$	$\omega\ (l)$
j = Container					
Panamax	0.229 (2)	0.424 (0)	-0.480 (1)	0.630 (2)	0.374 (1)
Sub-panamax	0.542 (2)	0.537 (0)	-0.483 (0)	0.415 (2)	0.110 (1)
Handymax	0.363 (2)	0.412 (0)	-0.332 (0)	0.591 (2)	0.489 (1)
j = Drybulk					
Capesize	0.493 (2)	0.295 (0)	-0.422 (1)	0.599 (1)	0.364 (1)
Panamax	0.393 (2)	0.410 (0)	-0.422 (1)	0.575 (2)	0.411 (1)
Handymax	0.373 (2)	0.393 (0)	-0.421 (1)	0.591 (2)	0.425 (2)
Handysize	0.453 (2)	0.402 (0)	-0.372 (1)	0.546 (2)	0.443 (2)
j = Tanker					
VLCC	0.342 (2)	0.518 (0)	-0.366 (2)	0.552 (2)	0.420 (1)
Suezmax	0.065 (2)	0.447 (0)	-0.323 (2)	0.673 (2)	0.489 (0)
Aframax	0.059 (2)	0.451 (0)	-0.490 (0)	0.456 (2)	0.587 (1)

Table 2. Granger Causality Tests

The table reports causality tests between the monthly return on a broad stock market index of country i in excess of the one-month US Treasury bill $(R_{i,t})$ and the shipping sentiment index $SS_{j,t}^{\perp}$ for market j. For each sentiment index, we test the null hypothesis that stock index monthly excess return for country i does not Granger-cause shipping sentiment index of market j and, the null hypothesis that shipping sentiment index of market j does not Granger-cause stock index monthly excess return for country i. For each test, we report the χ^2-statistics and the optimal lag length is based on the Schwartz criterion. Superscripts a, b, c denote rejection of the null hypothesis at the 1%, 5%, 10% significance levels, respectively. Country abbreviations are as follows: Brazil (BRZ), Canada (CAN), France (FRA), Germany (GER), India (IND), Italy (ITA), Japan (JPN), P.R. of China (PRC), Russian Federation (RUS), United Kingdom (UK) and United States (USA).

	BRA	CAN	FRA	GER	IND	ITA	JPN	PRC	RUS	UK	USA
j = Container											
$H_0: R_{i,t} \nrightarrow SS_{j,t}^{\perp}$	0.763	0.120	0.256	0.654	0.102	0.899	0.022	0.317	0.979	0.001	0.126
$H_0: SS_{j,t}^{\perp} \nrightarrow R_{i,t}$	2.122[c]	2.752[c]	3.091[c]	3.421[c]	5.188[b]	2.207	4.981[b]	7.431[a]	2.994[c]	3.475[c]	4.224[b]
j = Drybulk											
$H_0: R_{i,t} \nrightarrow SS_{j,t}^{\perp}$	0.354	4.663[b]	1.560	2.303	1.487	0.936	2.070	0.834	0.879	0.207	1.752
$H_0: SS_{j,t}^{\perp} \nrightarrow R_{i,t}$	2.777[c]	3.584[c]	3.375[c]	2.585[c]	8.917[a]	1.667	8.441[a]	12.047[a]	5.248[b]	4.964[b]	5.709[b]
j = Tanker											
$H_0: R_{i,t} \nrightarrow SS_{j,t}^{\perp}$	0.376	1.118	1.578	1.567	0.852	2.377	1.277	0.585	1.339	0.023	0.635
$H_0: SS_{j,t}^{\perp} \nrightarrow R_{i,t}$	10.596[a]	18.291[a]	10.554[a]	6.938[a]	14.651[a]	9.165[a]	10.764[a]	8.005[a]	13.524[a]	10.671[a]	16.318[a]

Table 3. Estimation Results for Stock Index Returns

The table reports OLS estimates of $\beta_{i,j}$ (denoted by $\hat{\beta}_{i,j}$) for the regression models: $R_{i,t} = \beta_{i,0} + \beta_{i,j} SS^{\perp}_{j,t-1} + u_{i,t}$ (Eq. 5) and $R_{i,t} = \beta_{i,0} + \beta_{i,j} SS^{\perp}_{j,t-1} + \beta^{b}_{i,r} RMRF_{r,t-1} + \beta^{s}_{i,r} SMB_{r,t-1} + \beta^{h}_{i,r} HML_{r,t-1} + \beta^{w}_{i,r} WML_{r,t-1} + u_{i,t}$ (Eq. 6). $\hat{\beta}_{i,j}$ coefficients of Eq. 5 (Eq. 6) are presented in columns 2, 6 and 10 (4, 8 and 12). $R_{i,t}$ is the stock index monthly excess return for country i and $SS^{\perp}_{j,t-1}$ is the lagged shipping sentiment for market j. The control variables are the Fama and French (2012) factors for developed markets constructed using six value-weight portfolios formed on size and book-to-market: $RMRF_{r,t-1}$ is region's r value-weighted return on the market portfolio minus the one-month US Treasury bill rate, $SMB_{r,t-1}$ is the equal-weight average of the returns on the three small stock portfolios minus the average of the returns on the three big stock portfolios for region r, $HML_{r,t-1}$ is the equal-weight average of the returns for the two high B/M portfolios minus the average of the returns for the two low B/M portfolio for region r and $WML_{r,t-1}$ is the equal-weight average of the returns for the two winner portfolios minus the average of the returns for the two loser portfolios for region r. Newey-West t-statistics in parentheses are for testing $H_0: \beta_{i,j} = 0$ against $H_\alpha: \beta_{i,j} < 0$. Pooled estimates impose the restrictions that $\beta_{i,j} = \bar{\beta}_j, \beta^{b}_{i,r} = \bar{\beta}^{b}_{r}, \ \beta^{s}_{i,r} = \bar{\beta}^{s}_{r}, \ \beta^{h}_{i,r} = \bar{\beta}^{h}_{r}$ and $\beta^{w}_{i,r} = \bar{\beta}^{w}_{r}$. Superscripts a, b, c indicate significance at the 1%, 5%, 10% levels, respectively. * indicates significance at the 10% level or better of the test $H_0: \beta_{i,j} = \beta^{b}_{i,r} = \beta^{s}_{i,r} = \beta^{h}_{i,r} = \beta^{w}_{i,r} = 0$. All p-values for the tests are estimated by the wild bootstrap procedure in Rapach, Strauss, and Zhou (2013).

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	j = container				j = drybulk				j = tanker			
r = Global/ i =	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)
Brazil	-0.643^{b}	1.15	0.003	22.80*	-0.544^{b}	1.33	-0.027	22.81*	-1.260^{a}	4.66	-0.614^{b}	23.82*
	(-1.851)		(0.235)		(-2.455)		(-0.147)		(-4.277)		(-3.369)	
Canada	-0.394^{b}	1.53	0.029	43.04*	-0.354^{b}	1.98	-0.019	43.04*	-0.933^{a}	9.03	-0.433^{b}	44.82*
	(-1.887)		(0.343)		(-2.788)		(-0.269)		(-3.570)		(-3.634)	
France	-0.480^{b}	1.65	-0.009	33.17*	-0.388^{c}	1.73	-0.001	33.17*	-0.827^{a}	5.15	-0.259	33.64*
	(-1.973)		(-0.021)		(-3.337)		(-0.007)		(-3.244)		(-2.122)	
Germany	-0.433^{c}	1.07	0.105	30.83*	-0.385^{c}	1.36	0.054	30.80*	-0.776^{a}	3.61	-0.154	30.90*
	(-1.609)		(0.622)		(-3.468)		(0.352)		(-2.524)		(-0.751)	
India	-0.726^{b}	1.84	-0.318	16.40*	-0.788^{b}	3.48	-0.461	17.17*	-1.183^{a}	5.14	-0.738^{b}	17.91*
	(-1.804)		(-1.347)		(-4.745)		(-3.333)		(-3.418)		(-4.117)	
Italy	-0.468	1.08	0.036	25.30*	-0.320	0.81	0.113	25.39*	-0.901^{a}	4.21	-0.306	25.74*
	(-1.817)		(0.299)		(-1.829)		(0.677)		(-3.076)		(-2.618)	
Japan	-0.635^{b}	2.16	-0.180	25.30*	-0.648^{b}	3.61	-0.259	25.39*	-0.876^{a}	4.33	-0.293	25.74*
	(-2.257)		(-0.874)		(-9.313)		(-2.027)		(-3.864)		(-2.210)	
P.R of China	-1.044^{a}	3.86	-0.871^{b}	7.13*	-0.995^{a}	5.62	-0.859^{b}	8.51*	-1.022^{a}	3.89	-0.778^{a}	6.68*
	(-2.672)		(-2.225)		(-4.778)		(-3.478)		(-2.954)		(-2.948)	
Russian Fed.	-1.131^{b}	1.46	-0.448	14.63*	-1.288^{a}	3.04	-0.685^{c}	15.22*	-2.642^{a}	8.39	-1.942^{a}	18.59*
	(-1.953)		(-1.649)		(-4.155)		(-2.872)		(-4.105)		(-4.245)	
United Kingdom	-0.357^{c}	1.31	0.019	33.96*	-0.338^{c}	1.89	-0.008	33.96*	-0.597^{b}	3.85	-0.112	34.08*
	(-1.722)		(0.241)		(-4.689)		(-0.101)		(-2.489)		(-1.315)	
United States	-0.437^{b}	1.98	-0.036	37.10*	-0.394^{c}	2.59	-0.054	37.13*	-0.785^{a}	6.72	-0.290^{c}	37.93*
	(-1.704)		(-0.267)		(-4.634)		(-0.774)		(-2.767)		(-3.476)	
Pooled	-0.613^{a}	1.57	-0.152	18.50*	-0.586^{b}	2.25	-0.201	18.65*	-1.073^{a}	4.84	-0.538^{a}	19.51*
	(-2.406)		(-0.710)		(-2.317)		(-0.911)		(-3.686)		(-2.987)	

Table 4. Out-of-Sample Predictive Ability of Shipping Sentiment

Panel A of the table reports the Campbell and Thompson (2008) out-of-sample R^2 statistic (R^2_{OS}) in columns 2, 4 and 6. The statistic measures the mean-squared prediction error (MSPE) reduction when the forecasts estimated by the unrestricted model given by Eq. 7 are compared to the historical average forecast estimated by the restricted model. Panel B reports the R^2_{OS} when the Campbell and Thompson (2008) truncation (CTT) approach is implemented: $\hat{R}_{i,t+1} = max(0, \hat{\beta}_{i,0} + \hat{\beta}_{i,j} SS^{\perp}_{j,t})$, where $\hat{\beta}_{i,0}$ and $\hat{\beta}_{i,j}$ are the estimates of $\beta_{i,0}$ and $\beta_{i,j}$ of Eq. 7. Bold figures highlight cases where the $R^2_{OS,\mathrm{CTT}}$ is higher than its no constraint R^2_{OS} counterpart. Columns 3, 5 and 7 report the $R^2_{OS,P}$ statistics for the pooled version of Eq. 7 that imposes the restriction $\beta_{i,j} = \bar{\beta}_j$. The out-of-sample forecasts are based on recursive estimation windows. * indicates significance at the 10% level or better of the test H_0: $R^2_{OS} = 0$ against H_α: $R^2_{OS} > 0$ according to the Clark and West (2007) MSPE-adjusted statistic. "Average" is the average R^2_{OS} ($R^2_{OS,P}$) statistic.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: No constraint	j = container		j = drybulk		j = tanker	
i =	R^2_{OS} (%)	$R^2_{OS,P}$ (%)	R^2_{OS} (%)	$R^2_{OS,P}$ (%)	R^2_{OS} (%)	$R^2_{OS,P}$ (%)
Brazil	-0.47	1.39*	0.10	2.19*	4.09*	5.30*
Canada	-1.50	-1.24	0.67	1.12*	7.02*	8.13*
France	0.19	-0.84	-0.26	-0.37	4.79*	3.27*
Germany	-0.07	-0.87	-0.78	-0.11	3.35*	1.57*
India	-1.73	1.01	0.63	3.03*	3.64*	6.02*
Italy	-0.98	-1.09	-3.72	-1.72	4.50*	4.65*
Japan	-1.76	0.34	2.26*	2.76*	2.21*	2.81*
P.R of China	3.32*	2.34*	4.12*	5.71*	4.45*	3.53*
Russian Fed.	0.90*	2.38*	4.36*	4.25*	3.31*	9.66*
United Kingdom	0.02	-0.68	0.05	-1.13	1.46*	-3.12*
United States	0.91	0.22*	0.85	0.61	5.08*	4.63*
Average	-0.11	0.27	0.75	1.49	3.99	4.22
Panel B: CTT approach						
Brazil	**1.00***	**1.74***	**0.17**	0.88	**4.14***	4.38*
Canada	**0.69***	**0.29***	0.29	-0.69	2.69*	1.63*
France	-0.11	**-0.68**	-1.62	-2.07	3.32*	2.17*
Germany	**0.48**	**0.07**	-1.47	-1.43	2.95*	**2.81***
India	**1.22***	**2.67***	**2.08***	2.11	**5.34***	5.89*
Italy	-2.10	-1.94	-4.50	-3.22	1.45	1.36
Japan	**1.49***	**1.91***	1.29*	0.98	**2.37***	2.17*
P.R of China	-2.36*	1.55*	-0.45	3.53*	1.93*	2.21*
Russian Fed.	**2.31***	**2.40***	1.99*	1.89*	1.46*	5.06*
United Kingdom	**1.48***	**1.55***	-0.49	-2.01	**4.76***	**3.70***
United States	**1.62***	**1.26***	-0.42	-1.89	4.51*	3.26*
Average	**0.52**	**0.98**	-0.28	-0.17	3.17	3.15

Table 5. Mean-Variance Asset Allocation Results

The table reports the portfolio performance measures for a mean-variance investor with a risk aversion coefficient (γ) of 1, 3 and 5 respectively. The investor allocates capital between stock indices and risk-free Treasury bills on a monthly basis using the out-of-sample forecasts of the excess stock market returns based on lagged shipping investor sentiment. $SS^{\perp}_{container}$, $SS^{\perp}_{drybulk}$ and $SS^{\perp}_{tanker}$ are the container, drybulk and tanker sentiment indices respectively. CER gain is the annualized certainty equivalent return gain for the investor and the Sharpe Ratio is the mean excess return of the portfolio divided by its standard deviation. The portfolio weights are estimated recursively using data available at the forecast formation time t. Transaction costs are set at 50 bps. * attached to the CER gain and Sharpe ratio indicate significance at the 10% level or better based on the stationary bootstrap p-values. The out-of-sample evaluation period is April 2004 to April 2014.

