BANK OF GREECE
EUROSYSTEM

Working Paper

Is there accuracy of forward freight agreements in forecasting future freight rates? an empirical investigation

Evangelia Kasimati
Nikolaos Veraros

230

JUNE 2017

WORKINGPAPERWORKINGPAPERWORKINGPAPERWO

BANK OF GREECE
Economic Analysis and Research Department – Special Studies Division
21, E. Venizelos Avenue
GR-102 50 Athens
Tel: +30210-320 3610
Fax: +30210-320 2432

www.bankofgreece.gr

Printed in Athens, Greece
at the Bank of Greece Printing Works.

ISSN 1109-6691

IS THERE ACCURACY OF FORWARD FREIGHT AGREEMENTS IN FORECASTING FUTURE FREIGHT RATES? AN EMPIRICAL INVESTIGATION

Evangelia Kasimati
Bank of Greece

Nikolaos Veraros
King's College London

Abstract

Participants in the maritime industry place much interest in the Forward Freight Agreements (FFA/FFAs), being an indispensable tool for hedging shipping freight risk. Our paper innovates by directly comparing the FFA predictions with their actual future settlement prices as well as by examining contracts going forward as far as next calendar year. We combine straightforward comparison measurements with cointegration analysis to test for the accuracy and efficiency of the FFA projections. We find that FFAs display limited usefulness in predicting future freights, only slightly superior than simple naïve models. The shorter the contract period and the smaller the vessel the better the forecast. We also find FFAs being relatively good predictors of future market direction but missing the turning points of the market cycles.

JEL-classifications: C32, G13, G14

Keywords: Shipping, Freight Rates, Forward Freight Agreements, Forecasting, Vector Error Correction Models

Acknowledgments: The authors would like to particularly thank Stefanos Frangos from Navios who kindly provided access to all required data series. The views in this paper are those of the author and do not necessarily reflect those of the Bank of Greece.

Correspondence:
Evangelia Kasimati,
Bank of Greece,
Eleftheriou Venizelou 21,
Athens 10250, Greece,
Tel: +30 210 3202649,
ekasimati@bankofgreece.gr

1. Introduction

Forward Freight Agreements (FFA/FFAs) represent swap contracts on freight rates for ocean going vessels. FFAs provide a mechanism for hedging freight rate risk in the shipping market. Main users include shipowners needing to hedge the future income of their vessels as well as shippers wishing to hedge the cost of carrying products. Other participants also include traders opening speculative positions based on their views about the future developments in the shipping values. The prices of the FFA contracts are expressed either as dollars per day or as dollars per ton, effectively simulating the charter income of a vessel. FFA settlements take place at the end of each month. The settlement price is calculated as the average spot prices of either the whole month, or the last 7 days, depending on the contract type. In our analysis we focus on the FFAs for dry-bulk vessels, as opposed to tankers, being the most popular both in actual trading and academic investigation. The main specifications of the four different types of dry-bulk vessels to which FFAs refer are summarized in Table 1. For each vessel type there are different standardized routes for which separate FFA contracts apply, as well as one contract for the average of all routes. Figure 1 demonstrates the development in the FFA volumes since 2007, in comparison with the fluctuations in the level of the spot freight rates as expressed through the Baltic Dry Index (BDI), internationally accredited as the most acknowledged indicator of the market.

[Insert Table 1 here]

[Insert Figure 1 here]

The trading of the FFAs takes place over the counter (OTC). Following the collapse in the freight market in the aftermath of the Lehman crisis in late 2008, the clearing of the FFAs has increasingly switched to clearing houses, mainly LCH.Clearnet, so as to minimize the counterparty risk (see percentage of cleared volumes at the bottom part of Figure 1). The reporting of the FFA closing prices (being the BFAs), as well as the determination of their monthly settlement prices is performed by The Baltic Exchange (http://www.balticexchange.com) , through corresponding assessments by a panelist of shipbrokers (The Baltic Exchange, 2009). The Baltic Exchange, based in London, is a membership organization originally established in 1744 which includes as members shipbrokers, freight and derivative brokers, trading houses and shipowners. It serves as the world's most reliable independent source of maritime information for trading and settlement of physical assets and derivative contracts.

The participants in the FFA market effectively take a bet on the future freight rates. For example, by going short at the contract of the next calendar year the trader will make a profit if the spot rates to be realized in the next calendar (twelve monthly settlement prices) prove to be lower than the FFA price that she contracted today. This brings forward the crucial question of the extent to which the FFAs are good predictors of the future freight rates, which is exactly the topic that we examine in this paper.

The main feature of the FFA market is that the underlying service (i.e. carriage of goods through a vessel) cannot be stored. Accordingly, a cost-of-carry relationship between spot and forward prices cannot be established, thus leaving room for more degrees of freedom in the relationship between the two. In addition, the two markets have totally different structures. In the spot market we effectively transact on actual vessels' chartering whereas in the FFAs it is merely paper trading. Accordingly, in the spot market the transaction costs (not only limited to cash expenses but also including barriers to entry, amount and time horizon of committed resources, as well as ability to operate) are significantly higher than in the FFA market (Kavussanos, 2006).

2. Literature review

In the related academic literature, Alizadeh and Nomikos (2003) examined the extent to which the forward prices were good predictors of future spot prices. They specifically focused on the ability of the FFAs to correctly predict the future direction of the market (up or down). They concluded that it was only for one-month predictions that FFA forecasts were better than tossing a coin. For longer periods (two and three months) the success in the forecast was not significantly different from 50%.

Kavussanos & Visvikis (2004) investigated the lead-lag relation between spot and forward rates both in returns and volatilities based on a VECM. Through a cointegration analysis, a bi-directional relation was identified; nevertheless, FFA returns seemed to incorporate faster the new information than spot. Authors attributed this to the tendency of the market participants to prefer trading in the market with the lower transaction costs. On the investigation of the volatility behaviour through an extended VECM-GARCH model, the results also indicated that FFAs played a leading role in incorporating new information, although to a lesser extent when compared to the returns analysis. The overall conclusion was that FFA prices contained useful information on future spot prices and therefore could be used as price discovery vehicles.

Kavussanos *et al.* (2004) investigated the unbiasedness hyphothesis for FFA forecasts through the use of cointegration analysis. They found that for forecasts of one and two months the predictions were unbiased, whereas for three months the results were mixed.

Batchelor *et al.* (2007) compared different models, including ARIMA, VAR and VECMs both restricted and unrestricted, to assess the forecasting performance of the FFAs, They split the data set between in-sample and out-of-sample. As in Kavussanos and Visvikis (2006) they found that VECMs fitted better the data. In their in-sample analysis, Granger causality suggested a bi-directional relation, however forward rates seemed to lead spot rates for longer periods. In their out-of-sample analysis they forecasted up to 20 days forward for both spot and forward rates. They inferred that forward rates were much harder to predict than spot rates. In addition, their VECMs outperformed the other models in the forecast of the spot rates, but not in the forecast of the forward rates.

Recent literature addresses more specialized issues, thus being of more limited usefulness for our paper. Kavoussanos *et al.* (2014) used high frequency data to synchronize developments in the commodities markets with freight rates of vessels used to carry the particular commodities. They concluded that new information appeared first in the returns and volatilities of the commodities future markets before spilling over into the FFA market. Alizadeh *et al.* (2015) studied the importance of liquidity measures on the FFA returns. Such measures included bid –ask spread and the ratio of returns on volume traded. By adopting a pooled cross sectional time series regression they concluded that liquidity measures played a significant role in determining the returns of near term FFA contracts. Tezuka *et al.* (2012) extended a model developed by Bessembinder and Lemmon (2002) for the electricity market in order to derive the equilibrium spot and forward prices in shipping, in an effort, *inter alia,* to examine the reasons why bias existed between expected spot and forward prices for forecast periods longer than two months.

Groder (2010) has been among the very few researchers to compare the FFA prices directly to their future settlements. He reconstructed his data set into monthly observations and examined FFA forecasts for up to three months forward. He was also among the few to employ Root Mean Square Error to test the accuracy of the predictions. He concluded that FFAs were not superior forecasts than simple random-walk models.

Zhang et al. (2014) and Zhang (2015) enhanced the existing academic literature by investigating the two way lead – lag relationship between spot and forward markets and between spot and time charter markets. Their spot freight forecasting method for analysis and prediction of dry bulk market based on the price discovery of FFA and time charter (TC). Their research results show that FFA and TC through the VECM models perform better than spot through VAR models. This implies that both FFA and TC rates can improve the forecasting performance of spot rates. The TC market has a price discovery function that is similar to the FFA market, while a smaller ship type and a longer charter period lead to a stronger price discovery function. Both papers concluded that, the integration of FFA and TC rates can further improve the forecasting performance of spot freight rates.

As discussed above, the existing academic literature relies heavily on testing the unbiasedness hypothesis in the forward freight market. In an efficient market with unbiased predictions all new information should be immediately incorporated into the FFA prices. Unbiasedness has been determined in almost all of the studies, at least for the short-term forecasts. However, as Groder (2010) reasonably argues, an unbiased prediction is not necessarily a correct, or even a good, prediction. Unbiased forecasts could still diverge hugely from their future settlement prices, thus becoming practically useless for the market participants. A measure of the deviation between the forward prices and their future spot realizations is indispensable for evaluating the forecasting ability of the FFAs.

