BI Norwegian Business School - campus Oslo

GRA 19502

Master Thesis

Component of continuous assessment: Thesis Master of Science

A Study of the Volatility in the Dry Bulk Market

I

Start: 02.03.2017 09.00

Finish: 01.09.2017 12.00

Name: **Maksym Malovichko**

Yu-Ju Chou

Programme:

Master of Science in Business

Major in Finance

Hand-in date:

09.08.2017

Content

Abstract

This thesis aims to measure and analyze the volatility of the dry bulk freight spot market. Empirical research is conducted by using samples of daily observations from 1985 to 2016. We find the return series are stationary and heteroscedastic. We use AR-GARCH type of models and compare different model specifications. We conclude that bigger ships are riskier and their underlying dynamics are more complex. Shocks are very persistent in the dry bulk freight market, but decrease with the vessel size. We find positive news to have higher impact on the volatility, as predicted by the short-term supply-and-demand model. In Capesize market higher risk leads to lower return. When the market gets extremely volatile, our models systematically underestimate the volatility as the vessel size increases.

Acknowledgements

We would like to thank our thesis supervisor Professor Kjell Jørgensen for his guidance and support. He has encouraged us to experiment with new ideas, while keeping our focus on the most important issues. We also appreciate the resources and advice from the whole faculty, who made the process clear and well designed to stimulate our work. We specifically express our gratitude and admiration to the staff at the Baltic Exchange. Without their hard work and great professionalism our research would not be possible. Lastly, we have been lucky to always receive support and motivation from family, friends, and our peers at BI. They made the journey interesting and fulfilling.

1 Introduction

Volume of seaborne trade accounts for about 80% of total world merchandise trade. Two thirds of the seaborne trade volume is dry cargo, raw materials including primarily ore, coal, grains that will be further processed to make all kinds of end products. The freight shippers pay to charter bulk carriers, ships built to transport dry cargo, has long been closely watched by the shipping industry and the financial market as it is perceived as a leading indicator of global economic state. In February 2016, BDI, an index that measures dry cargo freight, fell to its historical low, a time described by some as “the worst market since the Viking age”. Unfortunate shipowners, charterers, and shipping banks either suffered big write-off of their asset value or went into distress. Currently hit hard by the downturn, the dry bulk shipping market is characterized by its high volatility with the example of a 94% dive between May and December 2008. As bulk carriers account for about 43% of world fleet and carry two thirds of the seaborne cargo, a better understanding of the volatility of the freight will not only help the struggling dry bulk industry make future investment decisions and improve risk management, but also provide insight into global economy.

The goal of this thesis is to understand the nature of risk in the dry bulk freight market by properly measuring it and finding its qualitative and quantitative impact. We aim to capture possible dynamics of heteroscedasticity, asymmetric effects, and risk premium. We are also interested in the model performance when the volatility is the highest. In previous studies, low number of observations and data quality were major issues We use time-series data of daily dry bulk freight index, provided by the Baltic Exchange which covers four major types of dry bulk vessels: Capesize, Panamax, Supramax and Handysize. We have more than 20 years of daily observations, which shall help us to analyze the underlying processes more confidently.

We approach the topic as follows. First, we explore the nature and development of dry bulk freight market. Previous empirical studies and theoretical models are thoroughly examined as foundations and inspirations for our research. Second, we use the most recent data to conduct empirical research of the volatility in the dry bulk freight market, based on the models and results from the previous studies. Third, we derive practical implications for market participants and enhance academic studies of this topic.

2 Literature Review

2.1 Overview of the Dry Bulk Market

The dry bulk carriers have unique characteristics and are governed by unique market mechanisms. In the new millennium, two trends stood out in the dry bulk sector. First, it has attracted more attention from players outside shipping who search for new investment class and leading economic indicators. Second, chartering chains have grown longer and more fragile. One single company that fails to perform could trigger a series of disastrous events (Gratsos et al., 2012). The following sections provide a systematic overview of the most important features of this market and their studies.

Bulk carriers are built to transport homogenous dry bulk commodities in large quantities by sea. Five major bulk commodities iron ore, coal, grain, bauxite/alumina, phosphate rock account for about 60% of total dry bulk trade (UNCATD, 2015). Although each vessel has its own specification, for the purpose of conducting analysis, they are usually grouped with other similar vessels by their capacity (tonnage) for carrying cargoes (Stopford, 2009; Alizadeh and Nomikos, 2009; UNCATD, 2015). Table 2.a shows a common way to group different bulk carriers.

Group	Tonnage
Capesize	100,000 dwt plus
Panamax	60,000–99,999 dwt
Handymax	40,000–59,999 dwt
Handysize	10,000–39,999 dwt

Table 2.a: Four vessel groups (UNCATD, 2015, ix)

Each group of vessels has its unique trading advantage depending on the parcel size of the cargo, cargo handling, distance, routes, and ports. When choosing the vessel to hire or build, trade-offs are general made between three factors: economies of scale, the parcel sizes of the available cargoes, port draught and cargo handling facilities (Stopford, 2009; Alizadeh and Nomikos, 2009).