	$SS^{\perp}_{container}$		$SS^{\perp}_{drybulk}$		$SS^{\perp}_{tanker}$	
Risk Aversion $\gamma = 1$	CER gain (%)	Sharpe ratio	CER gain (%)	Sharpe ratio	CER gain (%)	Sharpe ratio
Brazil	0.95	0.29	0.29	0.04	3.94*	0.68*
Canada	-0.56	-0.18	1.20	0.46	4.08*	0.69*
France	0.66	0.16	-0.16	-0.11	2.83*	0.55*
Germany	0.59	0.32	-0.16	-0.02	2.40*	0.65*
India	-1.62	-0.34	2.47*	0.33	4.75*	0.63*
Italy	-0.82	-0.22	-2.30	-0.38	2.04	0.29
Japan	-0.34	-0.08	0.47	0.13	1.68*	0.30*
P.R of China	3.08	-0.28	4.60*	0.47*	3.95*	0.57*
Russian Federation	2.94	0.37	3.94*	0.44*	9.82*	0.59*
United Kingdom	0.27	0.17*	-0.28	-0.11	1.68*	0.58*
United States	1.14*	0.47	0.28	0.08	3.27*	0.69*
Risk Aversion $\gamma = 3$						
Brazil	0.42	0.29	0.20	0.04	1.49*	0.68*
Canada	-0.28	-0.18	0.51	0.46	1.57*	0.69*
France	0.28	0.16	-0.03	-0.11	1.07*	0.55*
Germany	0.26	0.32	-0.02	-0.02	0.96*	0.65*
India	-0.69	-0.34	0.90*	0.33	1.78*	0.63*
Italy	-0.32	-0.22	-0.85	-0.38	0.81	0.29
Japan	-0.19	-0.08	0.23	0.13	0.66*	0.30*
P.R of China	1.33	0.28	1.54*	0.47*	1.23*	0.57*
Russian Federation	1.19	0.37	1.29	0.44*	3.79*	0.60*
United Kingdom	0.13	0.17*	-0.09	-0.11	0.68*	0.58*
United States	0.47*	0.47	0.14	0.08	1.31*	0.69*
Risk Aversion $\gamma = 5$						
Brazil	0.31	0.29	0.18	0.04	1.00*	0.68*
Canada	-0.23	-0.18	0.37	0.46	1.07*	0.69*
France	0.20	0.15	0.00	-0.11	0.72*	0.55*
Germany	0.19	0.32	0.00	-0.02	1.07*	0.65*
India	-0.51	-0.34	0.59*	0.33	1.18*	0.63*
Italy	-0.22	-0.22	-0.56	-0.38	0.57	0.29
Japan	-0.16	-0.08	0.18	0.13	0.45*	0.30*
P.R of China	0.98	0.28	0.92*	0.47*	0.68*	0.57*
Russian Federation	0.84	0.37	0.76*	0.44*	2.53*	0.60*
United Kingdom	0.10	0.17*	-0.06	-0.11	0.48*	0.58*
United States	0.34*	0.47	0.11	0.08	0.91*	0.68*

Table 6. Shipping Sentiment and Alternative Sentiment Indices

The table reports OLS estimates of $\beta_{US,j}$ and β_{US}^{k} (denoted by $\hat{\beta}_{US,j}$ and $\hat{\beta}_{US}^{k}$) for the regression models: $R_{US,t} = \beta_{US,0} + \beta_{US,j}SS_{j,t-1}^{\perp} + \beta_{US}^{k}S_{t-1}^{k} + u_{US,t}$ $k = BW, PLS$, (Eq. 12) and $R_{US,t} = \beta_{US,0} + \beta_{US,j}SS_{j,t-1}^{\perp} + \beta_{US}^{k}S_{t-1}^{k} + \beta_{US}^{b}RMRF_{US,t-1} + \beta_{US}^{s}SMB_{US,t-1} + \beta_{US}^{h}HML_{US,t-1} + \beta_{US}^{w}WML_{US,t-1} + u_{US,t}$ (Eq.13). $\hat{\beta}_{US,j}$ ($\hat{\beta}_{US}^{k}$) coefficients of Eq. 12 and Eq. 13 are presented in columns 2 (3) and 5 (6) respectively. $R_{US,t}$ is the excess return on the US stock market index, S_{t-1}^{BW} the lagged orthogonalized Baker and Wurgler (2006) investor sentiment index and S_{t-1}^{PLS} the lagged orthogonalized Huang *et al*. (2015) aligned investor sentiment index. The control variables $RMRF_{US,t-1}$, $SMB_{US,t-1}$ and $HML_{US,t-1}$ are the Fama and French (1993) factors and $WML_{US,t-1}$ is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios. Newey-West *t*-statistics in parentheses are for testing $H_0: \beta_{US,j} = 0$ against $H_\alpha: \beta_{US,j} < 0$. Superscripts a, b, c indicate significance at the 1%, 5%, 10% levels, respectively. * indicates significance at the 10% level or better of the test $H_0: \beta_{US,j} = \beta_{US}^{k} = \beta_{US}^{b} = \beta_{US}^{s} = \beta_{US}^{h} = \beta_{US}^{w} = 0$. All *p*-values for the tests are estimated following Rapach, Strauss, and Zhou (2013) wild bootstrap.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
$k = BW$ / $j =$	$\hat{\beta}_{US,j}$ (%)	$\hat{\beta}_{US}^{k}$	R^2 (%)	$\hat{\beta}_{US,j}$ (%)	$\hat{\beta}_{US}^{k}$	R^2 (%)
Container	-0.287	-1.107^{b}	3.84*	-0.059	-0.663^{c}	31.44*
	(-1.182)	(-3.911)		(-0.552)	(-2.162)	
Drybulk	-0.343	-1.170^{b}	4.79*	-0.102	-0.639^{c}	32.19*
	(-2.368)	(-3.413)		(-2.001)	(-1.610)	
Tanker	-0.706^{b}	-0.920^{b}	8.05*	-0.346^{c}	-0.504	33.20*
	(-4.583)	(-2.268)		(-3.027)	(-1.300)	
$k = PLS$ / $j =$						
Container	-0.357	-0.860^{b}	4.56*	-0.099	-0.537^{c}	31.72*
	(-1.455)	(-2.470)		(-0.753)	(-8.155)	
Drybulk	-0.359^{c}	-0.881^{b}	5.16*	-0.107	-0.508^{c}	32.38*
	(-2.390)	(-2.368)		(-1.547)	(-4.496)	
Tanker	-0.701^{b}	-0.661^{b}	8.11*	-0.337^{c}	-0.401^{c}	33.31*
	(-3.149)	(-3.297)		(-2.826)	(-4.521)	

Table 7. Estimation Results for Baltic Dry Index (BDI) and West Texas Intermediate (WTI) Oil Price

The table reports OLS estimates of β_i^{BDI} and β_i^{WTI} (denoted by $\hat{\beta}_i^{BDI}$ and $\hat{\beta}_i^{WTI}$) for the regression models: $R_{i,t} = \beta_{i,0} + \beta_i^k R_{t-1}^k + u_{i,t}$, $k = BDI, WTI$, (Eq. 14). $R_{i,t}$ is the stock index monthly excess return for country i, R_{t-1}^{BDI} is the lagged monthly return on the Baltic Dry Index (BDI) and R_{t-1}^{WTI} is the lagged monthly price change of West Texas Intermediate (WTI) oil price. $\hat{\beta}_i^{BDI}$ and $\hat{\beta}_i^{WTI}$ coefficients are presented in columns 2 and 4 respectively. Newey-West *t*-statistics in parentheses are for testing H_0: $\beta_i^{BDI} = 0$ against H_α: $\beta_i^{BDI} > 0$; and H_0: $\beta_i^{WTI} = 0$ against H_α: $\beta_i^{WTI} < 0$. Pooled estimates impose the restrictions that $\beta_i^{BDI} = \bar{\beta}^{BDI}$ and $\beta_i^{WTI} = \bar{\beta}^{WTI}$. Superscripts a, b, c indicate significance at the 1%, 5%, 10% levels, respectively. All *p*-values for the tests are estimated by the wild bootstrap procedure in Rapach, Strauss, and Zhou (2013).

(1)	(2)	(3)	(4)	(5)
$i =$	$\hat{\beta}_i^{BDI}$ (%)	R^2 (%)	$\hat{\beta}_i^{WTI}$ (%)	R^2 (%)
Brazil	2.021	0.22	-8.124	0.77
	(1.059)		(-1.188)	
Canada	3.508[a]	2.32	0.490	0.01
	(3.624)		(0.083)	
France	3.215[b]	1.41	-1.814	0.10
	(2.823)		(-0.299)	
Germany	2.772[c]	0.84	-2.360	0.13
	-2.545)		(-0.286)	
India	4.082[c]	1.11	0.690	0.01
	(2.593)		(0.067)	
Italy	4.803[a]	2.18	-2.239	0.10
	(2.788)		(-0.291)	
Japan	2.642[c]	0.71	3.980	0.36
	(2.279)		(0.703)	
P.R of China	-3.592	0.87	1.782	0.05
	(-1.648)		(0.441)	
Russian Fed.	5.862[c]	0.75	2.603	0.03
	(2.566)		(0.214)	
United Kingdom	2.464[c]	1.19	-3.101	0.42
	(2.378)		(-0.628)	
United States	2.402[b]	1.14	-0.763	0.03
	(2.052)		(-0.121)	
Pooled	2.744[c]	0.66	-0.805	0.10
	(1.602)		(-0.182)	

Table 8. Estimation Results Excluding Financial Crisis

The table reports OLS estimates of $\beta_{i,j}$, β_i^{BDI} and β_i^{WTI} (denoted by $\hat{\beta}_{i,j}$, $\hat{\beta}_i^{BDI}$ and $\hat{\beta}_i^{WTI}$) for the regression models: $R_{i,t} = \beta_{i,0} + \beta_{i,j} SS_{j,t-1}^{\perp} + u_{i,t}$ (Eq. 5) and $R_{i,t} = \beta_{i,0} + \beta_i^k R_{t-1}^k + u_{i,t},\ \ k = BDI, WTI,$ (Eq. 14) . $R_{i,t}$ is the stock index monthly excess return for country i, $SS_{j,t-1}^{\perp}$ is the lagged shipping sentiment for market j, R_{t-1}^{BDI} is the lagged monthly return on the Baltic Dry Index (BDI) and R_{t-1}^{WTI} is the lagged monthly price change of West Texas Intermediate (WTI) oil price. $\hat{\beta}_{i,j}$ coefficients are presented in columns 2, 4 and 6; $\hat{\beta}_i^{BDI}$ and $\hat{\beta}_i^{WTI}$ coefficients are presented in columns 6 and 8 respectively. Newey-West *t*-statistics in parentheses are for testing H_0: $\beta_{i,j} = 0$ against H_α: $\beta_{i,j} < 0$; H_0: $\beta_i^{BDI} = 0$ against H_α: $\beta_i^{BDI} > 0$; and H_0: $\beta_i^{WTI} = 0$ against H_α: $\beta_i^{WTI} < 0$. Pooled estimates impose the restrictions that $\beta_{i,j} = \bar{\beta}_j, \beta_i^{BDI} = \bar{\beta}^{BDI}$ and $\beta_i^{WTI} = \bar{\beta}^{WTI}$. Superscripts a, b, c indicate significance at the 1%, 5%, 10% levels, respectively. All *p*-values for the tests are estimated by the wild bootstrap procedure in Rapach, Strauss, and Zhou (2013). The exclusion period is December 2007 to June 2009.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(10)	(11)	(12)	(13)
	j = container		j = drybulk		j = tanker					
i =	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_i^{BDI}$ (%)	R^2 (%)	$\hat{\beta}_i^{WTI}$ (%)	R^2 (%)
Brazil	-0.385	0.31	-0.325	0.17	-1.065[a]	2.32	3.992	0.63	-14.967[b]	2.32
	(-0.981)		(-0.763)		(-2.843)		(1.429)		(-3.103)	
Canada	-0.135	0.16	-0.048	0.02	-0.531[a]	2.49	3.457[b]	2.05	-5.624[c]	1.42
	(-0.494)		(-0.188)		(-2.508)		(2.586)		(-1.822)	
France	-0.335	0.67	-0.242	0.27	-0.659[a]	2.55	2.963	1.00	-7.658[b]	1.74
	(-1.052)		(-0.793)		(-2.482)		(1.817)		(-1.918)	
Germany	-0.224	0.22	-0.109	0.04	-0.485[c]	1.03	2.430	0.50	-9.358[b]	1.95
	(-0.729)		(-0.392)		(-1.665)		(1.548)		(-1.965)	
India	-0.310	0.29	-0.691[c]	1.12	-0.778[a]	1.81	2.917	0.49	-8.733[c]	1.15
	(-1.030)		(-1.975)		(-2.554)		(1.280)		(-1.463)	
Italy	-0.277	0.32	-0.032	0.00	-0.608[b]	1.52	-0.772	0.04	-10.047[b]	2.11
	(-0.870)		(-0.914)		(-1.883)		(-0.431)		(-2.101)	
Japan	-0.497[c]	1.07	-0.769[b]	1.98	-0.570[c]	1.39	0.852	0.06	-2.883	0.18
	(-1.792)		(-2.146)		(-2.135)		(0.453)		(-0.543)	
P.R of China	-0.290	0.26	-0.180	0.08	-0.528[c]	0.85	-0.884	0.05	0.794	0.01
	(-0.676)		(-0.406)		(-1.812)		(-0.344)		(0.144)	
Russian Fed.	-0.810[c]	0.56	-1.581[b]	1.67	-2.185[a]	4.06	5.112	0.43	-13.408	0.77
	(-1.199)		(-2.216)		(-2.999)		(1.332)		(-1.642)	
United Kingdom	-0.230	0.49	-0.353[c]	0.90	-0.433[b]	1.72	2.725[b]	1.32	-6.867[b]	2.19
	(-0.883)		(-1.324)		(-1.686)		(2.269)		(-1.876)	
United States	-0.245	0.61	-0.223	0.39	-0.422[b]	1.79	1.587	0.49	-7.534[b]	2.89
	(-1.098)		(-0.910)		(-1.816)		(1.334)		(-2.084)	
Pooled	-0.340[c]	0.48	-0.414[c]	0.54	-0.751[a]	1.88	2.615	0.53	-7.844[b]	1.09
	(-2.336)		(-2.482)		(-3.327)		(1.540)		(-2.021)	

Figure 1. **Market and Sector Sentiment Indices 1996-2014**

The figure depicts the three market sentiment indices (container, drybulk and tanker) and their respective sector sentiment indices.