The innovation of our paper is that we specifically focus on measuring the accuracy of the FFA predictions by comparing forward prices with their settlements realized in subsequent periods. This contrasts with most of the existing studies in which the data set of the researchers includes pairs of same-day spot and FFA prices, mostly trying to identify lead-lag relations in prices reported simultaneously. Compared to Groder (2010) we extend our comparisons up to the next calendar year as opposed to only 3 month forward. We also employ the longest time framework with better quality of data compared to early studies which relied mostly on brokers' inputs. Finally, our study covers all four types of dry-bulk vessels, thus allowing for a full range of useful insights for the market participant.

We employ tests including Root Mean Square Errors (RMSE), Mean Absolute Errors (MAE) and subsequent calculations of Theil coefficients for analyzing the relative accuracy in the predictions. Due to the limited use of such tests in the existing literature of the freight derivatives, we draw extensively from the literature of evaluating forecasts in macroeconomic variables. Such studies evaluate short-term and long-term forecasts on main macroenonomic variables from organizations like the IMF, UN, OECD, EC and WB. Artis (1996) and Artis and Marcellino (2001) tested the extent to which the mean forecast error was zero (unbiased), evaluated the accuracy of forecasts through the RMSE test, and calculated derived Theil tests to compare against naïve models. They also tested the efficiency of the predictions by regressing actuals over forecasts in a so called "realization regression". Artis (1996) further tested the directional accuracy and the ability to forecast the cycle. A similar approach has been employed by Kreinin (2000) who also tested the ability to forecast turning points by specifically studying the predictions in upturns and downturns. Finally Hong and Tan (2014), in a most recent study, employed RMSE, MAE, and Mean Absolute Percentage Errors (MAPE) to compare the accuracy of macroeconomic predictions by UN, IMF, and WB.

The rest of our paper is organized as follows. Section III describes the data. Section IV presents the empirical results on forecast accuracy, efficiency, direction and cycle turning points. Finally, section V discusses our conclusions.

3. Data description

In our analysis we use BFAs for all four types of dry-bulk vessels (i.e. CAPES, PMX, SMX, HS). We select forward freight contracts for the average freights of the reported routes, settling with the full month spot rates, since these are the most liquid and highly traded. Our data covers the interval from Jan 2006 until Feb 2014. To our knowledge this is the largest time span ever used in a similar study. In addition, it includes a period of high volatility with significant swings in the freight market, thus testing the FFAs predictive power under challenging conditions. Data is in levels and is derived from the official source of the Baltic Exchange as opposed to other studies which used individual brokers' assumptions (Kavussanos and Visvikis, 2004, Kavussanos *et al.,* 2004, Batchelor *et al.,* 2007). We use BFAs for next month (short-term forecast), next quarter (medium-term forecast), and next calendar year (long-term forecast). To our knowledge this is the first study to evaluate the

FFAs forecasting ability for forward periods longer than the next three months. Apparently, for market participants, the longer the forecasting period, the more valuable the forecast. Similarly to Groder (2010), we reorganize our original daily data into monthly ones, as the monthly settlement prices are our basis for comparing predictions to realizations. For our naïve scenarios, against which we compare the FFA predicting performance, we use the historical spot prices for the last month, last quarter and last 12 months so as to match our corresponding forecast horizons. We also experiment using the last month historical spot prices to forecast the longer periods (next quarter and next calendar) in order to test the assumption that the most recent market information is the best prediction of the future. As an example, on January 2011, our prediction for the next quarter SMX freights consists of the average observations within the month of January of the BFAs for the particular vessel type for Q2 2011. Our realization is the average of the actual monthly settlement prices for April, May and June 2011, and our naïve model is the average spot prices of November 2010, December 2010 and January 2011, or alternatively only Janueary 2011 being the most updated historical information. We compare the accuracy of the BFA prediction of the future settlement with that of the historical spot prices. We have excluded as outliers the observations related to settlements in Q4 2008 due to the short-term complete collapse in the freight rates provoked by the Lehman Brothers crisis (see Figure 1). Haralambides and Thanopoulou (2014) distinguished that particular crisis from previous shipping cycles arguing that the shipping market after the event was in a completely different territory. During that period letters of credit were hard to obtain from the international banking system, thus halting worldwide trade (Porter *et al*., 2011). Chor and Manova (2012) particularly measured the peak in the trade contraction in USA in November 2008 when import and exports fell by ca 23% and 14% respectively on a monthly basis. They also referred to a broad-based increase in the price of various trade-related credit instruments in the period Oct 2008 – Jan 2009, adversely affecting international trade. Our adjustment resulted in a loss of only three observations in our data set. Table 2 summarizes the main descriptive statistics of our data.

[Insert Table 2 here]

4. Empirical results

Forecast accuracy

We start our analysis with the measurement of the forecast errors of the FFAs, being the most important aspect for the maritime practitioner. On the data described in the previous section we define:

f_t= forecast (FFA or naïve) at month *t* of the spot freight in the future

g_t=realization of the spot freight related to forecast f_t

$e_t=(f_t-g_t)/g_t$

We specify below the particular mathematical functions for measuring the forecast errors

$$MAPE = \sum_{t=1}^{T} \frac{|e_t|}{T} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(f_t - g_t)^2}{T}} \tag{2}$$

$$THEIL1 = \frac{\sqrt{\frac{\sum_{t=1}^{T}(f_t - g_t)^2}{T}}}{\sqrt{\frac{\sum_{t=1}^{T} f_t^2}{T}} + \sqrt{\frac{\sum_{t=1}^{T} g_t^2}{T}}} \tag{3}$$

$$THEIL2 = \frac{THEIL1\ (BFA)}{THEIL1\ (NAIVE)} \tag{4}$$

Similar to Hong and Tan (2014) we calculate the MAPE using observations representing percentage differences between prediction and realization. This facilitates the comparison between different types of vessels commanding difference levels of freight, with the larger vessels securing higher freights and displaying larger volatility.

For the RMSE test we use observations representing absolute differences (similar to Groder 2010). This allows for the subsequent estimation of the Theil scores comparing the forecasting performance of alternative models. The RMSE is a commonly used test for forecast evaluation as it assumes a symmetric loss function, which seems reasonable given the high monetary value of the FFA contracts (Batchelor *et al.,* 2007). Finally, similar to Artis (1996) and Groder (2010), we use Theil tests to compare the forecast performance between different vessel types (Theil 1), as well as between FFA and the naïve models (Theil 2). Theil 1 takes values between 0 and 1, with the better predictions being closer to zero. For Theil 2, values lower than 1 indicate that FFAs demonstrate lower forecast errors than Naïve. We run the above tests both for the FFAs and the naïve model and we summarize our results in Table 3, along with some basic descriptive statistics of the time series.

[Insert Table 3 here]

The predictability for longer periods (next calendar vs. next month) and larger vessels (CAPES vs HS) is rather limited. This is in line with Kavussanos *et al.* (2004) who found biased estimations for FFA forecast longer to two-months. The CAPES demonstrate means of the forecast errors which are always statistically different from zero, suggesting bias in the prediction. In addition, MAPE, being effectively the average percentage error, scores between 42% to 58%, depending on the forecasting period. On the other hand, smaller vessels like the SMX and HS demonstrate means equal to zero for both next month and next quarter, whereas their MAPE values are in the region of only 11% for the next month and 20% for the next quarter. Based on the previous figures, CAPES demonstrate between 2.3x and 3.8x higher prediction error than PMX and SMX/HS respectively. The comparison of the FFAs with the naïve scenarios suggests that both their means and variance are not statistically different across all vessel types for next month and next quarter. This is in line to the findings of Groder (2010). Theil2 suggests only a slight superiority of the FFAs. Interesting enough is that when comparing means and variances the only category that FFAs rank clearly better than naïve is in the long-term projections (calendar). Nevertheless, long-term projections display by far the worse quality of forecasts.

No significant differences to the above results were identified when we adjusted our data so that they all start uniformly from September 2009, in order to be more comparable in terms of time span. Moreover, no improvement in the evaluation of the naïve predictions *vis a vis* the FFAs was observed when using the most recent (last month) historical observation to predict next quarter or calendar.[1] Conclusively, our findings indicate that the

[1] Results for using data series uniformly starting from September 2009 and naïve model using exclusively last month freights are omitted to save space. Available upon request to the authors

practitioner is not likely to derive much value by relying on the FFAs as predictions of future spot rates.

Forecast Efficiency

We then turn our analysis to the study of the efficiency of the predictions. According to Artis (1996), the following simple “realization regression” tests for the weak efficiency of the forecasts.

$$g_t = \beta_0 + \beta_1 f_t + \mu_t \qquad (5)$$

The extent of efficiency is examined through the joint test that $\beta_0 = 0$ and $\beta_1 = 1$, in which case g_t is 100% derived by the price of f_t. However, since most macroeconomic time series are found to be non-stationary, the cointegration framework developed by Johansen (1988, 1991) is considered a more adequate specification. Kavussanos *et al.* (2004, 2006) and Batchelor *et al.* (2007) concluded that a VECM structure is the most efficient in testing the FFAs efficiency in the presence of non-stationary data. Accordingly, we adopt the following basic form:

$$\Delta X_t = \mu + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \Pi X_{t-1} + \varepsilon_t; \qquad \varepsilon_t \sim N(O, \Sigma) \qquad (6)$$

where X_t is a 2 x 1 vector including $(g_t, f_{t,t-n})'$, μ is a 2 x 1 vector including linear trend, and intercept, Δ denotes first differences, ε_t is a 2 x 1 vector including the residuals, Σ is the variance covariance matrix of the residuals, Γ_i is the coefficient estimates for the short-term elements of the model and Π for the long term which is effectively the cointegration relation. Our analysis focuses on the long-term coefficients Π through testing the restriction that $\beta_0 = 0$ and $\beta_1 = 1$ in the following cointegration relation:

$$\beta' X_{t-1} = (1\ \beta_0\ \beta_1)(g_{t-1}\ 1\ f_{t-1,t-n-1})' \qquad (7)$$

where $f_{t-1,\, t-n-1}$ is the forward price at time t-n for delivery at time t lagged by one period, and g_{t-1} is the spot price at the maturity of the contract lagged by one period.

We start our analysis by testing for the presence of unit roots in our time series in order to determine the existence of stationarity (Diebold, 2004). As expected, Table 4 demonstrates that under both Augmented Dickey-Fuller and Phillips-Perron tests all variables exhibit the existence of unit roots, thus suggesting the absence of stationarity. The only marginal result stands for spot monthly CAPES which demonstrates borderline rejection at 5% significance, depending on which of the two test we employ.

[Insert Table 4 here]

We then isolate the cointegration relation within the broader VECM in order to test for the joint restriction of $\square_0$ = 0 and $\square_1$ =1. Table 5 presents our results. No cointegration relation is identified for all Calendar models, thus indicating that the next calendar forecasts are not efficient. This is mostly in line with the earlier findings in the previous section of our paper. One cointegration relation is identified for the rest of the models, with the joint test easily accepted for all but Quarter CAPES which scores marginally.

An explanation for the non-efficiency of the calendar contracts could be a possible structural break as Figure 2 indicates. Based on the visual observation of the graph and the findings of Chor and Manova (2012) already discussed in earlier section, we experiment with setting the breakpoint at 2008M11. Our Chow test (Diebold and Chen, 1996) reconfirms the existence of two different regimes around this date (Table 6, Panel A). The only exception is HS, mainly due to the fact that its time series started only 10 months before the breakpoint date. However, when we run our models for the second regime, displaying reduced volatility, we again fail to identify cointegration relation, as the results in Table 6 Panel B indicate.

[Insert Table 5 here]

[Insert Figure 2]

[Insert Table 6 here]

Forecast direction

Similarly to the methodology of Artis (1996) we try to assess the ability of the FFAs to correctly predict the direction of the market in future realizations. As a basis for establishing the direction of the market we use the current spot freight. For example, if in January 2008, the FFA is higher than the actual January 2008 spot we assign a positive sign to the forecast (meaning that FFA predicts an increase compared to current freight levels). If the future settlement of the particular FFA contract proves to be also higher than the Jan 2008 spot we assign again a positive sign to the realization, and we conclude that the FFA predicted correctly the direction of the market. This is also in line with the approach of Alizadeh and Nomikos (2003). We perform this analysis for both the FFAs and naïve models. We summarize our results in Table 7, together with the respective statistical tests.

[Insert Table 7 here]

Contrary to the findings of Alizadeh and Nomikos (2003) in our study the FFAs score better in long-term (calendar) predictions. In all cases FFAs predict better than the naïve models. In addition, in all but one case, the percentage of correct predictions is higher than 50%. We have marked with gray color the observations which reject the random walk hypothesis (null) at the 5% level of statistical significant (seven out of twelve). The naïve model cannot reject the null hypothesis in all cases with the exception of only two in calendar forecasts. The longer term of historical trend embodied in 12-month historical average prices seems valuable in predicting longer term forward freights. Nevertheless, if we compare those long-term naïve forecasts to the corresponding FFAs, they are clearly inferior.

Forecast cycle

We then turn our analysis to evaluating the ability of the FFAs to forecast the changes in the market cycles, i.e. to time correctly the turning points in the market cycles (shifting from rising to falling market and *vice versa*). We limit the analysis to long-term and medium-term forecasts. The intense swings in freight rates in the first half of our sample period provides an excellent basis for testing this ability. Similarly to Artis (1996) we visually display our analysis through Figures 3 to 6.

[Insert Figures 3 – 6 here]

In the long-term (calendar) forecasts, the FFAs fail to predict the market change in the high volatility period of 2006 – 2009. This applies even when we focus on the most recent observations of the forecasts (i.e. last three months before the forecasted calendar, represented by crossed points in the graphs) expected to incorporate the most updated information for the prediction. Although these observation generally display the lowest errors from the realizations, they are again mostly on the wrong side of the cycle, i.e. they predict a continuous rise when the market actually starts falling. In the medium term (quarterly) the CAPES and PMX still fail to capture the change over the cycle in the period

2006 - 2009, even if we limit the comparison to the most updated 3rd month observation. For SMX and HS the results are more mixed.

5. Conclusions

FFAs have historically been the main instrument providing hedging ability to the shipping market participants, thus their trading behaviour is of high importance for the maritime industry.

Our analysis innovates by focusing on comparing directly the FFA predictions with their actual future realizations. Our period under investigation extends to 9 years, being the largest time span ever used in a similar study, effectively covering most of the development of the market in its current state. We also include the high volatility period of 2007 – 2008. In addition, we differentiate our results among the four different dry-bulk vessel types, thus allowing industry participants to draw conclusions about the specific assets that they operate.

We concluded that the forecast ability of the FFAs is of limited value, particularly when referring to long-term predictions which would be clearly of higher importance for the practitioner. The quality of the forecasts is improved for shorter periods of projection and smaller size of vessels. Although there seems to be a small superiority compared to naïve models, which are simply based on historical values, this is not found to be statistically significant. In line with the existing academic literature, long-term FFA predictions are found to be biased. FFAs seem to be relatively efficient in forecasting the direction (up or down) of the movement in the future freights, however they miss the turning points of the cycles. The limited ability of the FFAs to predict future spot freights seems in line with the inability of expressing period freight rates as a kind of average of expected future spot rates, evidenced long ago by Hale and Vanags (1989).

According to Groder (2010), the limited forecast ability of the FFAs could be the result of either the low ability of the market participants to predict future freights or the emergence of new unpredictable information between now and the maturity dates of the contracts.

We would also add that given the low ability for immediately trading in the actual spot market (i.e. charter a vessel) the industry participants perceive the FFA market as an efficient alternative to trade on current market developments and not so much on future market expectations. As an example of our last statement we experimented with a simple trading rule stating that if the price of the spot has changed for three consecutive days in the same direction, it will also follow through in the fourth day. In an efficiency market, the probability of success would not be greater than 50% (random walk). Our test provided an astonishing 85% probability of success in the spot market, compared to only 56% in the FFAs market. This clearly demonstrates that spot prices are very difficult to trade (i.e. physically charter a vessel), thus leaving the FFA market as the efficient vehicle to immediately consolidate the current changes in the freight market fundamentals. This also further reconfirms existing literature that identifies a faster incorporation of the new information in the FFA market as opposed to the spot. Further academic research would be useful to better analyze the mixture of current spot market developments and future freight expectations in the behaviour of the FFA prices. In addition, taking into account Patton's research (2011), future research could estimate ARMA-GARCH models to perform out of sample forecasts, checking out whether these forecasts give better outcomes if FFAs are also included in the models. The latter might show whether FFAs hold additional information for future prices, compared to the price history only. The fact that FFAs are not good predictors is a well-documented result in all future/derivative markets (i.e. stocks, exchange rates, etc) since new information arrives every day or minute. Maybe this is explained by the fact that as FFAs are a less liquid bet (since they are OTC instruments) compared to stocks or exchange rates and can hold some information for some days or minutes. Finally, there are other factors (such as transaction costs or other asymmetries) that could improve the forecast.

References

Alizadeh, A., K. Kappou, D. Tsouknidis, and I. Visvikis. 2015. "Liquidity effects and FFA returns in the international shipping derivatives market." *Transportation Research Part E: Logistics and Transportation Review* 76: 58-75.

Alizadeh, A., and N. Nomikos. 2003. "Do FFAs provide good forecasts?" *Lloyd's Shipping Economist*, December, 32 -34.

Artis, M. 1996. "How accurate are the IMF's short-term forecasts? Another examination of the World Economic Outlook." *IMF Working Paper,* No 96/89.

Artis, M. and M. Marchellino. 2001. "Fiscal Forecasting: The track record of the IMF, OECD, and EC." *Econometrics Journal* 0: 1-23.

Batchelor, R., A. Akizadeh and I. Visvikis. 2007. "Forecasting spot and forward prices in the international freight market" *International Journal of Forecasting* 23: 101-114.

Bessembinder, H. and M. L. Lemmon. 2002. "Equilibrium pricing and optimal hedging in electricity forward markets." *Journal of Finance* 57: 1347-1382.

Chor, D. and K. Manova. 2012. "Off the cliff and back? Credit conditions and international trade during the global financial crisis." *Journal of International Economics* 87: 117-133

Diebold, F. X. 2004. *Elements of Forecasting*, 3rd ed. Canada: Thomson South Western.

Diebold F. and C. Chen. 1996. "Testing structural stability with endogenous breakpoint: A size comparison of analytic and bootstrap procedure." *Journal of Econometrics* 70: 221 -241.

Groder, P. 2010. "Forecasting shipping freight rates." Working Paper, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1597113.

Hale, C. and A. Vanags. 1989. "Spot and period rates in the dry bulk market." *Journal of Transport Economics and Policy* 23: 281-291.

Haralambides H. and H. Thanopoulou. 2014. "The Economic Crisis of 2008 and World Shipping: Unheeded Warnings." *Spoudai* 64: 5-13

Hong, P. and Z. Tan. 2014. "A comparative study of the forecasting performance of three international organizations" *Journal of Policy Modeling* 36: 745-757

Johansen, S. 1988. "Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control."* 12: 231-254.

Johansen, S. 1991. "Estimation and Hypothesis testing of cointegration vectors in Gaussian vector autoregressive models." *Econometrica* 59: 1551-1580.