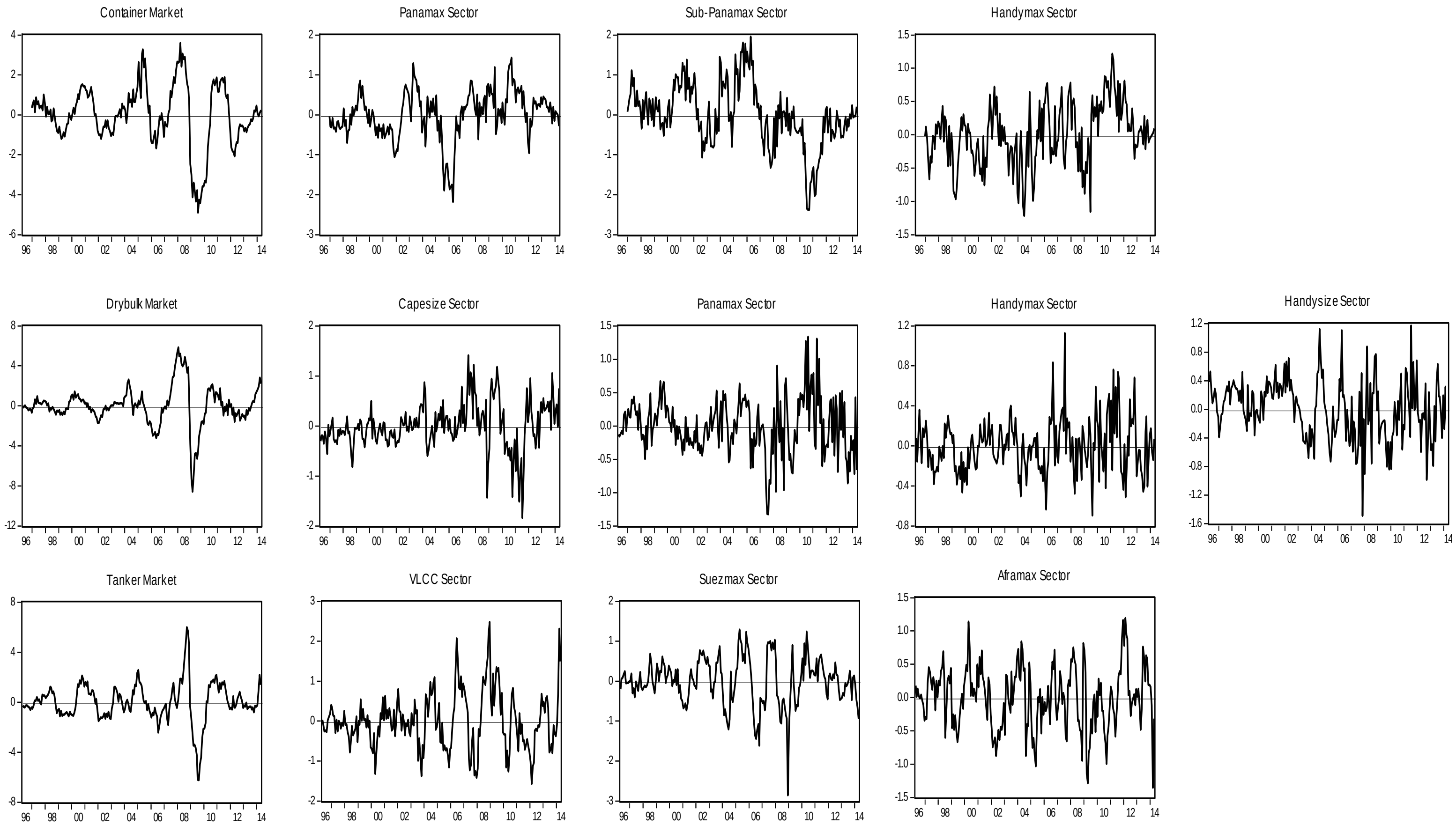


Figure 2. **Out-of-sample forecasts based on lagged shipping sentiment vs. historical average forecasts**

The figure depicts the cumulative differences of squared forecast errors estimated by the unrestricted models given by Eqs. (16)-(18) relative to the historical average forecast errors. The out-of-sample forecasts are based on recursive estimation windows. Vertical shaded areas represent country specific business cycle recessions according to the Economic Cycle Research Institute (www.businesscycle.com).

Graph A. Excess return forecasts for stock indices based on lagged container sentiment index model vs. historical average forecast model

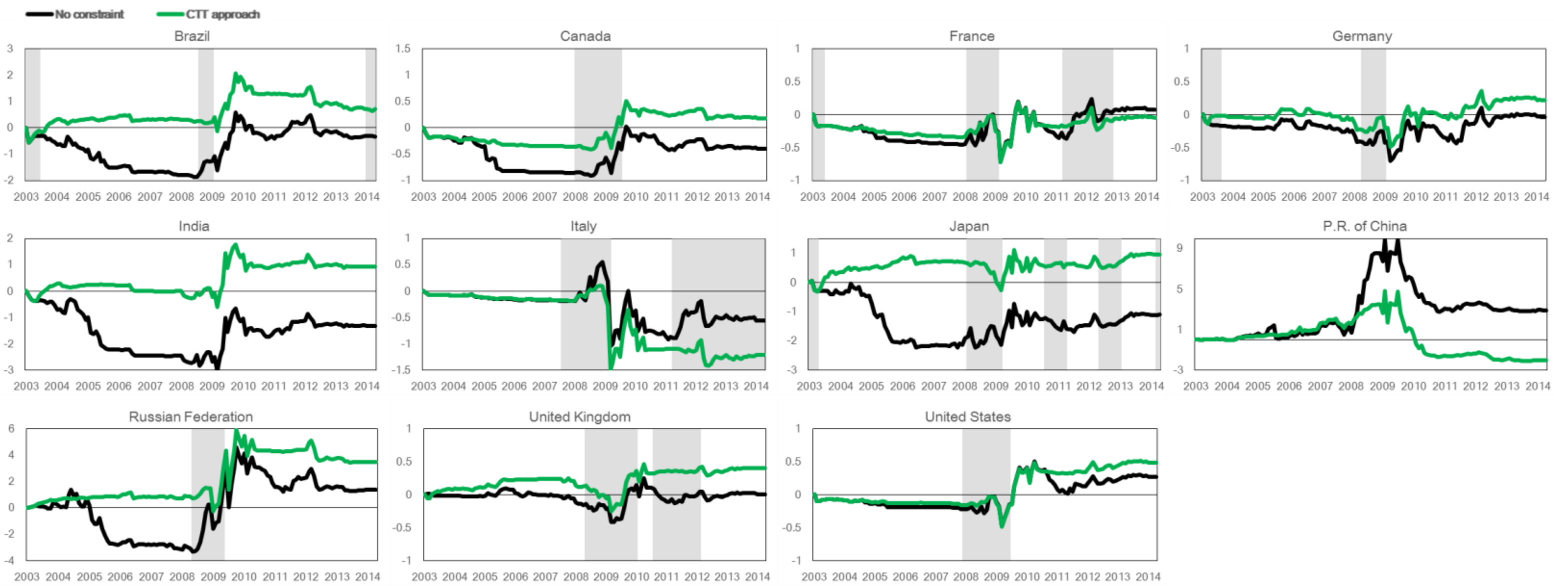


Figure 2. (cont'd)

Graph B. Excess return forecasts for stock indices based on lagged drybulk sentiment index model vs. historical average forecast model

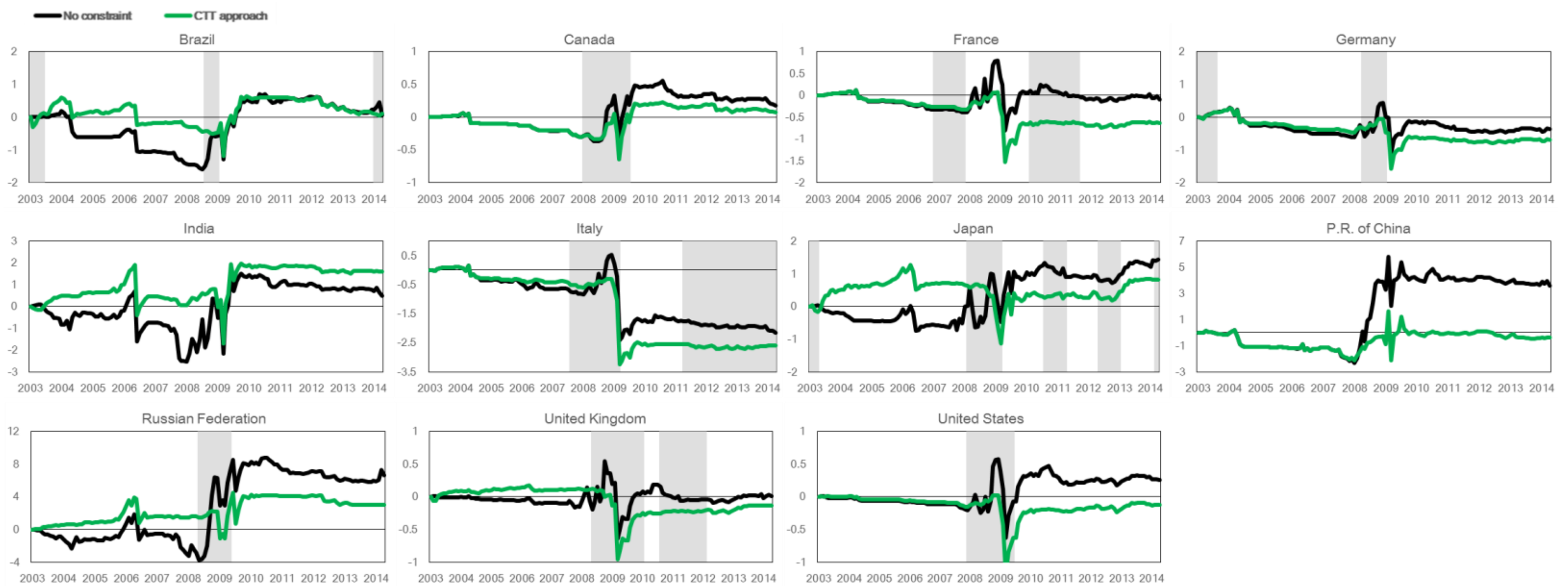
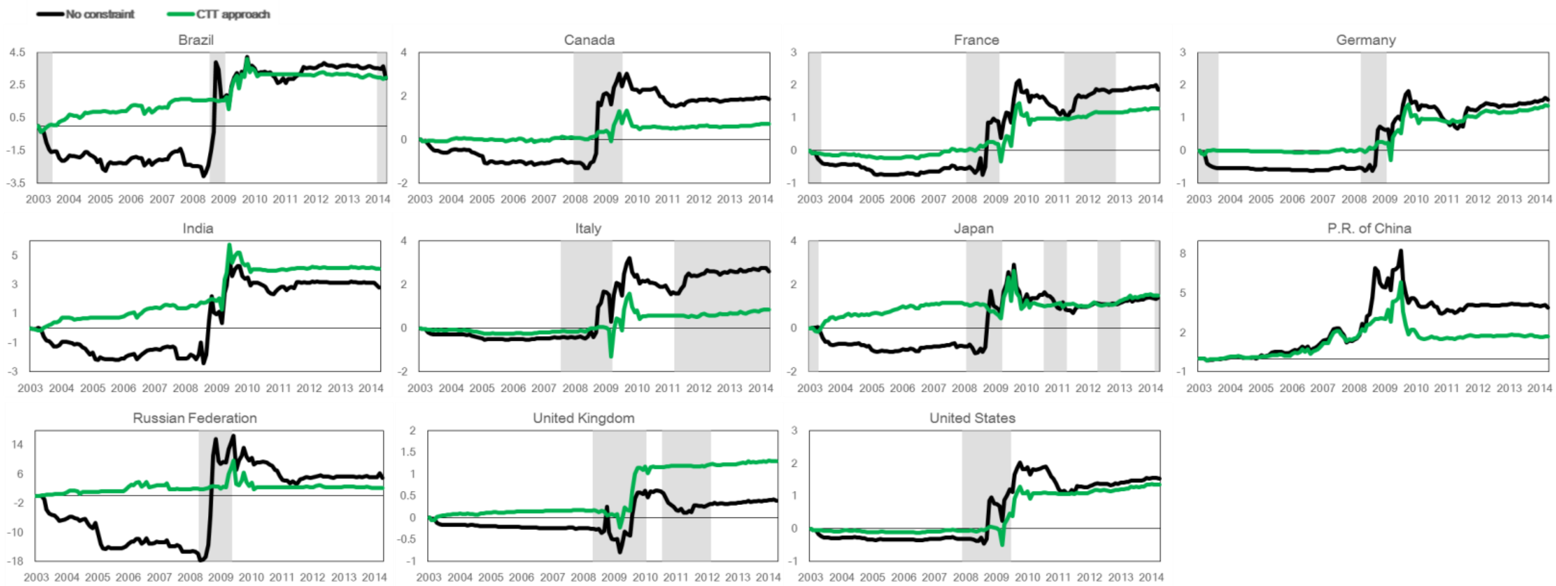


Figure 2. **(cont'd)**

Graph C. Excess return forecasts for stock indices based on lagged tanker sentiment index model vs. historical average forecast model



Supplementary Appendix

Shipping Investor Sentiment and International Stock Return Predictability

Abstract: This appendix contains unreported results discussed in the paper "Shipping Investor Sentiment and International Stock Return Predictability". The paper includes references to the corresponding sections and tables in this appendix. Appendix A provides a summary for the different types of seaborne transportation cargoes and vessel type characteristics. Table A.1 outlines the major vessel types, sizes and corresponding cargoes transported. Appendix B describes the sentiment proxies employed for the construction of the shipping sentiment indices. Table A.2 reports the correlation coefficients between the sentiment proxies and total sentiment indices, and Table A.3 reports the correlation of total, market and sector sentiment indices. Appendix C and Table A.4 reports the OLS estimates $\beta_{i,j}$ (denoted by $\hat{\beta}_{i,j}$) of Eq. 5 and Eq. 6 when the stock market excess returns are categorized into regions and the relevant Fama and French (2012) regional factors are employed as control variables. Appendix D and Table A.5 reports the results of testing the unique predictive power of shipping sentiment when 14 economic predictors of stock returns are included in the model. Table A.6, of Appendix D, evaluates the out-of-sample performance of the Baker and Wurgler (2006) and Huang *et al*. (2015) sentiment indices, and the 14 economic predictors of Welch and Goyal (2008) in terms of prediction errors based on the Campbell and Thompson (2008) out-of-sample R^2 statistic (denoted by R^2_{OS}) and the Clark and West (2007) adjusted mean-squared prediction error (MSPE-adjusted) statistic. Appendix E and Table A.7 evaluates the out-of-sample performance of the Baltic Dry Index (BDI) and West Texas Intermediate (WTI) oil prices. To end, Appendix F presents the figures of the R^2_{OS} and $R^2_{OS,\text{CTT}}$ statistics and the corresponding *p*-values for various forecast evaluation periods running from April 2003 through April 2014, with the end of the out-of-sample period fixed at April 2014.

Appendix A: Types of seaborne transportation cargoes and vessel type characteristics

Seaborne trade encompasses the transportation of many different commodities. For example, raw materials such as oil, iron ore and coal; agricultural products such as grains, sugar and wheat; industrial materials such as cement, chemicals and rubber; and manufactured products such as cars, consumer products and heavy machinery. The shipping transportation model can be split into three main categories: bulk parcels, specialized parcels and general cargo parcels, depending on the parcel size distribution (PSD) function of the commodity and the service requirements of the cargo parcel; where a parcel is an individual quantity of cargo for transportation. Bulk cargoes can be split into two main categories: 1) liquid bulk (such as oil products and liquid chemicals) that requires tanker transportation; and, 2) drybulk (such as iron ore, grain and steel products) that requires conventional drybulk transportation. Additionally, specialized cargoes (such as motor vehicles and refrigerated food) require transportation by vessels designed precisely for a specialized cargo type. Finally, general cargoes (such as manufactured products of all types) require mainly containership transportation. This paper is focused, on the liquid, drybulk and general cargoes seaborne transportation. In particular, container vessels that their carrying capacity is measured in TEUs (Twenty-foot Equivalent Unit – the international standard measure for containers and containership capacity); and, drybulk and tanker vessels that their carrying capacity is measured in deadweight ton (dwt – a unit of carrying capacity including cargo, fuel, oil, water, stores and crew and is measured in in metric tons of 1,000 kilograms).