Kavussanos, M. G. and I. D. Visvikis. 2004. "Market interactions in returns and volatilities between spot and forward sipping freight markets." *Journal of Banking & Finance* 28: 2015-2049.

Kavussanos, M. G. and I. D. Visvikis. 2006. "Shipping freight derivatives: a survey of recent evidence." *Maritime Policy & Management* 33: 233-255.

Kavussanos, M. G., I. D. Visvikis, and D. N. Dimitrakopoulos. 2014. "Economic spillovers between related derivatives markets: The case of commodity and freight markets." *Transportation Research Part E: Logistics and Transportation Review* 68: 79-102.

Kavussanos, M. G., I. D. Visvikis and D. Menachof. 2004. "The unbiasedness hypothesis in the freight forward market: evidence from cointegration tests" *Review of Derivatives Research* 7: 241-266.

Kreinin, M. 2000. "Accuracy of OECD and IMF projections." *Journal of Policy Modeling* 22: 61-79.

Patton, A.J. 2011. "Volatility forecast comparison using imperfect volatility proxies." *Journal of Econometrics* 160: 246-256.

Porter, R. B., R. R. Glauder and T. J. Healey. 2011. "*New Directions in Financial Services Regulation*." USA: The MIT Press.

Tezuka, K., I. Masahiro and M. Ishizaka. 2012. "An equilibrium price model of spot and forward shipping freight markets." *Transportation Research Part E: Logistics and Transportation Review* 48: 730-742.

The Baltic Exchange. 2009. "*Manual for Panelists. A Guide to Freight Reporting and Index Production*." London: The Baltic Exchange.

The Baltic Exchange. 2014. "*The Baltic Code*." London: The Baltic Exchange.

Zhang H. and Zeng Q. 2015. "A study of the relationships between the time charter and spot freight rates." *Applied Economics*, 47:9, 955-965

Zhang J., Zeng Q. and Zhao X. 2014. "Forecasting spot freight rates based on forward freight agreement and time charter contract", *Applied Economics*, 46:29, 3639-3648.

Table 1. Types of Dry-Bulk Vessels Used as Underlying for FFA contracts

Vessel	Deadweight (tons)	Geared	Age
Capesize (CAPE)	172,000	No	<10yrs
Panamax (PMX)	74,000	No	<12yrs
Supramax (SMX)	52,454	Yes	<10yrs
Handysize (HS)	28,000	Yes	<15yrs

Source: The Baltic Exchange (2009, 2014)

Table 2 Data Descriptive Statistics

	BFAs CAPE month	CAPE quarter	CAPE calendar	PMX month	PMX quarter	PMX calendar	SMX month	SMX quarter	SMX calendar	HS month	HS quarter	HS calendar
Mean	18,958	42,724	37,254	13,906	23,953	21,267	13,776	21,462	19,310	10,384	15,415	13,914
Median	15,432	27,103	24,627	12,112	16,710	15,038	12,091	15,435	13,739	9,491	10,642	10,010
Maximum	58,802	162,646	130,608	31,602	82,081	69,080	28,504	65,016	57,234	19,469	42,544	37,596
Minimum	6,172	6,774	8,999	5,428	5,854	6,715	6,847	6,838	7,739	6,043	5,517	6,760
Std. Dev.	11,563	43,283	33,141	7,184	19,795	15,706	5,349	15,489	12,505	3,403	10,858	8,403
Skewness	1.141	1.616	1.602	1.046	1.541	1.600	1.052	1.445	1.459	1.056	1.418	1.481
Kurtosis	3.976	4.448	4.430	2.963	4.257	4.541	3.147	3.885	4.035	3.158	3.565	3.853
Jarque-Bera	16.161	54.878	49.218	11.482	48.466	50.433	11.687	39.966	36.352	11.767	32.426	32.063
Probability	0.000	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.000

	Average Spot CAPE last month	CAPE last 3 months	CAPE last 12 months	PMX last month	PMX last 3 months	PMX last 12 months	SMX last month	SMX last 3 months	SMX last 12 months	HS last month	HS last 3 months	HS last 12 months
Mean	19,221	44,950	48,931	13,686	24,170	26,138	13,886	22,121	23,879	10,424	15,855	17,135
Median	13,315	28,728	38,647	11,943	16,947	21,404	12,474	16,554	20,360	9,686	10,940	13,978
Maximum	70,281	180,458	152,621	34,592	82,605	70,560	31,370	66,243	57,880	21,336	44,311	39,687
Minimum	3,501	4,432	7,023	4,181	5,091	6,980	6,729	5,263	8,570	5,693	4,314	7,184
Std. Dev.	14,239	45,549	40,327	7,928	20,022	17,617	5,860	15,520	14,036	3,861	10,978	9,937
Skewness	1.204	1.541	1.140	1.106	1.423	1.148	1.067	1.386	1.070	1.077	1.312	0.987
Kurtosis	4.315	4.444	3.252	3.271	4.062	3.318	3.402	3.905	3.075	3.440	3.450	2.692
Jarque-Bera	19.756	50.681	21.065	13.025	40.380	21.483	12.381	37.206	17.385	12.685	27.453	13.461
Probability	0.000	0.000	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.001

Table 3. Summary Statistics of the FFAs and Naïve Forecasting Errors

	CAPES					
	Next month (short-term)		Next quarter (mid-term)		Next calendar (long-term)	
	BFA	Naïve	BFA	Naïve	BFA	Naïve
Mean of forecast errors[1]	0.216	0.148	0.124	0.209	0.227	0.733
p-value, (Mean = 0)	0.002	0.055	0.049	0.025	0.004	0.000
St. Dev[1]	0.527	0.602	0.627	0.924	0.707	1.126
Observations	63	63	102	102	84	84
Start date	Sep-09	Sep-09	Jan-06	Jan-06	Jan-06	Jan-06
Skewness[1]	0.858	1.566	1.433	1.998	0.590	0.246
Kurtosis[1]	3.132	5.737	5.134	6.735	2.619	1.619
Jarque-Bera[1]	7.783	45.406	54.237	127.152	5.380	7.527
p-value	0.020	0.000	0.000	0.000	0.068	0.023
MAPE[1]	42%	43%	46%	59%	58%	105%
RMSE[2]	8,354	9,404	20,877	23,140	37,355	47,931
Theil 1[2]	0.183	0.198	0.172	0.187	0.413	0.435
Theil 2[2]	0.921		0.920		0.950	
Mean equal test (vs. Naïve) p-value [1,3]	0.504		0.443		0.001	
Variance equal test (vs. Naïve) p-value [1,4]	0.747		0.149		0.000	

	PANAMAX					
	Next month (short-term)		Next quarter (mid-term)		Next calendar (long-term)	
	BFA	Naïve	BFA	Naïve	BFA	Naïve
Mean of forecast errors[1]	0.089	0.038	0.043	0.043	0.182	0.558
p-value, (Mean = 0)	0.006	0.188	0.225	0.325	0.018	0.000
St. Dev[1]	0.245	0.228	0.355	0.435	0.688	0.955
Observations	63	63	102	102	84	84
Start date	Sep-09	Sep-09	Jan-06	Jan-06	Jan-06	Jan-06
Skewness[1]	1.200	0.508	1.136	2.078	1.000	0.660
Kurtosis[1]	5.095	2.625	4.508	11.317	3.933	2.460
Jarque-Bera[1]	26.638	3.077	31.604	367.440	17.043	7.114
p-value	0.000	0.215	0.000	0.000	0.000	0.029
MAPE[1]	18%	19%	28%	31%	53%	81%
RMSE[2]	3,038	3,061	8,709	9,481	18,649	22,810
Theil 1[2]	0.097	0.097	0.142	0.154	0.390	0.409
Theil 2[2]	0.997		0.920		0.954	
Mean equal test (vs. Naïve) p-value [1,3]	0.233		0.994		0.004	
Variance equal test (vs. Naïve) p-value [1,4]	0.875		0.474		0.003	

	SUPRAMAX					
	Next month (short-term)		Next quarter (mid term)		Next calendar (long term)	
	BFA	Naïve	BFA	Naïve	BFA	Naïve
Mean of forecast errors[1]	0.025	0.023	-0.032	0.015	0.099	0.474
p-value, (Mean = 0)	0.161	0.249	0.163	0.665	0.113	0.000
St. Dev[1]	0.140	0.155	0.232	0.339	0.550	0.789
Observations	63	63	102	102	79	79
Start date	Sep-09	Sep-09	Jan-06	Jan-06	Jun-06	Jun-06
Skewness[1]	0.717	0.460	0.479	2.063	1.375	0.921
Kurtosis[1]	3.352	3.060	2.360	11.457	4.653	3.072
Jarque-Bera[1]	5.717	2.232	5.637	376.293	33.871	11.188
p-value	0.057	0.328	0.060	0.000	0.000	0.004
MAPE[1]	11%	12%	20%	24%	38%	65%
RMSE[2]	1,997	2,180	6,148	6,960	11,928	16,858
Theil 1[2]	0.067	0.073	0.117	0.131	0.304	0.363
Theil 2[2]	0.925		0.888		0.836	
Mean equal test (vs. Naïve) p-value [1,3]	0.931		0.251		0.001	
Variance equal test (vs. Naïve) p-value [1,4]	0.574		0.154		0.004	