Table A.1. Vessel Types, Sizes and Seaborne Transportation Cargoes

Vessel type	Carrying Capacity	Transported Cargoes
Container market		
Panamax	3,000 TEU >	Manufactured products of all types
Sub-panamax	2,999 – 2,000 TEU	Manufactured products of all types
Handymax	1,999 – 1,000 TEU	Manufactured products of all types
Drybulk market		
Capesize	100,000 dwt >	Iron ore and coal
Panamax	99,999 – 60,000 dwt	Coal; grains and minor bulks
Handymax	59,999 – 40,000 dwt	Grains and minor bulks
Handysize	39,999 – 10,000 dwt	Minor Bulks
Tanker market		
VLCC (Very Large Crude Carrier)	200,000 dwt >	Crude oil
Suezmax	199,999 – 120,000 dwt	Crude oil and clean petroleum products
Aframax	119,999 – 80,000 dwt	Crude oil and clean petroleum products

Appendix B: Shipping Sentiment Indices

Studies by Baker and Wurgler (2006) and Baker, Wurgler, and Yuan (2012) use market price-based proxies such as closed-end fund discounts, IPOs' volume and their first-day returns, volume turnover, equity share of new issues, dividend premium and volatility premium. Other studies employ micro-trading data such as trading positions of large speculators, large hedgers and small traders in the US futures markets (Wang, 2001) or broker data and transaction data (Kumar and Lee, 2006; Barber, Odean, and Zhu, 2009). Investor surveys (Lee, Jiang, and Indro, 2002; Brown and Cliff, 2004; Menkhoff and Rebitzky, 2008) and consumer confidence indices (Lemmon and Portniaguina, 2006; Schmeling, 2009) are also employed as proxies for sentiment. Finally, investor sentiment is also linked to close-end fund discounts (Hwang, 2011, Neal and Wheatley, 1998; Swaminathan, 1996) and mutual funds (Ben-Rephael, Kandel, and Wohl, 2012). What follows is a description of the proxies employed for the construction of the shipping sentiment indices and unreported results discussed in the paper.

The first proxy employed is net contracting (NC) which measures the number of orders for newbuilding vessels contracted with shipyards after accounting for order cancellations and scrapped vessels. The use of net contracting proxy is motivated as follows. First, we assume that the demand for new vessels in the shipping market resembles the demand for new equity issues in the financial markets. This is by analogy to Baker and Wurgler (2006, 2007) and Baker, Wurgler, and Yuan (2012) who use the number of IPOs as a sentiment proxy; as the demand for IPOs is considered to be extremely sensitive to investor sentiment. Second, we assume that shipping participants tend to follow the herd and invest in new capacity when the orderbook and valuations are generally high. Usually, high shipping earnings are associated with high second-hand vessel prices and orderbook, but forecast low future returns (Greenwood and Hanson, 2015). Furthermore, Greenwood and Hanson (2015) argue that over-investment in new capacity during booms is due to shipowners being overconfident and incorrectly believing that investments will continue to reap high returns[1]. Therefore, we assume that high-sentiment periods are characterized by high vessel orders with cancellations and scrapping of vessels being at low levels. The monthly net contracting is given by,

$$NC_{j,q,t} = \left(order_{j,q,t} - order_{j,q,t-1} + del_{j,q,t}\right) - scrap_{j,q,t}, \qquad (A.1)$$

where $order_{j,q,t}$ is the orderbook, i.e., the number of vessels awaiting construction or being constructed in market j and sector q during month t, $del_{j,q,t}$ the number of vessel deliveries and $scrap_{j,q,t}$ the number of vessels being scrapped. By construction our proxy takes into account order cancellations which reflect investment sentiment and conditions in shipping markets, thus, it measures net investment in new capacity.

The second proxy is the money committed (MC) in each market and sector,

[1] Greenwood and Hanson (2015) attribute this behaviour partly to "competition neglect" by shipowners, which is caused by the time lag involved in the shipbuilding process (Kahneman, 2011).

$$MC_{j,q,t} = order_{j,q,t} \times newPx_{j,q,t}, \tag{A.2}$$

where $newPx_{j,q,t}$ is the price of newbuilding vessels in market *j* and sector *q* during month *t*. This proxy reflects the funds committed for the purchase of newbuilding vessels. Broader measures of financing activity are used in the literature as sentiment proxies. For example, Baker and Wurgler (2000) suggest that the share of equity issues in the total of equity and debt issues is a measure of financing activity that can capture sentiment. The main source of capital for shipping projects is bank finance, with a historical average debt-to-equity ratio of seventy-to-thirty for a straightforward shipping project. Furthermore, expansion phases in the freight rate market (i.e., high sentiment) have been historically fuelled by the liberal availability of debt finance for newbuildings, as the providers of credit are overconfident in financing shipping projects. At the same time, there is a corresponding willingness of investors to become excessively geared. This stems from the fact that we are dealing with a niche market where providers of credit follow their competitors to avoid losing any market share. Therefore, we expect MC to be positively related to investor sentiment.

The third sentiment proxy is the price-to-earnings ratio (PE) for vessels,

$$PE_{j,q,t} = schPx_{j,q,t}/earn_{j,q,t}, \tag{A.3}$$

where $schPx_{j,q,t}$ is the price of five-year old second-hand vessels and $earn_{j,q,t}$ the annualized earnings (one-year time-charter rates[2]) in market j and sector q during month t. The price-to-earnings ratio as a measure of sentiment has been previously considered in the literature and has been identified as a predictor of stock returns (Campbell and Shiller, 1998; Fisher and Statman, 2006; Kurov, 2008). Generally, high PE ratios reflect the relative degree of overvaluation in asset prices and high investor sentiment. In the shipping industry, the estimate of earnings is forward-looking and reflects the expected earnings to be received from operating a vessel for one year from the point of valuation. As such, when current vessel prices are high relative to the one-year forward-looking earnings (i.e., high PE ratio), investors expect vessel prices to drop in the future in anticipation of limited earnings growth. We thus expect high PE ratios to be associated with low sentiment levels.

The fourth proxy we consider is the second-hand to newbuilding price ratio (SNB),

$$SNB_{j,q,t} = schPx_{j,q,t}/newPx_{j,q,t}. \tag{A.4}$$

Newbuilding vessels have longer useful economic lives than identical second-hand vessels of certain age (e.g., five or ten-year old vessels), which in general implies higher capital outlays. However, during prosperous freight market conditions and high sentiment periods, investors prefer to take advantage of the prevailing market conditions immediately. As a result, they favour the purchase of second-hand vessels to avoid the time lag in the construction process of newbuildings[3].

[2] Fixed daily freight rate, measured in US$/day, received by the shipowner for chartering (leasing or letting-out) a vessel for a one-year period.

[3] The building of new vessels is characterized by significant construction lags. The actual construction time, which is on average two years, may often be lengthened considerably by the lack of available berth capacity in shipyards or due to order backlog. For example, Kalouptsidi (2014) quantifies the impact of time-to-build on shipping investments and estimates that the average construction time almost doubled between 2001 and 2008.

This preference consequently creates an immediate delivery premium that may occasionally drive second-hand vessel prices above newbuilding vessel prices. The selection of SNB as a sentiment proxy is by analogy (inverse) to Baker and Wurgler (2004) use of dividend premium. Dividends are generally perceived by investors as a characteristic for safety (Baker and Wurgler, 2006). When dividends are at premium, companies are more likely to pay them and less so when they are at discount (Fama and French, 2001). Therefore, companies appear to cater prevailing sentiment for or against "safety" in their decision of dividend distributions. Similarly, SNB reflects the preference of market agents for second-hand vessels to newbuilding vessels and captures the immediate delivery premium, which is related to the level of optimism or pessimism regarding the current market conditions.

Our last proxy reflects the relative liquidity in the shipping industry. The use of liquidity as a sentiment proxy follows from Baker and Stein (2004) who suggest turnover as a candidate proxy for investor sentiment. They argue that, under short-sales constraints, irrational investors are more likely to participate in the market and add liquidity when they are optimistic. Short-sales constraints are even more profound in the shipping industry as it is difficult and costly for participants to establish short positions on vessels. Baker and Wurgler (2006, 2007) capture market liquidity by the ratio of trading volume to the number of shares listed on the New York Stock Exchange, whereas Baker, Wurgler, and Yuan (2012) use the ratio of total dollar volume over a year to the total market capitalization at the end of the previous year. However, liquidity is an elusive notion (Amihud, 2002; Pastor and Stambaugh, 2003) which has been represented by various empirical measures in the literature[4]. We represent shipping market liquidity by the

[4] Proxies for liquidity, among others, include: (i) turnover (Amihud and Mandelson, 1986), (ii) dollar volume (Chordia, Subrahmanyam, and Anshuman, 2001), (iii) share volume (Brennan and Subrahmanyam, 1995), (iv) Roll-implicit spread estimator (Roll, 1984), (v) illiquidity ratio (Amihud, 2002) and (vi) proportion of zero returns measure (Lesmond, Ogden, and Trzcinka, 1999).

turnover ratio[5] (TURN) which measures the activity in the sale and purchase market for second-hand vessels in terms of the total number of vessels available in the market,

$$TURN_{j,q,t} = M^{-1} \sum_{p=t-M+1}^{t} Sale_{j,q,p}/Fleet_{j,q,p}, \tag{A.5}$$

where $Fleet_{j,q,p}$ is the total number of available vessels in sector q of market j and month p, $Sale_{j,q,p}$ the number of vessels sold and $M = 12$ months up to time t. We anticipate high turnover periods to be related to high sentiment.

For each market and sector we construct a first-stage index comprising fifteen loadings given by the current, one-month lagged and two-month lagged orthogonalized proxies (denoted by $\perp$). This way, a lead-lag relationship between the proxies is allowed since a number of proxies may reflect a shift in sentiment earlier than others (Brown and Cliff, 2004; Baker and Wurgler, 2006). To decide which proxies shall be included in the total shipping sentiment (SS) index, we estimate the correlation between the first-stage index and the current and lagged proxies. The proxies with the highest correlation qualify as the final sentiment proxies and the first principal component of the selected proxies provides the total sentiment indices.

Further, the high correlation[6] between the first-stage indices and total sentiment indices suggests that dropping the remaining ten proxies does not lead to any significant loss of information. Table A.2. also indicates that the sentiment proxies – across all markets and sectors – are highly correlated with the total sentiment index, but the correlation between the proxies is

[5] Our choice of liquidity measure is driven by data availability on a monthly basis.

[6] Correlation between the first-stage indices and $SS_{j,q,t}^{total^{\perp}}$ lies within 84% - 97% in the container market, 91% - 97% in the drybulk market and 75% - 97% in the tanker market.

relatively low. This implies that the proxies contain idiosyncratic information in reflecting investor sentiment and the risk of using variables that reveal the same information is low. Finally, the high correlation coefficients between the total sentiment indices reported in Table A.3. suggest that sentiment may flow from one sector to another. We circumvent this issue and separate the overall market sentiment from the sector-specific sentiment as discussed in the paper.

Table A.2. Correlation of Index Components and Total Sentiment Indices

$SS_{j,q,t}^{total^{\perp}}$ is the first principal component of the five orthogonalized sentiment proxies in market *j* and sector *q* during month *t*. NC is the net contracting (Eq. A.1), MC the money committed (Eq. A.2), PE the price-to-earnings ratio (Eq. A.3), SNB the second-hand-to-newbuilding price ratio (Eq. A.4) and TURN the turnover ratio (Eq. A.5). The orthogonalized proxies labelled by $\perp$ are the residuals from the regression of each of the five raw sentiment proxies on the G7 industrial production growth and two-recession period dummies for the G7 and Major 5 Asia countries. Superscripts a, b, c indicate significance at the 1%, 5%, 10% levels, respectively.

	Correlations with $SS_{j,q,t}^{total^{\perp}}$	Correlations between proxies				
j = Container q = Panamax	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t-1}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t-1}^{\perp}$
$NC_{t-2}^{\perp}$	0.51[a]	1.00				
$MC_{t}^{\perp}$	0.57[a]	0.23[a]	1.00			
$PE_{t-1}^{\perp}$	-0.65[a]	-0.15[b]	-0.14[b]	1.00		
$SNB_{t-2}^{\perp}$	0.85[a]	0.16[b]	0.41[a]	-0.50[a]	1.00	
$TURN_{t-1}^{\perp}$	0.50[a]	0.13[b]	0.14[b]	-0.27[a]	0.17[b]	1.00
j = Container q = Sub-panamax	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t-1}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t-1}^{\perp}$
$NC_{t-2}^{\perp}$	0.56[a]	1.00				
$MC_{t}^{\perp}$	0.75[a]	0.28[a]	1.00			
$PE_{t-1}^{\perp}$	-0.68[a]	-0.12[c]	-0.39[a]	1.00		
$SNB_{t-2}^{\perp}$	0.58[a]	0.27[a]	0.40[a]	-0.11[c]	1.00	
$TURN_{t-1}^{\perp}$	0.45[a]	0.13[c]	0.12[c]	-0.19[a]	0.42[a]	1.00
j = Container q = Handymax	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t-1}^{\perp}$
$NC_{t-2}^{\perp}$	0.53[a]	1.00				
$MC_{t}^{\perp}$	0.61[a]	0.33[a]	1.00			
$PE_{t}^{\perp}$	-0.49[a]	-0.15[b]	-0.11[c]	1.00		
$SNB_{t-2}^{\perp}$	0.88[a]	0.33[a]	0.42[a]	-0.36[a]	1.00	
$TURN_{t-1}^{\perp}$	0.43[a]	0.12[c]	0.11[c]	-0.22[a]	0.41[a]	1.00
j = Drybulk q = Capesize	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t-1}^{\perp}$	$SNB_{t-1}^{\perp}$	$TURN_{t-1}^{\perp}$
$NC_{t-2}^{\perp}$	0.69[a]	1.00				
$MC_{t}^{\perp}$	0.41[a]	0.37[a]	1.00			
$PE_{t-1}^{\perp}$	-0.59[a]	-0.18[a]	-0.12[c]	1.00		
$SNB_{t-1}^{\perp}$	0.84[a]	0.39[a]	0.36[a]	-0.36[a]	1.00	
$TURN_{t-1}^{\perp}$	0.51[a]	0.17[b]	0.27[a]	-0.31[a]	0.36[a]	1.00
j = Drybulk q = Panamax	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t-1}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t-1}^{\perp}$
$NC_{t-2}^{\perp}$	0.64[a]	1.00				
$MC_{t}^{\perp}$	0.64[a]	0.35[a]	1.00			
$PE_{t-1}^{\perp}$	-0.66[a]	-0.29[a]	-0.11[c]	1.00		
$SNB_{t-2}^{\perp}$	0.90[a]	0.44[a]	0.64[a]	-0.51[a]	1.00	
$TURN_{t-1}^{\perp}$	0.64[a]	0.36[a]	0.11[c]	-0.43[a]	0.56[a]	1.00
j = Drybulk q = Handymax	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t-1}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t-2}^{\perp}$
$NC_{t-2}^{\perp}$	0.57[a]	1.00				
$MC_{t}^{\perp}$	0.60[a]	0.38[a]	1.00			
$PE_{t-1}^{\perp}$	-0.64[a]	-0.16[b]	-0.11[c]	1.00		
$SNB_{t-2}^{\perp}$	0.90[a]	0.34[a]	0.60[a]	-0.50[a]	1.00	
$TURN_{t-2}^{\perp}$	0.65[a]	0.19[a]	0.12[c]	-0.44[a]	0.51[a]	1.00
j = Drybulk q = Handysize	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t-1}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t-2}^{\perp}$
$NC_{t-2}^{\perp}$	0.58[a]	1.00				
$MC_{t}^{\perp}$	0.62[a]	0.36[a]	1.00			
$PE_{t-1}^{\perp}$	-0.57[a]	-0.21[a]	-0.12[c]	1.00		
$SNB_{t-2}^{\perp}$	0.84[a]	0.37[a]	0.62[a]	-0.32[a]	1.00	
$TURN_{t-2}^{\perp}$	0.68[a]	0.21[a]	0.11[c]	-0.47[a]	0.44[a]	1.00

Table A.2. (continued)

	Correlations with $SS_{j,q,t}^{total^{\perp}}$	Correlations between proxies				
j = Tanker q = VLCC	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t-2}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t-1}^{\perp}$
$NC_{t-2}^{\perp}$	0.50[a]	1.00				
$MC_{t}^{\perp}$	0.76[a]	0.20[a]	1.00			
$PE_{t-2}^{\perp}$	-0.53[a]	-0.14[b]	-0.34[a]	1.00		
$SNB_{t-2}^{\perp}$	0.80[a]	0.35[a]	0.45[a]	-0.25[a]	1.00	
$TURN_{t-1}^{\perp}$	0.61[a]	0.13[b]	0.34[a]	-0.12[c]	0.41[a]	1.00
j = Tanker q = Suezmax	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t-2}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t}^{\perp}$
$NC_{t-2}^{\perp}$	0.48[a]	1.00				
$MC_{t}^{\perp}$	0.55[a]	0.12[c]	1.00			
$PE_{t-2}^{\perp}$	-0.40[a]	-0.13[b]	-0.11[c]	1.00		
$SNB_{t-2}^{\perp}$	0.83[a]	0.11[c]	0.29[a]	-0.19[a]	1.00	
$TURN_{t}^{\perp}$	0.60[a]	0.11[c]	0.17[b]	-0.16[b]	0.30[a]	1.00
j = Tanker q = Aframax	$SS_{j,q,t}^{total^{\perp}}$	$NC_{t-2}^{\perp}$	$MC_{t}^{\perp}$	$PE_{t}^{\perp}$	$SNB_{t-2}^{\perp}$	$TURN_{t-1}^{\perp}$
$NC_{t-2}^{\perp}$	0.48[a]	1.00				
$MC_{t}^{\perp}$	0.62[a]	0.12[c]	1.00			
$PE_{t}^{\perp}$	-0.68[a]	-0.14[b]	-0.20[a]	1.00		
$SNB_{t-2}^{\perp}$	0.63[a]	0.13[c]	0.23[a]	-0.20[a]	1.00	
$TURN_{t-1}^{\perp}$	0.81[a]	0.12[c]	0.37[a]	-0.43[a]	0.34[a]	1.00

Table A.3. Correlation of Total, Market and Sector Sentiment Indices

$SS_{j,t}^{market^{\perp}}$ is the first principal component of all $SS_{j,q,t}^{total^{\perp}}$ in each market *j*. $SS_{j,q,t}^{sector^{\perp}}$ are the residuals from regressing $SS_{j,q,t}^{total^{\perp}}$ on $SS_{j,t}^{market^{\perp}}$ for each sector *q* in market *j*. Superscripts a, b, c indicate significance at the 1%, 5%, 10% levels, respectively.

Panel A: market and total sentiment indices

	Correlations with $SS_{j,t}^{market^{\perp}}$	Correlations between $SS_{j,q,t}^{total^{\perp}}$			
j = Container	$SS_{j,t}^{market^{\perp}}$	$SS_{j,pana,t}^{total^{\perp}}$	$SS_{j,subpana,t}^{total^{\perp}}$	$SS_{j,hmax,t}^{total^{\perp}}$	
$SS_{j,pana,t}^{total^{\perp}}$	0.89[a]	1.00			
$SS_{j,subpana,t}^{total^{\perp}}$	0.82[a]	0.54[a]	1.00		
$SS_{j,hmax,t}^{total^{\perp}}$	0.95[a]	0.83[a]	0.69[a]	1.00	
j = Drybulk	$SS_{j,t}^{market^{\perp}}$	$SS_{j,cape,t}^{total^{\perp}}$	$SS_{j,pana,t}^{total^{\perp}}$	$SS_{j,hmax,t}^{total^{\perp}}$	$SS_{j,hsize,t}^{total^{\perp}}$
$SS_{j,cape,t}^{total^{\perp}}$	0.94[a]	1.00			
$SS_{j,pana,t}^{total^{\perp}}$	0.96[a]	0.86[a]	1.00		
$SS_{j,pana,t}^{total^{\perp}}$	0.98[a]	0.91[a]	0.94[a]	1.00	
$SS_{j,hsize,t}^{total^{\perp}}$	0.96[a]	0.87[a]	0.91[a]	0.94[a]	1.00
j = Tanker	$SS_{j,t}^{market^{\perp}}$	$SS_{j,vlcc,t}^{total^{\perp}}$	$SS_{j,suez,t}^{total^{\perp}}$	$SS_{j,afra,t}^{total^{\perp}}$	
$SS_{j,vlcc,t}^{total^{\perp}}$	0.88[a]	1.00			
$SS_{j,suez,t}^{total^{\perp}}$	0.88[a]	0.61[a]	1.00		
$SS_{j,afra,t}^{total^{\perp}}$	0.94[a]	0.76[a]	0.77[a]	1.00	

Panel B: market, total and sector sentiment indices

	Correlations with $SS_{j,t}^{market^{\perp}}$ or $SS_{j,q,t}^{total^{\perp}}$		Correlations between $SS_{j,q,t}^{sector^{\perp}}$			
j = Container	$SS_{j,t}^{market^{\perp}}$	$SS_{j,q,t}^{total^{\perp}}$	$SS_{j,pana,t}^{sector^{\perp}}$	$SS_{j,subpana,t}^{sector^{\perp}}$	$SS_{j,hmax,t}^{sector^{\perp}}$	
$SS_{j,pana,t}^{sector^{\perp}}$	0.00	0.45[a]	1.00			
$SS_{j,subpana,t}^{sector^{\perp}}$	0.00	0.35[a]	-0.68[a]	1.00		
$SS_{j,hmax,t}^{sector^{\perp}}$	0.00	0.15[b]	-0.12[b]	-0.52[a]	1.00	
j = Drybulk	$SS_{j,t}^{market^{\perp}}$	$SS_{j,q,t}^{total^{\perp}}$	$SS_{j,cape,t}^{sector^{\perp}}$	$SS_{j,pana,t}^{sector^{\perp}}$	$SS_{j,hmax,t}^{sector^{\perp}}$	$SS_{j,hsize,t}^{sector^{\perp}}$
$SS_{j,cape,t}^{sector^{\perp}}$	0.00	0.33[a]	1.00			
$SS_{j,pana,t}^{sector^{\perp}}$	0.00	0.27[a]	-0.52[a]	1.00		
$SS_{j,hmax,t}^{sector^{\perp}}$	0.00	0.18[a]	-0.34[a]	-0.11[b]	1.00	
$SS_{j,hsize,t}^{sector^{\perp}}$	0.00	0.26[a]	-0.46[a]	-0.32[a]	-0.16[b]	1.00
j = Tanker	$SS_{j,t}^{market^{\perp}}$	$SS_{j,q,t}^{total^{\perp}}$	$SS_{j,vlcc,t}^{sector^{\perp}}$	$SS_{j,suez,t}^{sector^{\perp}}$	$SS_{j,afra,t}^{sector^{\perp}}$	
$SS_{j,vlcc,t}^{sector^{\perp}}$	0.00	0.48[a]	1.00			
$SS_{j,suez,t}^{sector^{\perp}}$	0.00	0.34[a]	-0.61[a]	1.00		
$SS_{j,afra,t}^{sector^{\perp}}$	0.00	0.19[a]	-0.39[a]	-0.37[a]	1.00	

Appendix C: Shipping Sentiment and Stock Market Returns

Table A.4. Regression Model Estimation Results for Stock Index Returns

The table reports OLS estimates of $\beta_{i,j}$ (denoted by $\hat{\beta}_{i,j}$) for the regression models: $R_{i,t} = \beta_{i,0} + \beta_{i,j}SS^{\perp}_{j,t-1} + u_{i,t}$ (Eq. 5) and $R_{i,t} = \beta_{i,0} + \beta_{i,j}SS^{\perp}_{j,t-1} + \beta^{b}_{i,r}RMRF_{r,t} + \beta^{s}_{i,r}SMB_{r,t} + \beta^{h}_{i,r}HML_{r,t} + \beta^{w}_{i,r}WML_{r,t} + u_{i,t}$ (Eq. 6). $\hat{\beta}_{i,j}$ coefficients of Eq. 5 (Eq. 6) are presented in columns 2, 6 and 10 (4, 8 and 12). $R_{i,t}$ is the stock index monthly excess return for country i and $SS^{\perp}_{j,t-1}$ is the lagged shipping sentiment for market j. The control variables are the Fama and French (2012) factors for developed markets constructed using six value-weight portfolios formed on size and book-to-market: $RMRF_{r,t}$ is region's r value-weighted return on the market portfolio minus the one-month US Treasury bill rate, $SMB_{r,t}$ is the equal-weight average of the returns on the three small stock portfolios minus the average of the returns on the three big stock portfolios for region r, $HML_{r,t}$ is the equal-weight average of the returns for the two high B/M portfolios minus the average of the returns for the two low B/M portfolio for region r and $WML_{r,t}$ is the equal-weight average of the returns for the two winner portfolios minus the average of the returns for the two loser portfolios for region r. Newey-West t-statistics in parentheses are for testing $H_0: \beta_{i,j} = 0$ against $H_\alpha: \beta_{i,j} < 0$. Pooled estimates impose the restrictions that $\beta_{i,j} = \bar{\beta}_j, \beta^{b}_{i,r} = \bar{\beta}^{b}_r,\ \beta^{s}_{i,r} = \bar{\beta}^{s}_r,\ \beta^{h}_{i,r} = \bar{\beta}^{h}_r$ and $\beta^{w}_{i,r} = \bar{\beta}^{w}_r$. Superscripts a, b, c indicate significance at the 1%, 5%, 10% levels, respectively. * indicates significance at the 10% level or better of the test $H_0: \beta_{i,j} = \beta^{b}_{i,r} = \beta^{s}_{i,r} = \beta^{h}_{i,r} = \beta^{w}_{i,r} = 0$. All p-values for the tests are estimated by the wild bootstrap procedure in Rapach, Strauss, and Zhou (2013).

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	j = container				j = drybulk				j = tanker			
r = Asia Pacific/ i =	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)	$\hat{\beta}_{i,j}$ (%)	R^2 (%)
India	-0.726^{b}	1.84	-0.065	16.43*	-0.788^{b}	3.48	-0.254	16.74*	-1.183^{a}	5.14	-0.473^{c}	17.13*
	(-1.804)		(-0.307)		(-4.745)		(-3.311)		(-3.418)		(-2.883)	
P.R of China	-1.044^{a}	3.86	-0.840^{b}	8.17*	-0.995^{a}	5.62	-0.830^{b}	9.40*	-1.022^{a}	3.89	-0.757^{a}	7.79*
	(-2.672)		(-2.021)		(-4.778)		(-3.770)		(-2.954)		(-2.367)	
Pooled	-0.885^{a}	2.76	-0.452^{c}	10.23*	-0.891^{a}	4.48	-0.542^{c}	11.05*	-1.102^{a}	4.50	-0.615^{a}	10.80*
	(-2.703)		(-1.439)		(-2.719)		(-1.623)		(-4.028)		(-2.548)	
r = European/ i =												
France	-0.480^{b}	1.65	0.060	36.85*	-0.388^{c}	1.73	0.071	36.88*	-0.827^{a}	5.15	-0.140	36.96*
	(-1.973)		(0.343)		(-3.337)		(0.379)		(-3.244)		(-0.684)	
Germany	-0.433^{c}	1.07	0.179	33.72*	-0.385^{c}	1.36	0.122	33.68*	-0.776^{a}	3.61	-0.021	33.56*
	(-1.609)		(0.665)		(-3.468)		(0.661)		(-2.524)		(-0.074)	
Italy	-0.468	1.08	0.093	30.10*	-0.320	0.81	0.199	30.34*	-0.901^{a}	4.21	-0.165	30.18*
	(-1.817)		(0.486)		(-1.829)		(1.597)		(-3.076)		(-1.089)	
Russian Federation	-1.131^{b}	1.46	-0.421	17.19*	-1.288^{a}	3.04	-0.634^{c}	17.68*	-2.642^{a}	8.39	-1.741^{a}	20.26*
	(-1.953)		(-1.451)		(-4.155)		(-2.292)		(-4.105)		(-3.199)	
United Kingdom	-0.357^{c}	1.31	0.025	32.31*	-0.338^{c}	1.89	0.001	32.31*	-0.597^{b}	3.85	-0.082	32.37*
	(-1.722)		(0.154)		(-4.689)		(0.009)		(-2.489)		(-0.514)	
Pooled	-0.357^{c}	1.31	-0.013	20.38*	-0.338^{c}	1.89	-0.048	20.39*	-0.597^{b}	3.85	-0.430	20.96*
	(-2.703)		(-0.053)		(-2.719)		(-0.203)		(-4.028)		(-2.063)	
r = Japanese/ i =												
Japan	-0.635^{b}	2.16	-0.304	25.65*	-0.648^{b}	3.61	-0.403^{b}	26.51*	-0.876^{a}	4.33	-0.525^{b}	26.62*
	(-2.257)		(-2.198)		(-9.313)		(-4.005)		(-3.864)		(-4.481)	
r = North American/ i =												
Canada	-0.394^{b}	1.53	-0.149	33.39*	-0.354^{b}	1.98	-0.153	33.53*	-0.933^{a}	9.03	-0.577^{a}	36.41*
	(-1.887)		(-0.855)		(-2.788)		(-2.438)		(-3.570)		(-4.024)	
United States	-0.437^{b}	1.98	-0.144	32.75*	-0.394^{c}	2.59	-0.131	32.81*	-0.785^{a}	6.72	-0.377^{b}	34.00*
	(-1.704)		(-1.348)		(-4.634)		(-2.540)		(-2.767)		(-3.268)	
Pooled	-0.415^{c}	1.74	-0.147	31.67*	-0.374^{b}	2.27	-0.142	31.77*	-0.859^{a}	7.85	-0.477^{a}	33.73*
	(-2.703)		(-0.812)		(-2.719)		(-0.803)		(-4.028)		(-2.368)	

Appendix D: Shipping Sentiment, US Investor Sentiment and Economic Predictors of Stock Returns

Following Huang *et al*. (2015), we test whether the predictive power of shipping sentiment remains significant after controlling for 14 economic variables of Welch and Goyal (2008), which are documented in the literature as stock return predictors. To this end, we run regressions of the type:

$$R_{US,t} = \beta_{US,0} + \beta_{US,j} SS_{j,t-1}^{\perp} + \phi^k \Pi_{t-1}^k + u_{US,t}, \qquad k = 1, \ldots, 14, \tag{A.6}$$

where Π_{t-1}^k is one of the 14 economic predictors described below.