	HANDYSIZE					
	Next month (short-term)		Next quarter (mid term)		Next calendar (long term)	
	BFA	Naïve	BFA	Naïve	BFA	Naïve
Mean of forecast errors[1]	0.028	0.016	0.001	0.040	0.205	0.556
p-value (Mean = 0)	0.117	0.287	0.981	0.294	0.002	0.000
St. Dev[1]	0.141	0.121	0.239	0.356	0.553	0.804
Observations	63	63	90	90	72	72
Start date	Sep-09	Sep-09	Jan-07	Jan-07	Jan-08	Jan-08
Skewness[1]	0.463	0.276	0.634	2.818	1.488	1.243
Kurtosis[1]	2.462	2.893	2.813	17.750	4.937	3.443
Jarque-Bera[1]	3.012	0.827	6.153	935.048	37.830	19.114
p-value	0.222	0.661	0.046	0.000	0.000	0.000
MAPE[1]	11%	10%	20%	24%	37%	63%
RMSE[2]	1,466	1,303	4,085	4,716	6,851	10,890
Theil 1[2]	0.067	0.059	0.110	0.126	0.280	0.368
Theil 2[2]	1.135		0.873		0.762	
Mean equal test (vs. Naïve) p-value [1,3]	0.614		0.389		0.003	
Variance equal test (vs. Naïve) p-value [1,4]	0.327		0.170		0.017	

[1] *observations as percentage differences*

[2] *observations as absolute differences*

[3] *Anova F-Stat*

[4] *Brown-Forsythe test*

Table 4. Results of Unit Root Tests

		Capesizes (CAPES)		
		Month	Quarter	Calendar
BFAs	ADF	-2.18 (0.213)	1.24 (0.998)	-2.64 (0.088)
	PP	-2.00 (0.285)	1.75 (0.999)	-2.12 (0.237)
Spot	ADF	-2.92 (0.048)	0.88 (0.995)	-0.11 (0.945)
	PP	-2.86 (0.055)	-1.11 (0.708)	-0.09 (0.946)
		Panamax (PMX)		
		Month	Quarter	Calendar
BFAs	ADF	-1.79 (0.379)	0.94 (0.996)	-2.93 (0.046)
	PP	-1.33 (0.612)	0.87 (0.995)	-2.13 (0.232)
Spot	ADF	-1.66 (0.446)	-0.78 (0.819)	-0.13 (0.942)
	PP	-1.46 (0.549)	-0.77 (0.823)	-0.13 (0.942)
		Supramax (SMX)		
		Month	Quarter	Calendar
BFAs	ADF	-1.55 (0.500)	0.98 (0.996)	-2.85 (0.056)
	PP	-1.31 (0.621)	0.81 (0.994)	-2.04 (0.271)
Spot	ADF	-1.58 (0.488)	-0.64 (0.856)	-0.33 (0.915)
	PP	-1.46 (0.547)	-0.59 (0.867)	-0.32 (0.915)
		Handysize (HS)		
		Month	Quarter	Calendar
BFAs	ADF	-1.21 (0.666)	1.41 (0.999)	-1.79 (0.380)
	PP	-1.18 (0.678)	1.07 (0.997)	-2.11 (0.243)
Spot	ADF	-1.33 (0.610)	-1.88 (0.340)	-1.23 (0.659)
	PP	-1.33 (0.610)	-0.56 (0.873)	-1.24 (0.652)

ADF: Augmented Dickey-Fuller

PP: Phillips-Perron

Null Hypo: unit root exists.

t-Stat (p-value)

Table 5. Johansen Cointegration Tests

		CE[a]	Trace Stat[b]	Coint. Coef. and restriction test[c] β_o	β_1	β_o=0 & β_1=1
CAPES	Month	1	31.22	1332.0	-1.0788	0.3939
			(0.001)	(1084.7)	(0.0537)	(0.8212)
	Quarter	1	66.81	30.3	-1.0699	5.6415
			(0.000)	(1813.8)	(0.0349)	(0.0595)
	Calendar	0	11.68			
			(0.478)			
PMX	Month	1	42.85	914.9	-1.0547	3.676
			(0.000)	(435.7)	(0.0289)	(0.1591)
	Quarter	1	37.22	122.1	-1.0007	0.0431
			(0.000)	(840.5)	(0.0309)	(0.9787)
	Calendar	0	12.54			
			(0.401)			
SMX	Month	1	23.35	378.3	-1.0371	1.7666
			(0.018)	(418.5)	(0.0292)	(0.4133)
	Quarter	1	30.92	-569.9	-0.9859	0.6354
			(0.001)	(766.0)	(0.0325)	(0.7278)
	Calendar	0	11.14			
			(0.528)			
HS	Month	1	33.19	535.8	-1.0541	2.1128
			(0.001)	(359.9)	(0.0333)	(0.3477)
	Quarter	1	38.46	-327.3	-0.9877	0.7153
			(0.000)	(420.9)	(0.0255)	(0.6993)
	Calendar	0	12.23			
			(0.428)			

[a] *number of Cointegration Equations*

[b] *Trace Statistic (Probability of accepting nul of no cointegration in parentheses)*

[c] *Normalized cointegrating coefficients at:* $g_t - \beta_1 f_t - \beta_o$, *(standard errors in parentheses)*

test of cointegration restriction LR Stat (probability in parentheses)

Table 6. Cointegration equations in the sub-sample 2008M11 – 2014M12

Panel A: Chow Test for structural break at 2008M11[a]

	CAPES	PMX	SMX	HS
F-Stat	212	154	65	1
Prob Chi-sq	0.00	0.00	0.00	0.33

[a] Null hypothesis: no breaks at specified breakpoint

Panel B: Cointegration equations in Calendar contracts for sub-sample [b]

CAPES	PMX	SMX	HS
16.02	14.72	18.15	n.a.
(0.173)	(0.243)	(0.095)	n.a.

[b] *Trace Statistic (Probability of accepting nul of no cointegration in parentheses)*

Table 7. Directional Accuracy of FFAs vs Naïve

CAPES

	Month		Quarter		Calendar	
	FFAs	Naïve	FFAs	Naïve	FFAs	Naïve
Correct Observ (%)	59%	51%	67%	55%	71%	62%
Binomial distribution [(1)]	0.0650	0.4007	0.0002	0.1380	0.0000	0.0107

PMX

	Month		Quarter		Calendar	
	FFAs	Naïve	FFAs	Naïve	FFAs	Naïve
Correct Observ (%)	62%	52%	55%	50%	68%	56%
Binomial distribution [(1)]	0.0215	0.3073	0.1380	0.4606	0.0003	0.1149

SMX

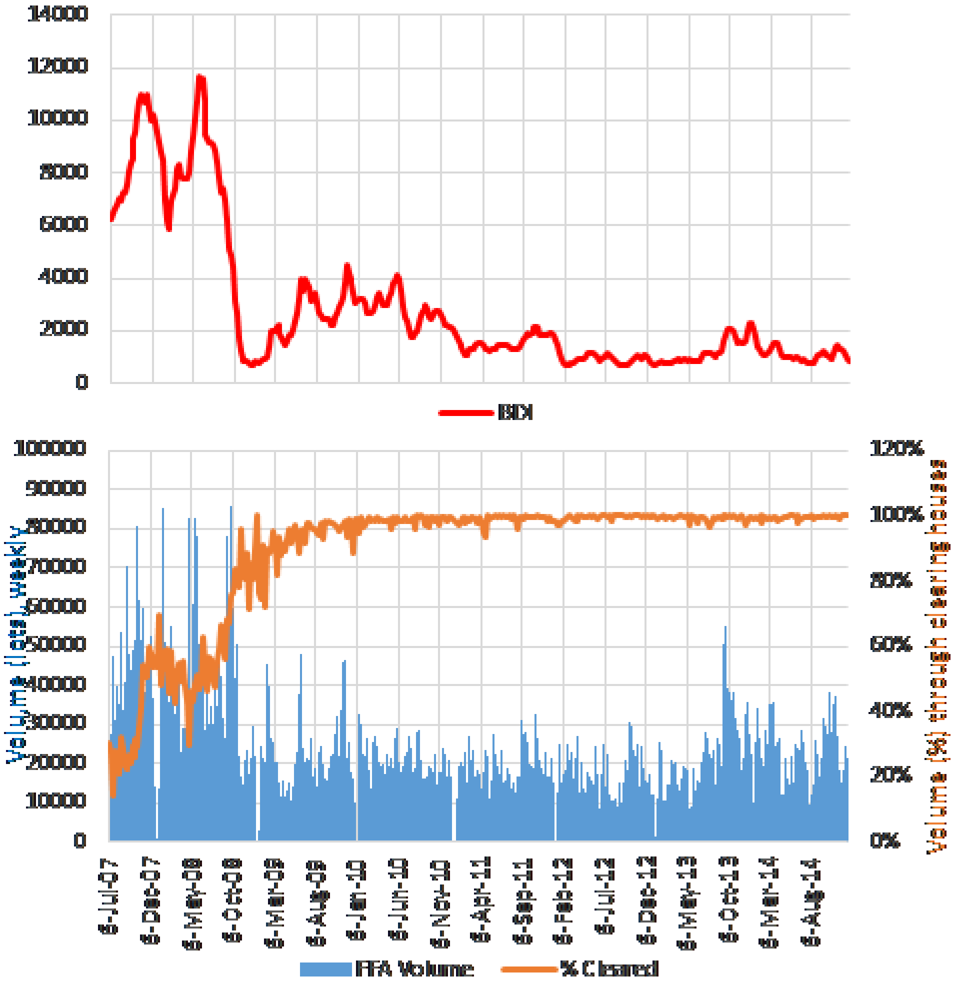
	Month		Quarter		Calendar	
	FFAs	Naïve	FFAs	Naïve	FFAs	Naïve
Correct Observ (%)	56%	48%	47%	47%	65%	47%
Binomial distribution [(1)]	0.1568	0.5993	0.6896	0.6896	0.0033	0.6735

HS

	Month		Quarter		Calendar	
	FFAs	Naïve	FFAs	Naïve	FFAs	Naïve
Correct Observ (%)	63%	41%	54%	44%	82%	61%
Binomial distribution [(1)]	0.0113	0.8963	0.1714	0.8286	0.0000	0.0222

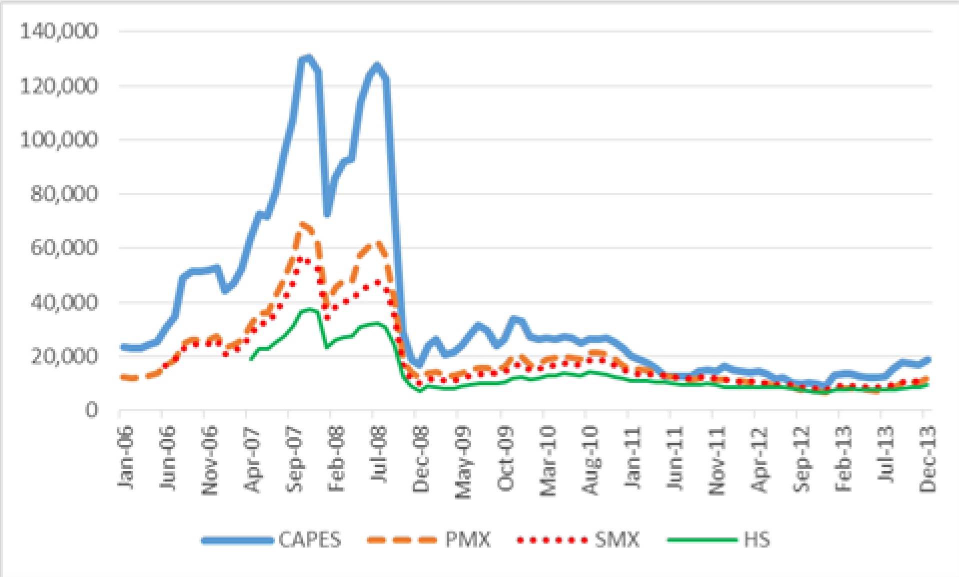
[(1)] *Cumulative binomial probability of values equal or higher of number of success under testing, assuming probability of success = 0.5*

Figure 1. FFA volumes vs BDI development



Source: Baltic Exchange

Figure 2. BFAs Next Calendar Contracts



Source: Baltic Exchange

Figure 3a. CAPES Realizations vs Projections (Calendar)

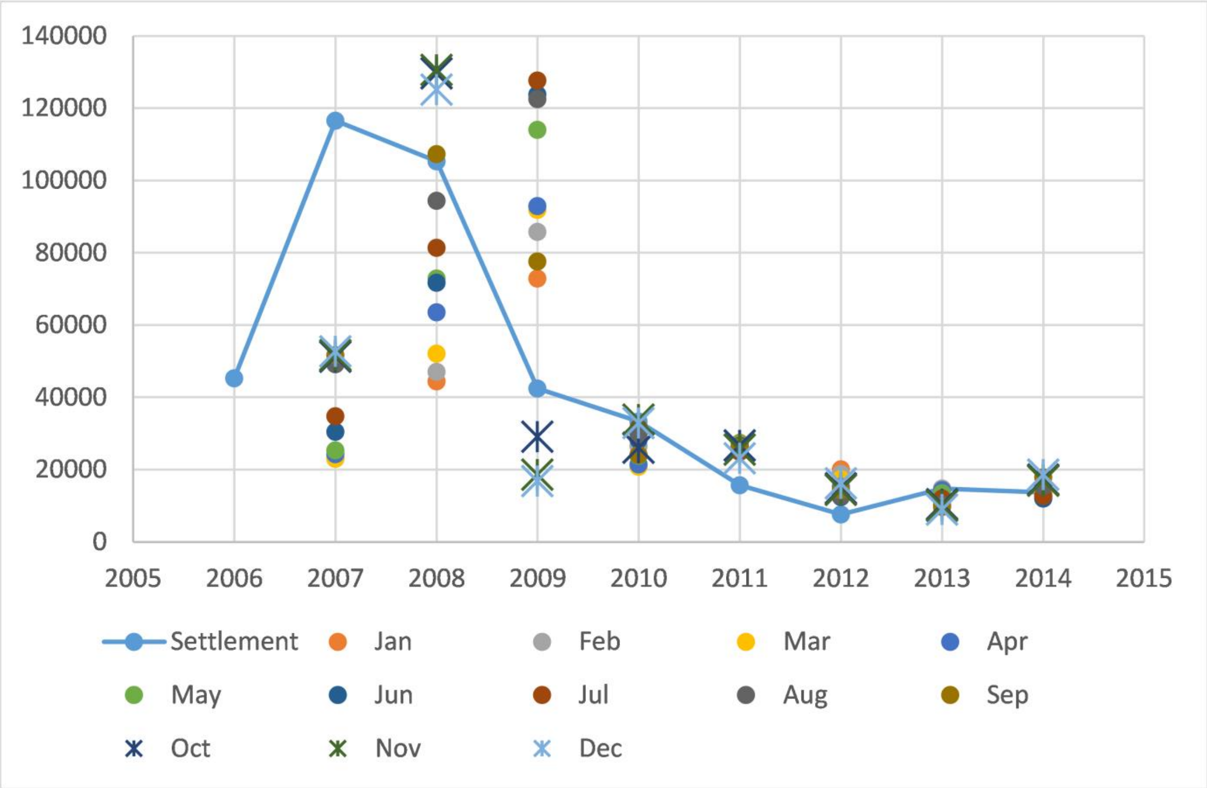


Figure 3b. CAPES Realizations vs Projections (Quarter)

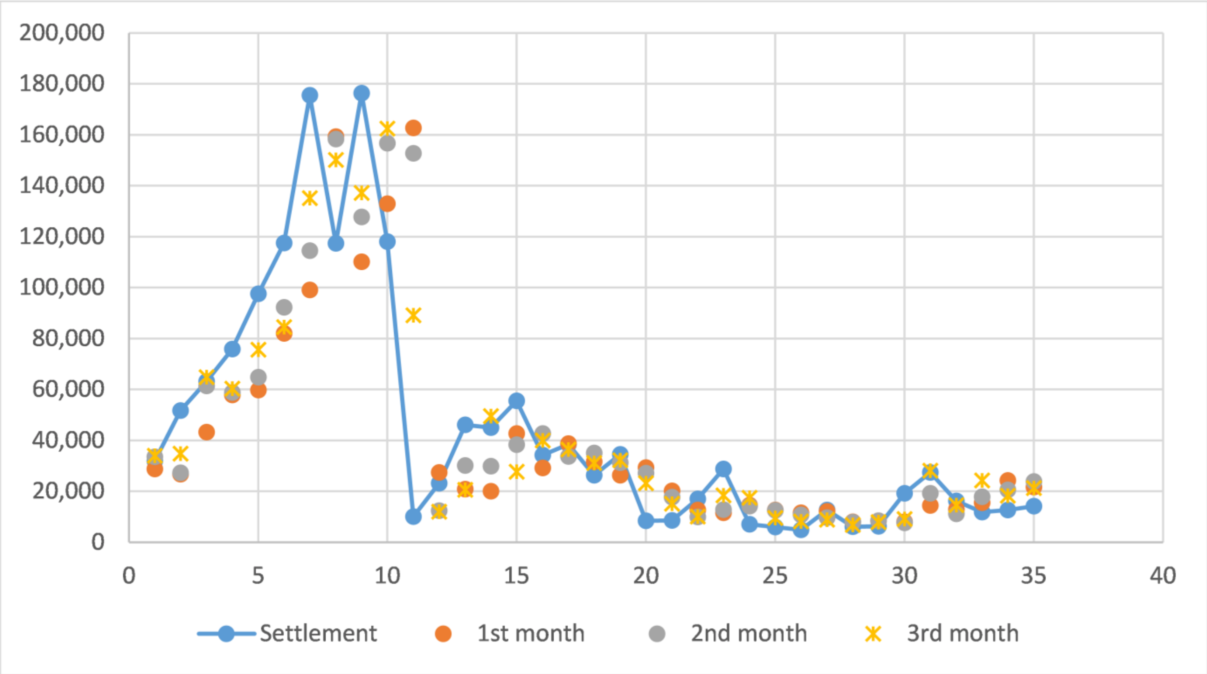


Figure 4a. PANAMAX Realizations vs. Projections (Calendar)