1. Dividend Price Ratio (**d/p**) is the difference between the log of a 12-month moving sum of dividends paid on the S&P 500 index and the log of prices (S&P 500).
2. Dividend Yield (**d/y**) is the difference between the log of dividends and the log of lagged prices.
3. Earnings Price Ratio (**e/p**) is the difference between the log of a 12-month moving sum of earnings on the S&P 500 and the log of prices.
4. Dividend Payout Ratio (**d/e**) is the difference between the log of dividends and the log of earnings.
5. Stock Variance (**svar**) is the sum of squared daily returns on the S&P 500.
6. Book to Market Ratio (**b/m**) is the ratio of book value to market value for the Dow Jones Industrial Average.
7. Net Equity Expansion (**ntis**) is the ratio of 12-month moving sums of net issues by NYSE listed stocks divided by the total market capitalization of NYSE stocks.
8. Treasury Bills (**tbl**) is the interest rate on the 3-month US T-bills.
9. Long Term Yield (**lty**) is the yield on the long-term US government bonds.

10. Long Term Rate of Return (**ltr**) is the return on the long-term US government bonds.
11. Term Spread (**tms**) is the difference between the long-term yield on government bonds and the T-bills rate.
12. Default Yield Spread (**dfy**) is the difference between BAA- and AAA- rated corporate bond yields.
13. Default Return Spread (**dfr**) is the difference between the return on long-term corporate bonds and the return on long-term government bonds.
14. Inflation (**infl**) is the one month lagged US Consumer Price Index.

Table A.5. Shipping Sentiment and Economic Predictors of Stock Returns

The table reports OLS estimates of $\beta_{US,j}$ (denoted by $\hat{\beta}_{US,j}$) for the regression model $R_{US,t} = \beta_{US,0} + \beta_{US,j}SS^{\perp}_{j,t-1} + \phi^k \Pi^k_{t-1} + u_{US,t}$, $k = 1, \ldots, 14$, (A.6). $\hat{\beta}_{US,j}$ coefficients of Eq. A.6 are presented in columns 2, 4 and 6 respectively. $R_{US,t}$ is the excess return on the US stock market index and Π^k_{t-1} is one of the 14 economic predictors: d/p, d/y, e/p, d/e, svar, b/m, ntis, tbl, lty, ltr, tms, dfy, dfr and infl. Newey-West *t*-statistics in parentheses are for testing H_0: $\beta_{US,j} = 0$ against H_α: $\beta_{US,j} < 0$. Superscripts a, b, c indicate significance at the 1%, 5%, 10% levels, respectively. * indicates significance at the 10% level or better of the test H_0: $\beta_{US,j} = \phi^k = 0$. All *p*-values for the tests are estimated following Rapach, Strauss, and Zhou (2013) wild bootstrap.

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	j = container		j = drybulk		j = tanker	
k =	$\hat{\beta}_{US,j}$ (%)	R^2 (%)	$\hat{\beta}_{US,j}$ (%)	R^2 (%)	$\hat{\beta}_{US,j}$ (%)	R^2 (%)
d/p	-0.450[c]	2.03	-0.409[c]	2.71	-0.789[a]	6.77*
	(-2.728)		(-2.728)		(-2.684)	
d/y	-0.341	2.85*	-0.346	3.60	-0.743[b]	7.59*
	(-2.353)		(-2.353)		(-2.481)	
e/p	-0.559[c]	2.53*	-0.498[b]	3.28	-0.969[a]	8.24*
	(-4.489)		(-4.489)		(-3.133)	
d/e	-0.586[c]	2.52	-0.503[b]	3.19	-0.941[b]	7.84*
	(-3.946)		(-3.946)		(-3.196)	
svar	-0.447[c]	4.06	-0.399	4.58	-0.727[b]	7.26*
	(-3.988)		(-3.988)		(-2.830)	
b/m	-0.443[c]	2.04	-0.411[c]	2.75	-0.794[a]	6.79*
	(-4.002)		(-4.002)		(-2.947)	
ntis	-0.407[b]	3.18	-0.371[c]	3.67	-0.747[a]	7.30*
	(-4.327)		(-4.327)		(-3.407)	
tbl	-0.440[c]	2.02	-0.403[c]	2.75	-0.787[b]	6.83*
	(-4.835)		(-4.835)		(-2.772)	
lty	-0.429[c]	2.08	-0.399[c]	2.80	-0.794[a]	7.04*
	(-4.343)		(-4.343)		(-2.944)	
ltr	-0.427[c]	2.18	-0.397[c]	2.85	-0.782[a]	6.81*
	(-4.436)		(-4.436)		(-2.624)	
tms	-0.456[c]	2.06	-0.406[c]	2.71	-0.792[a]	6.77*
	(-4.835)		(-4.835)		(-2.772)	
dfy	-0.567[c]	3.07*	-0.497[b]	3.77	-0.813[a]	7.27*
	(-3.467)		(-3.467)		(-3.147)	
dfr	-0.236	8.42*	-0.245	8.79*	-0.593[b]	11.38*
	(-4.556)		(-4.556)		(-3.118)	
infl	-0.430[c]	2.34*	-0.400[c]	3.04	-0.829[a]	7.75*
	(-3.695)		-(3.695)		(-3.005)	

Table A.6. Out-of-Sample Predictive Ability of Behavioral and Economic Variables

This table reports the Campbell and Thompson (2008) out-of-sample R^2 statistic (R^2_{OS}). The statistic measures the mean-squared prediction error (MSPE) reduction when the forecasts estimated by the unrestricted models $R_{i,t+1} = \beta_{i,0} + \beta_{i,j} SS^{\perp}_{j,t} + u_{i,t+1}$, $R_{US,t+1} = \beta_{US,0} + \beta^k_{US} S^k_t + u_{US,t+1}$ ($k = BW, PLS,$) and $R_{US,t+1} = \beta_{US,0} + \phi \Pi^k_t + u_{US,t+1}$ ($k = 1, \ldots, 14$) are compared to the historical average forecast estimated by the restricted model. The table also reports the R^2_{OS} when the Campbell and Thompson (2008) truncation (CTT) approach is implemented: $\hat{R}_{i,t+1} = max(0, \hat{\beta}_{i,0} + \hat{\beta}_{i,j} SS^{\perp}_{j,t})$, $\hat{R}_{i,t+1} = max(0, \hat{\beta}_{US,0} + \hat{\beta}^k_{US} R^k_t)$ and $\hat{R}_{i,t+1} = max(0, \hat{\beta}_{US,0} + \hat{\phi} \Pi^k_t)$, where $\hat{\beta}_{i,0}, \hat{\beta}_{i,j}, \hat{\beta}_{US,0}, \hat{\beta}^k_i$ and $\hat{\phi}$ are the estimates of $\beta_{i,0}, \beta_{i,j}, \beta_{US,0}, \beta^k_{US}$ and ϕ. Bold figures highlight cases where the $R^2_{OS,\text{CTT}}$ is higher than its no constraint R^2_{OS} counterpart. * indicates significance at the 10% level or better of the test H_0: $R^2_{OS} = 0$ against H_α: $R^2_{OS} > 0$ according to the Clark and West (2007) MSPE-adjusted statistic.

	R^2_{OS} (%) No constraint	R^2_{OS} (%) CTT approach
$SS^{\perp}_{container}$	0.91	**1.62***
$SS^{\perp}_{drybulk}$	0.85	-0.42
$SS^{\perp}_{tanker}$	5.08*	4.51*
S^{BW}	1.54*	1.24*
S^{PLS}	3.12*	1.56*
d/p	-3.63	**-2.54**
d/y	-8.53	**-8.23**
e/p	-4.44	**-0.80**
d/e	-5.44	**0.48**
svar	-1.30	**0.59***
b/m	-1.38	**-0.24**
ntis	-1.29	**-0.40**
tbl	-2.19	**0.06**
lty	-2.97	**0.27**
ltr	-1.36	**-0.35**
tms	-1.09	**-0.24**
dfy	-5.01	**0.11**
dfr	4.97*	1.47*
infl	-2.06	**-1.02**

Appendix E: Out-of Sample Testing for BDI and WTI

Table A.7. Out-of-Sample Predictive Ability of Shipping Sentiment

This table reports the Campbell and Thompson (2008) out-of-sample R^2 statistic (R^2_{OS}) in columns 2, 4, 6, 8 and 10. The statistic measures the mean-squared prediction error (MSPE) reduction when the forecasts estimated by the unrestricted models $R_{i,t+1} = \beta_{i,0} + \beta_{i,j} SS^{\perp}_{j,t} + u_{i,t+1}$ and $R_{i,t+1} = \beta_{i,0} + \beta^k_i R^k_t + u_{i,t+1}$ $(k = BDI, WTI)$ are compared to the historical average forecast estimated by the restricted model. Columns 3, 5, 7, 9 and 11 report the $R^2_{OS,P}$ statistics for the pooled version that imposes the restriction $\beta_{i,j} = \bar{\beta}_j$, $\beta^{BDI}_i = \bar{\beta}^{BDI}$ and $\beta^{WTI}_i = \bar{\beta}^{WTI}$. Bold figures highlight cases where the $R^2_{OS,\text{CTT}}$ is higher than its no constraint R^2_{OS} counterpart. The out-of-sample forecasts are based on recursive estimation windows. Values in parentheses report the Clark and West (2007) MSPE-adjusted statistic of the test H_0: $R^2_{OS} = 0$ against H_α: $R^2_{OS} > 0$. * indicates significance at the 10% level or better. "Average" is the average R^2_{OS} ($R^2_{OS,P}$) statistic.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(10)	(11)	(12)	(13)
Panel A: No constraint	j = container		j = drybulk		j = tanker		BDI		WTI	
i =	R^2_{OS} (%)	$R^2_{OS,P}$ (%)	R^2_{OS} (%)	$R^2_{OS,P}$ (%)	R^2_{OS} (%)	$R^2_{OS,P}$ (%)	R^2_{OS} (%)	$R^2_{OS,P}$ (%)	R^2_{OS} (%)	$R^2_{OS,P}$ (%)
Brazil	-0.47	1.39*	0.10	2.19*	4.09*	5.30*	-3.09	-0.96	-1.43	0.12
Canada	-1.50	-1.24	0.67	1.12*	7.02*	8.13*	1.94*	2.31*	-2.28	-3.71
France	0.19	-0.84	-0.26	-0.37	4.79*	3.27*	1.13*	0.86*	-3.81	-3.17
Germany	-0.07	-0.87	-0.78	-0.11	3.35*	1.57*	-0.29	-0.47	-5.30	-2.52
India	-1.73	1.01	0.63	3.03*	3.64*	6.02*	-0.24	0.33	-1.95	-2.07
Italy	-0.98	-1.09	-3.72	-1.72	4.50*	4.65*	2.73*	1.47*	-3.86	-1.78
Japan	-1.76	0.34	2.26*	2.76*	2.21*	2.81*	-0.71	-0.50	-0.55	-2.82
P.R of China	3.32*	2.34*	4.12*	5.71*	4.45*	3.53*	-2.13	-2.96	-0.74	-1.48
Russian Federation	0.90*	2.38*	4.36*	4.25*	3.31*	9.66*	-1.43	0.29	-3.23	-2.48
United Kingdom	0.02	-0.68	0.05	-1.13	1.46*	-3.12*	0.00	-0.18	-3.47	-3.68
United States	0.91	0.22*	0.85	0.61	5.08*	4.63*	-0.93	0.19	-3.95	-3.57
Average	-0.11	0.27	0.75	1.49	3.99	4.22	-0.27	0.03	-2.78	-2.47
Panel B: CTT approach										
Brazil	**1.00***	**1.74***	0.17	0.88	**4.14***	4.38*	**-0.02**	**0.55**	**-0.27**	**0.56**
Canada	**0.69***	**0.29***	0.29	-0.69	2.69*	1.63*	**1.96***	1.96*	**-1.87**	**-2.68**
France	-0.11	**-0.68**	-1.62	-2.07	3.32*	2.17*	0.98	0.69	**-2.72**	**-2.33**
Germany	**0.48**	**0.07**	-1.47	-1.43	2.95*	**2.81***	**1.42***	**0.84**	**-3.09**	**-1.58**
India	**1.22***	**2.67***	**2.08***	2.11	**5.34***	5.89*	**1.17***	**1.40***	**-0.20**	**-0.04**
Italy	-2.10	-1.94	-4.50	-3.22	1.45	1.36	1.34	0.58	**-2.38**	**-0.71**
Japan	**1.49***	**1.91***	1.29*	0.98	**2.37***	2.17*	**1.27***	**1.71***	**1.82***	**1.06**
P.R of China	-2.36*	1.55*	-0.45	3.53*	1.93*	2.21*	**-0.71**	**-1.13**	**-0.72**	**-1.23**
Russian Federation	**2.31***	**2.40***	1.99*	1.89*	1.46*	5.06*	**0.61**	**0.92***	**-1.23**	**-1.04**
United Kingdom	**1.48***	**1.55***	-0.49	-2.01	**4.76***	**3.70***	**1.04**	**1.92***	**-1.30**	**-1.33**
United States	**1.62***	**1.26***	-0.42	-1.89	4.51*	3.26*	**0.82**	**0.87**	**-2.90**	**-2.59**
Average	**0.52**	**0.98**	-0.28	-0.17	3.17	3.15	**0.90**	**0.94**	**-1.35**	**-1.08**

Figure A.1. Container Sentiment - R^2_{OS} No constraint

The figure reports the R^2_{OS} and the corresponding p-values when the estimation period is April 1996 to March 2003, and the beginning of the various forecast evaluation periods runs from April 2003 through April 2013. The end point of the out-of-sample period remains is always April 2014. Shaded areas represent the recent subprime and financial crises.