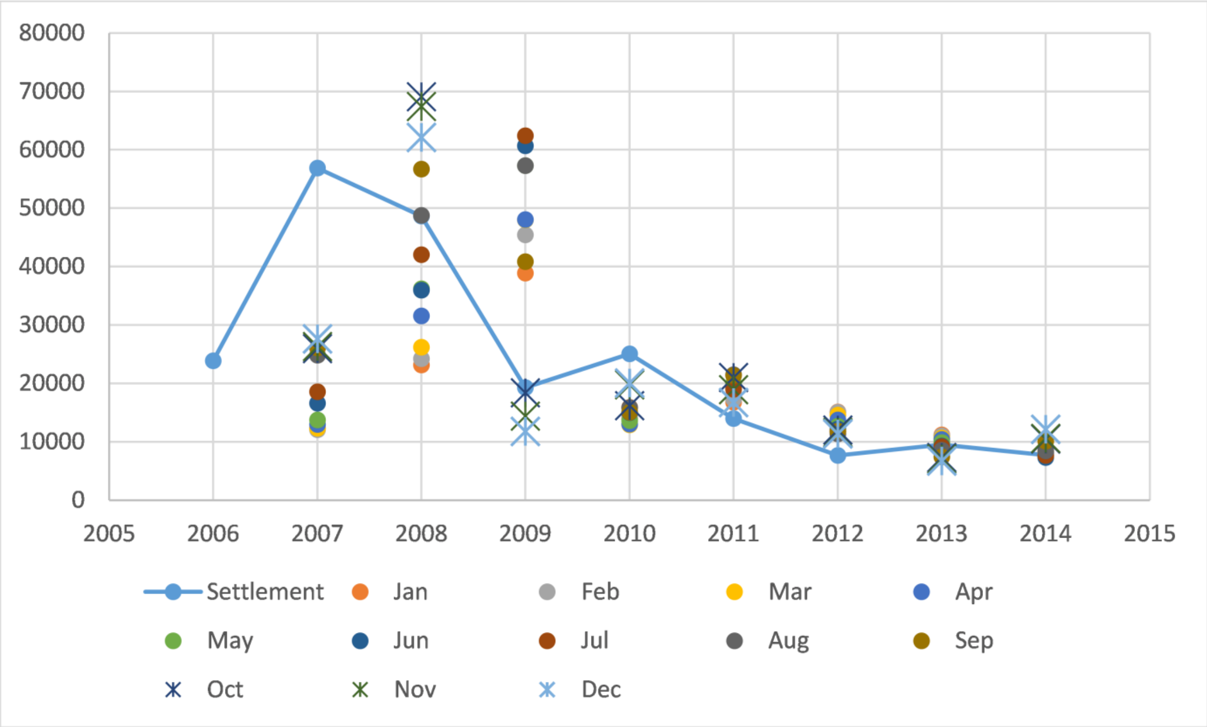


Figure 4b. PANAMAX Realizations vs. Projections (Quarter)

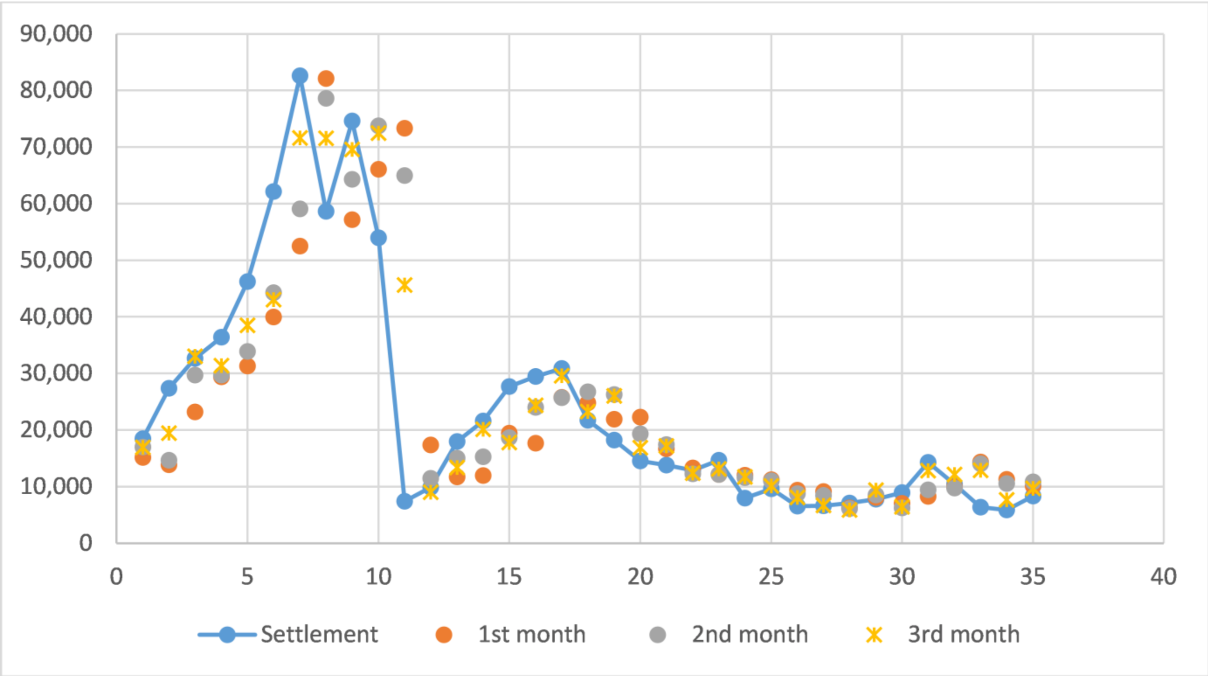


Figure 5a. SUPRAMAX Realizations vs Projections (Calendar)

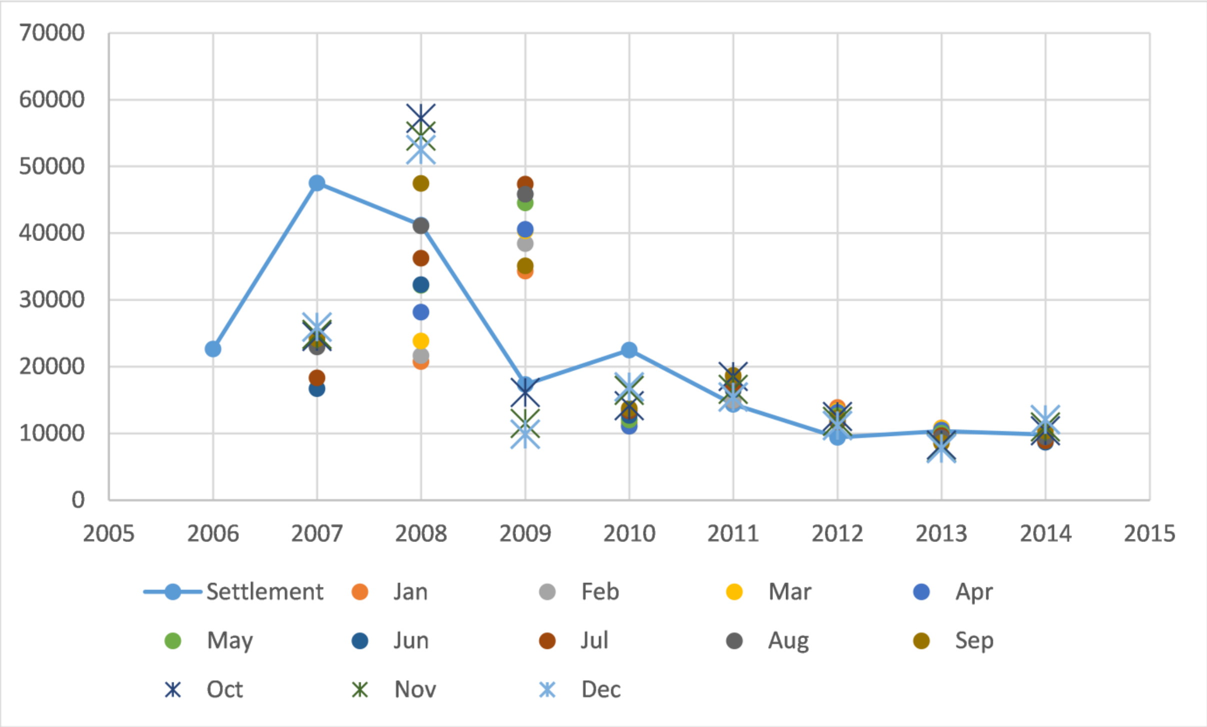


Figure 5b. SUPRAMAX Realizations vs Projections (Quarter)

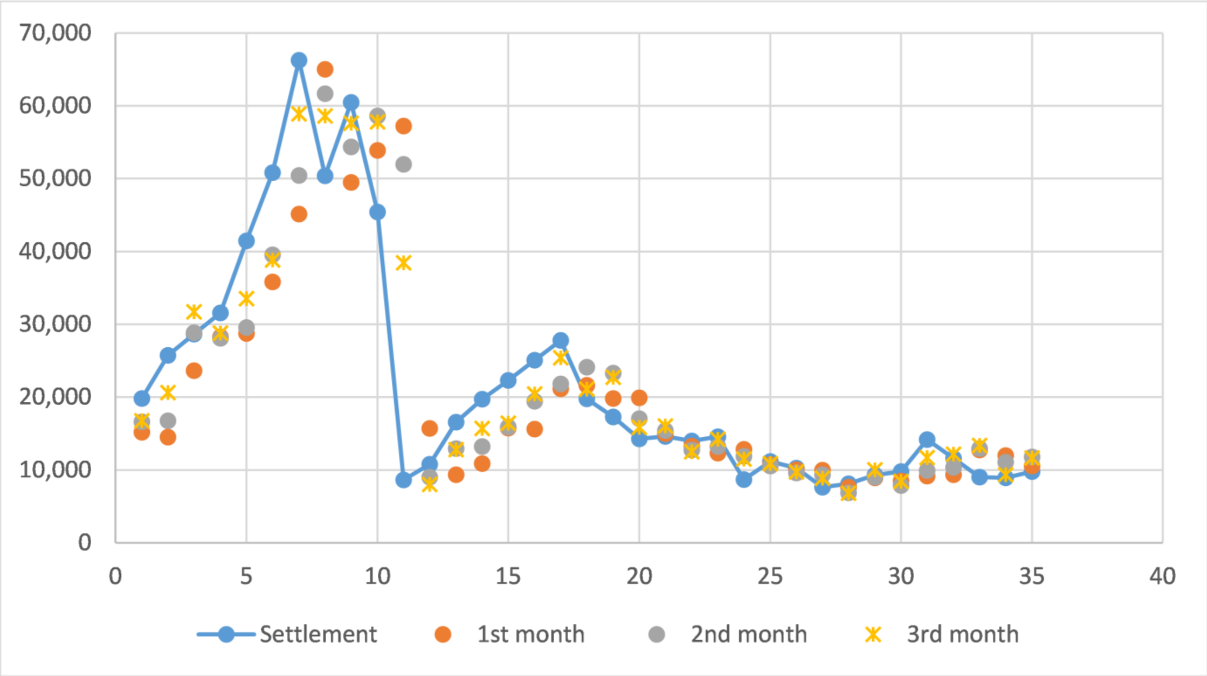


Figure 6a. HANDYSIZE Realizations vs Projections (Calendar)

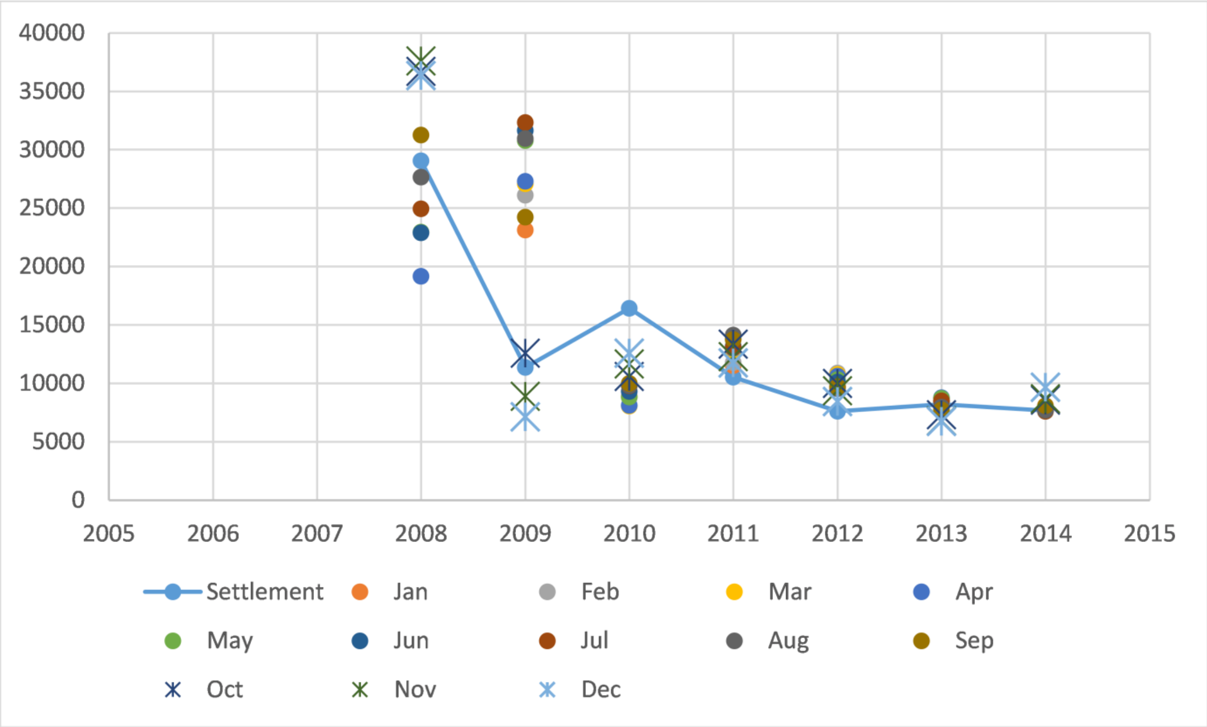
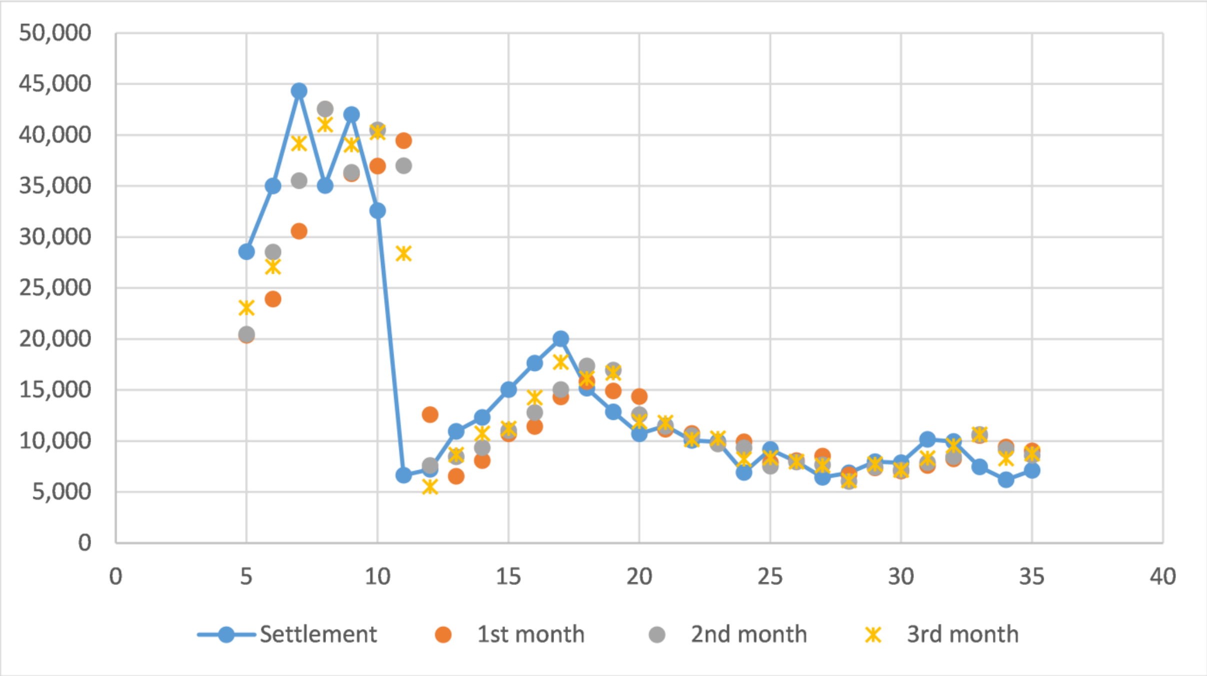


Figure 6b. HANDYSIZE Realizations vs Projections (Quarter)



BANK OF GREECE WORKING PAPERS

210. Malliaropulos, D., and P.M. Migiakis, "The Re-Pricing of Sovereign Risks Following the Global Financial Crisis", July 2016.

211. Asimakopoulos, I., P.K. Avramidis, D. Malliaropulos, N. G. Travlos, "Moral Hazard and Strategic Default: Evidence from Greek Corporate Loans", July 2016.

212. Tavlas, S.G., "New Perspectives on the Great Depression: a Review Essay", September 2016.

213. Papapetrou, E. and P. Tsalaporta, "Inter-Industry Wage Differentials in Greece: Rent-Sharing and Unobserved Heterogeneity Hypotheses", November 2016.

214. Gibson, D.H, S.G. Hall, and G.S. Tavlas, "Self-Fulfilling Dynamics: the Interactions of Sovereign Spreads, Sovereign Ratings and Bank Ratings during the Euro Financial Crisis", November 2016.

215. Baumann, U. and M. Vasardani, "The Slowdown in US Productivity Growth - What Explains it and Will it Persist?", November 2016.

216. Tsionas. G.M., "Alternative Bayesian Compression in Vector Autoregressions and Related Models", November 2016.

217. Tsionas. G.M., "Alternatives to Large VAR, VARMA and Multivariate Stochastic Volatility Models", December 2016.

218. Balfoussia, H. and D. Papageorgiou, "Insights on the Greek Economy from the 3D Macro Model", December 2016.

219. Anastasiou, D., H. Louri and M. Tsionas, "Non-Performing Loans in the Euro Area: are Core-Periphery Banking Markets Fragmented?," December 2016.

220. Charalambakis, E., Y. Dendramis, and E. Tzavalis, "On the Determinants of NPLS: Lessons from Greece", March 2017.

221. Christou, G. and P. Chronis, Markups and fiscal policy: analytical framework and an empirical investigation, April 2017.

222. Papadopoulos, S., P. Stavroulias, T. Sager and E. Baranoff, "A Ternary-State Early Warning System for the European Union", April 2017.

223. Papapetrou, E. and P. Tsalaporta, "Is There a Case for Intergenerational Transmission of Female Labour Force Participation And Educational Attainment? Evidence from Greece during the Crisis, April 2017.

224. Siakoulis V., "Fiscal policy effects on non-performing loan formation", May 2017.

225. Hondroyiannis G. and D. Papaoikonomou, "The Effect of Card Payments on VAT Revenue in Greece", May 2017.

226. Mamatzakis E.C. and A.N. Vu, "The interplay between quantitative easing and risk: the case of the Japanese banking", May 2017.

227. Kosma, T., E. Papapetrou, G. Pavlou, C. Tsochatzi and P. Zioutou, "Labour Market Adjustments and Reforms in Greece During the Crisis: Microeconomic Evidence from the Third Wave of the Wage Dynamics Survey", June 2017.

228. Gibson D.H, and G. Pavlou, “Exporting and Performance: Evidence from Greek Firms”, June 2017.

229. Papaspyrou S. T. “A New Approach to Governance and Integration in EMU for an Optimal Use of Economic Policy Framework - Priority to Financial Union”, June 2017.