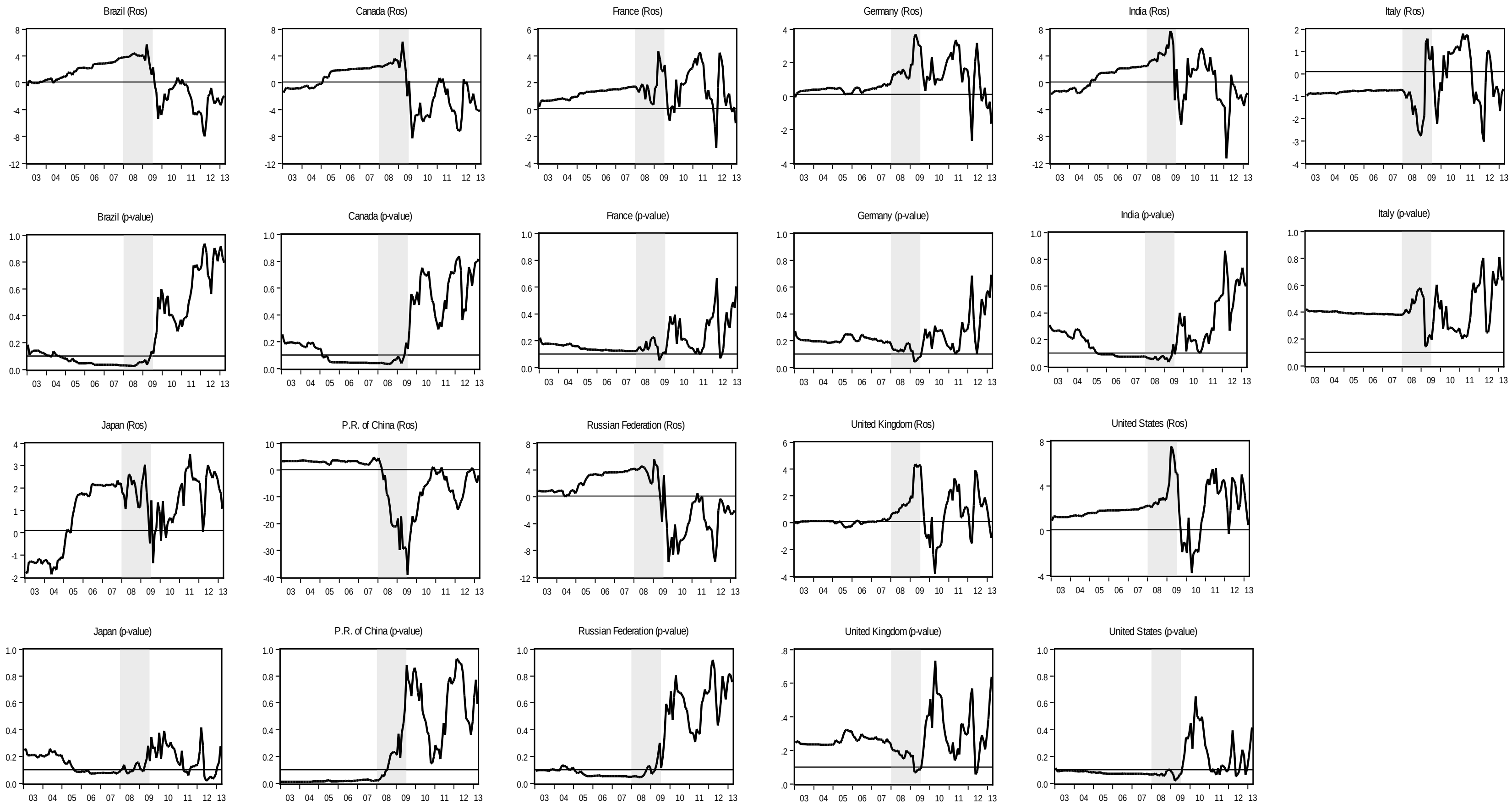


Figure A.2. Container Sentiment - $R^2_{OS,\mathrm{CTT}}$

The figure reports the $R^2_{OS,\mathrm{CTT}}$ and the corresponding p-values when the estimation period is April 1996 to March 2003, and the beginning of the various forecast evaluation periods runs from April 2003 through April 2013. The end point of the out-of-sample period remains is always April 2014. Shaded areas represent the recent subprime and financial crises.

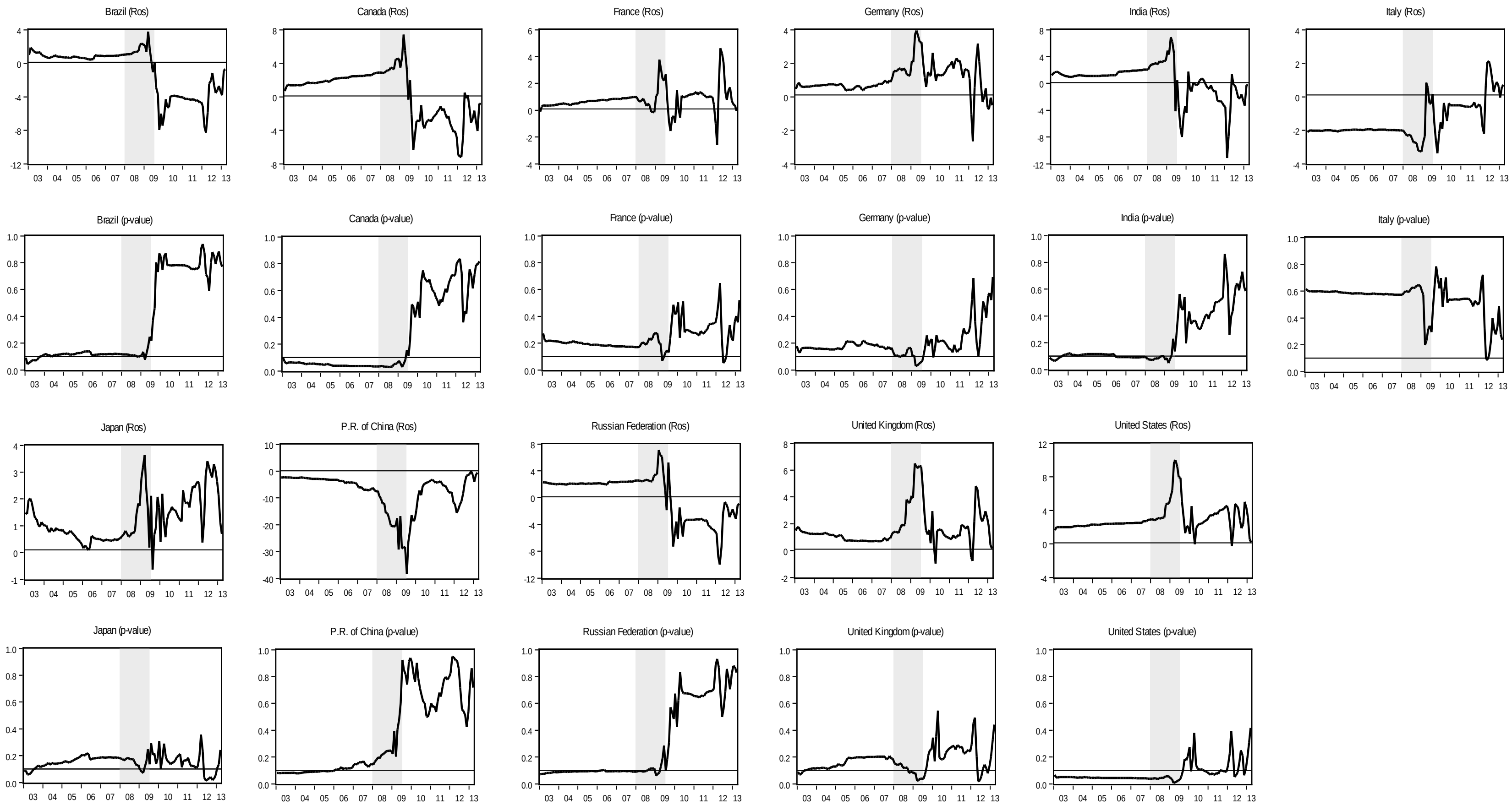


Figure A.3. Drybulk Sentiment - R^2_{OS} No constraint

The figure reports the R^2_{OS} and the corresponding *p*-values when the estimation period is April 1996 to March 2003, and the beginning of the various forecast evaluation periods runs from April 2003 through April 2013. The end point of the out-of-sample period remains is always April 2014. Shaded areas represent the recent subprime and financial crises.

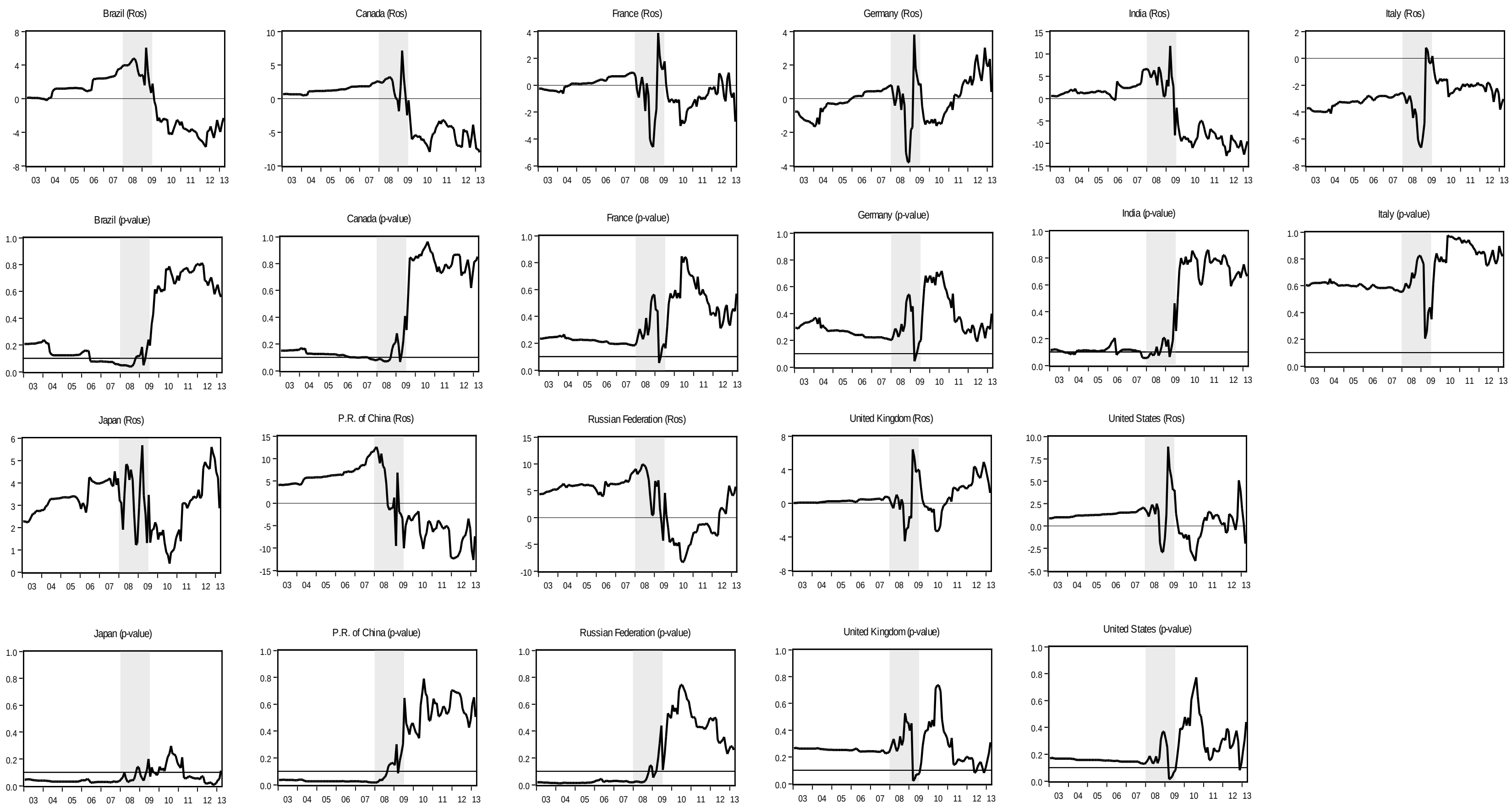


Figure A.4. Drybulk Sentiment - $R^2_{OS,\mathrm{CTT}}$

The figure reports the $R^2_{OS,\mathrm{CTT}}$ and the corresponding *p*-values when the estimation period is April 1996 to March 2003, and the beginning of the various forecast evaluation periods runs from April 2003 through April 2013. The end point of the out-of-sample period remains is always April 2014. Shaded areas represent the recent subprime and financial crises.

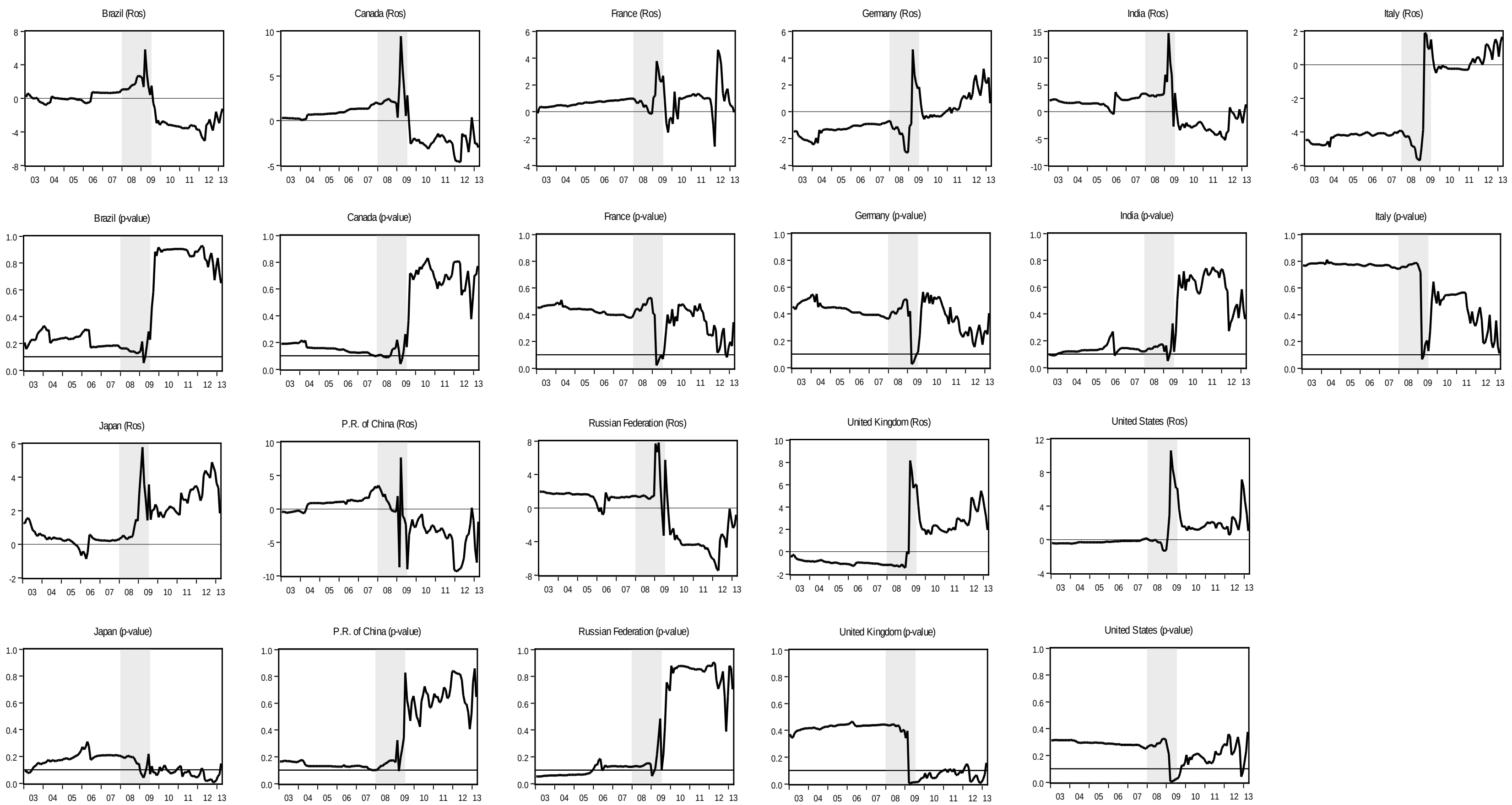


Figure A.5. Tanker Sentiment - R_{OS}^2 No constraint

The figure reports the R_{OS}^2 and the corresponding *p*-values when the estimation period is April 1996 to March 2003, and the beginning of the various forecast evaluation periods runs from April 2003 through April 2013. The end point of the out-of-sample period remains is always April 2014. Shaded areas represent the recent subprime and financial crises.

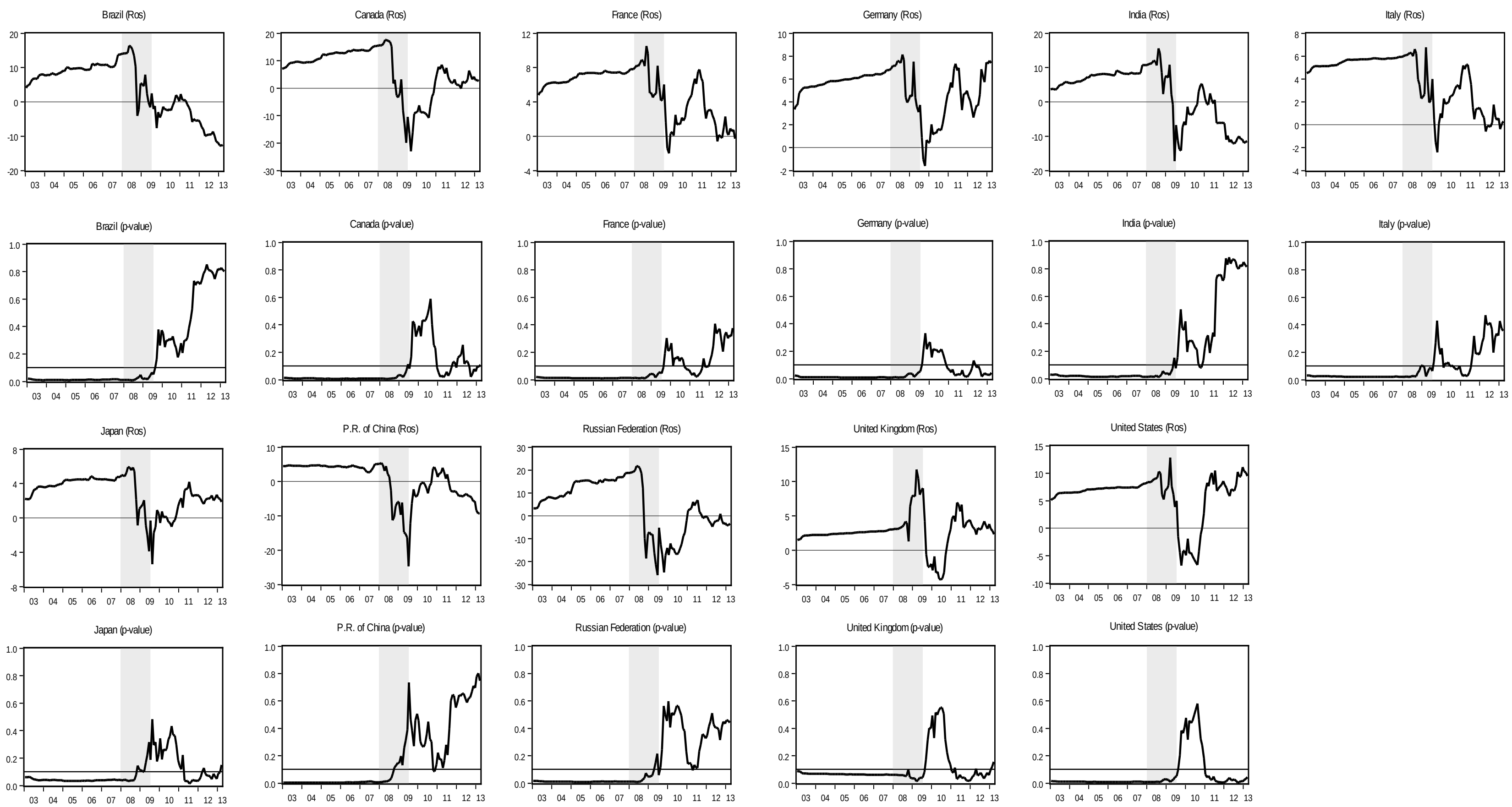
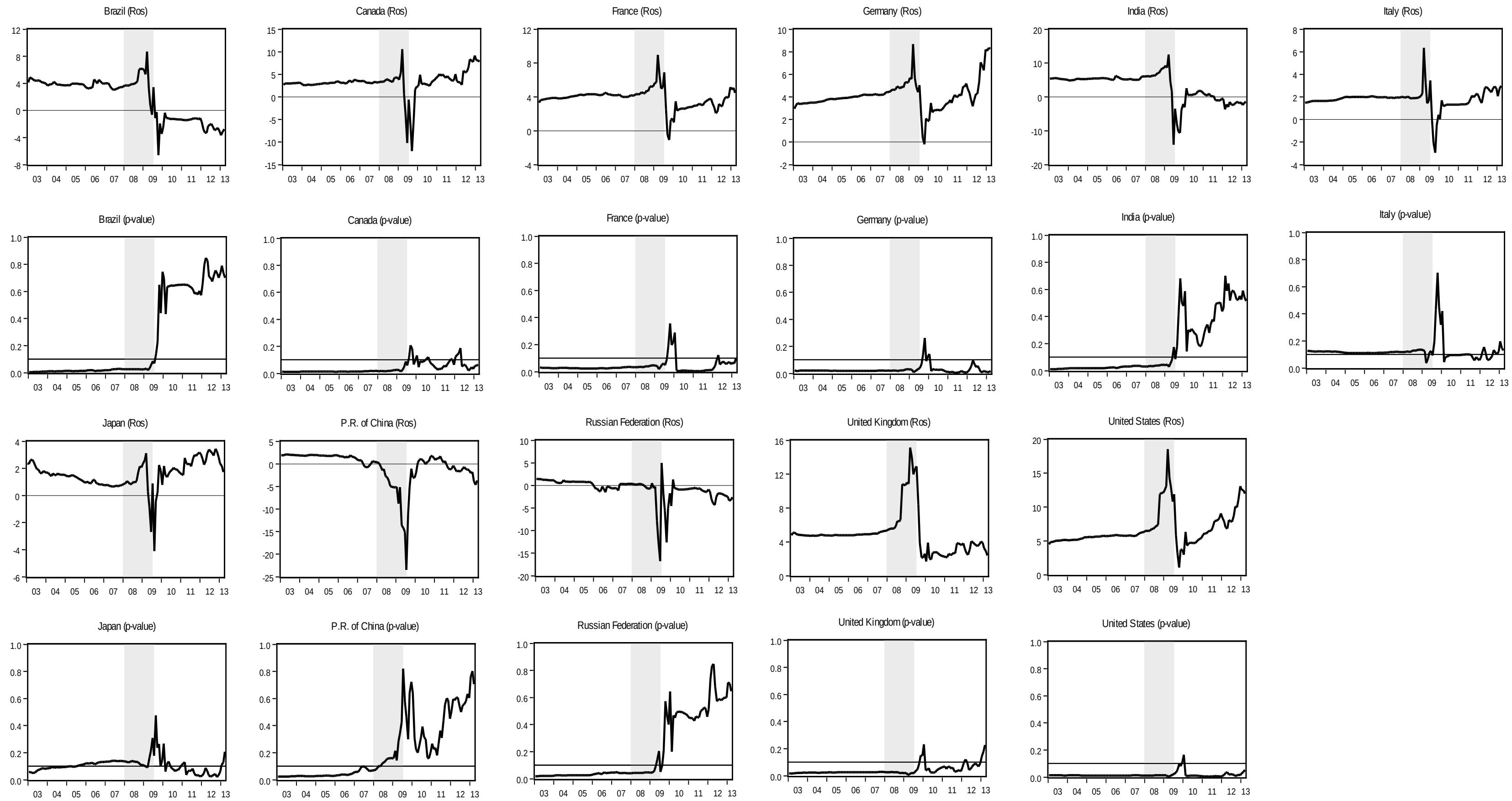


Figure A.6. Tanker Sentiment - $R^2_{OS,\text{CTT}}$

The figure reports the $R^2_{OS,\text{CTT}}$ and the corresponding *p*-values when the estimation period is April 1996 to March 2003, and the beginning of the various forecast evaluation periods runs from April 2003 through April 2013. The end point of the out-of-sample period remains is always April 2014. Shaded areas represent the recent subprime and financial crises.



References

Amihud, Y. "Illiquidity and stock returns: cross-section and time-series effects." Journal of Financial Markets, 5 (2002), 31-56.

Amihud, Y., and H. Mandelson. "Asset pricing and the bid-ask spread." Journal of Financial Economics, 17 (1986), 223-249.

Baker, M., and J. Stein. "Market liquidity as a sentiment indicator." Journal of Financial Markets, 7 (2004), 271-299.

Baker, M., and J. Wurgler. "The equity share in new issues and aggregate stock returns." Journal of Finance, 55 (2000), 2219-2257.

Baker, M., and J. Wurgler. "A catering theory of dividends." *Journal of Finance*, 59 (2004), 1125-1165.

Baker, M., and J. Wurgler. "Investor sentiment and the cross-section of stock returns." *Journal of Finance*, 61 (2006), 1645-1680.

Baker, M., and J. Wurgler. "Investor sentiment in the stock market." *Journal of Economic Perspectives*, 21 (2007), 129-151.

Baker, M.; J. Wurgler; and Y. Yuan. "Global, local and contagious investor sentiment." *Journal of Financial Economics*, 104 (2012), 272-287.

Barber, B.M.; T. Odean; and N. Zhu. "Do retail trades move markets?" *Review of Financial Studies*, 22 (2009), 151-186.

Ben-Rephael, A.; A. Kandel; and A. Wohl. "Measuring investor sentiment with mutual fund flows." *Journal of Financial Economics*, 104 (2012), 363-382.

Brennan, M.J., and A. Subrahmanyam. "Investment analysis and price formation in securities markets." *Journal of Financial Economics*, 38 (1995), 361-381.

Brown, G.W., and M.T. Cliff. “Investor sentiment and the near-term stock market.” *Journal of Empirical Finance*, 11 (2004), 1-27.

Campbell, J.Y., and R.J. Shiller. “Valuation ratios and the long-run market outlook.” *Journal of Portfolio Management*, 24 (1998), 11-26.

Campbell, J.Y., and S.B. Thompson. “Predicting the equity premium out of sample: can anything beat the historical average?” *Review of Financial Studies*, 21 (2008), 1509-1531.

Chordia, T.; A. Subrahmanyam; and V.R. Anshuman. “Trading activity and expected stock returns.” *Journal of Financial Economics*, 59 (2001), 3-32.

Clark, T.E., and K.D. West. “Approximately normal tests for equal predictive accuracy in nested models.” *Journal of Econometrics*, 138 (2007), 291-311.

Fama, E.F., and K.R. French. “Disappearing dividends: changing firm characteristics or lower propensity to pay?” *Journal of Financial Economics*, 60 (2001), 3-44.

Fama, E.F., and K.R. French. “Size, value and momentum in international stock returns.” *Journal of Financial Economics*, 105 (2012), 457-472.

Fisher, K., and M. Statman. “Market timing in regressions and reality.” *Journal of Financial Research*, 29 (2006), 293-304.

Greenwood, R., and S.G. Hanson. “Waves in ship prices and investment.” *Quarterly Journal of Economics*, 130 (2015), 55-109.

Huang, D.; F. Jiang; J. Tu; and G. Zhou. “Investor sentiment aligned: a powerful predictor of stock returns.” *Review of Financial Studies*, 28 (2015), 791-837.

Hwang, B.-H. “Country-specific sentiment and security prices.” *Journal of Financial Economics*, 100 (2011), 382-401.

Kahneman, D. “Thinking fast and slow.” New York: Macmillan, (2011) Kalouptsidi, M. “Time to build and shipping prices.” *American Economic Review*, 104 (2014), 564-608.

Kumar, A., and C. Lee. “Retail investor sentiment and return comovement.” *Journal of Finance*, 61 (2006), 2451-2486.

Kurov, A. “Investor sentiment, trading behaviour and institutional efficiency in index futures markets.” *The Financial Review*, 43 (2008), 107-127.

Lee, W.; C, Jiang; and D. Indro. “Stock market volatility, excess returns and the role of investor sentiment.” *Journal of Banking and Finance*, 26 (2002), 2277-2299.

Lemmon, M., and E. Portniaguina. “Consumer confidence and asset prices: some empirical evidence.” *Review of Financial Studies*, 19 (2006), 1499-1529.

Lesmond, D.A.; J.P. Ogden; and C.A. Trzcinka. “A new estimate of transaction costs.” *Review of Financial Studies*, 12 (1999), 1113-1141.

Menkhoff, L., and R. Rebitzky. “Investor sentiment in the US-dollar: longer-term, nonlinear orientation on PPP.” *Journal of Empirical Finance*, 15 (2008), 455-467.

Neal, R., and S.M. Wheatley. “Do measures of investor sentiment predict returns?” *Journal of Financial and Quantitative Analysis*, 33 (1998), 347-400.

Pastor, L., and R.F. Stambaugh. “Liquidity risk and expected stock returns.” *Journal of Political Economy*, 111 (2003), 642-685.

Rapach, D.E.; J.K. Strauss; and G. Zhou. “International stock return predictability: what is the role of the United States?” *Journal of Finance*, 68 (2013), 1633-1662.

Roll, R. “A simple implicit measure of the effective bid-ask spread in an efficient market.” *Journal of Finance*, 39 (1984), 1127-1139.

Schmeling, M. “Investor sentiment and stock returns: some international evidence.” *Journal of Empirical Finance*, 16 (2009), 394-408.

Swaminathan, B. “Time-varying expected small firm returns and closed-end fund discounts.” *Review of Financial Studies*, 9 (1996), 845-887.

Wang, C., “Investor sentiment and return predictability in agricultural futures markets.” *Journal of Futures Markets*, 21 (2001), 929-952.

Welch, I., and A. Goyal. “A comprehensive look at the empirical performance of equity premium prediction.” *Review of Financial Studies*, 21 (2008), 1455-1508.