2.1.1 Risks in the Dry Bulk Market

Making shipping investment is risky. In Cullinane's (1995) study on risk and return of investment in drybulk shipping, he referred to Gray's (1987) perspective on major commercial risks faced by shipowners: (1) Interest rate risk, (2) Exchange rate risk, (3) Bunker price risk, (4) Market risk. Out of four, market risk involves factors that could negatively affect the freight rate. It is industry specific and has the most direct impact on the revenues of shipowners. He argued that it is the most important risk for shipowners because uncertainties have a greater impact on revenues than costs. Stopford (2009) observed shipowners' anxiety about daily fluctuations of freight rates and went on to elaborate on "shipping risk" - risk about the return on shipping investment that comes from the cyclical nature of the shipping business. An increase in trade volume would result in a disproportion between supply and demand of shipping capacity and push up freight rate to restore the balance. As a result, shipowners may be tempted to increase fleet size hoping to capture more profit in a good market. In the end, a good market may eventually wind down as the supply of shipping capacity restores the balance. This uncertainty about the future of the shipping market motivates some companies to take the shipping risk and others to transfer the shipping risk. Our study specifically focuses on the shipping risk market participants face in the dry bulk freight market.

2.1.2 The Structure and Dynamics of the Dry Bulk Freight Market

The freight market is a market place where shipowners provide ships for hire and charterers/shippers hire the ships to transport cargoes. When a freight rate is agreed along with other terms on the “charter-party” (contract specifying all the terms) the ship is “fixed”. There are different charter types such as “Voyage Charter”, “Contract of Affreightment”, “Period Charter”, and “Bare boat charter” with different contract execution and risk transfer mechanism to suit the needs of different counterparties (Stopford, 2009). The most common approach to systematically understand the freight market is the supply and demand model that is often used in the commodities market. Table 2.b presents a general supply and demand model for the shipping market.

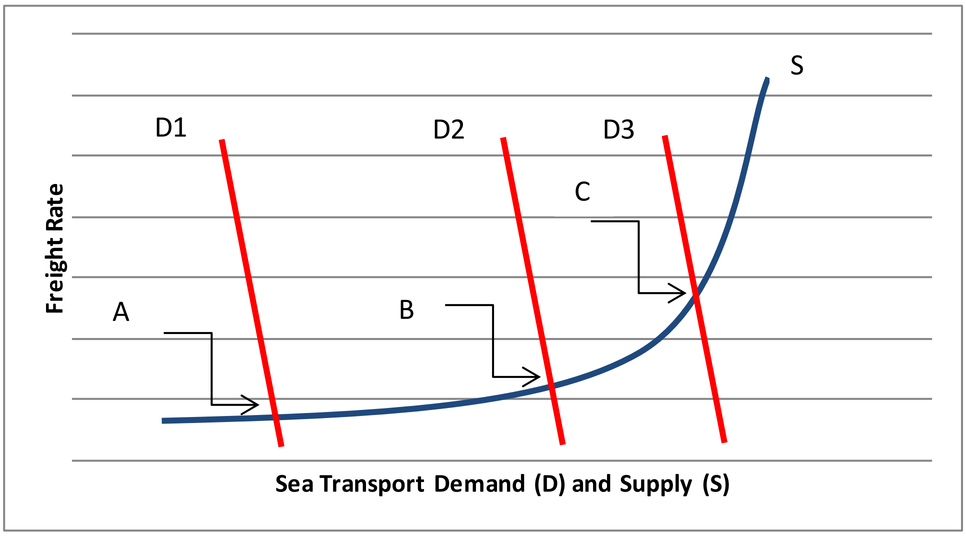
Demand	**Supply**
The world economy	World fleet
Seaborne commodity trades	Fleet productivity
Average haul	Shipbuilding production
Political events	Scrapping and losses
Transport costs	Freight rates

Table 2.b: Ten variables in the shipping market model (Stopford (2009), 136)

Any imbalance between supply and demand feed into the freight market, which acts as a control valve for the money paid from the shipper to the shipowner. This model also demonstrates key characteristics of the shipping market, as the demand is unpredictable and prone to fluctuate in the short run but the supply is slow to catch up. This process of capacity adjustment explains the volatile and cyclical nature of the shipping market (Randers and Göluke, 2007).

2.1.3 Short-Run Supply and Demand Model

Our study focuses on the short-run, where the supply adjustment is constrained by short-term measures, such as lay-up, reactivation, speed adjustment or switching markets. The short-run balance between supply and demand is illustrated in Graph 2.a. Supply function has a ‘J’ shape as new vessels cannot be added to the market immediately to respond to an increased demand, while vessels can slow steam or even be laid-up to respond to a decreased demand. Demand function in the short run is very inelastic as the cost of shipping is usually a very small portion of the cargo's total value and there is hardly any alternative way to transport the dry bulk cargo. When demand curve shifts from D1 to D2, the equilibrium freight rate only rises slightly from A to B as laid-up vessels start to resume operation and vessels operate in full speed. When demand curve shifts from D2 to D3, despite a smaller shift, the freight level rises significantly because more cargoes are fighting for the same number of vessels and the prices are determined by the oldest and least efficient vessels that require the highest costs to operate. Drobetz et al. (2012) argued that this model implies the asymmetric effect of positive shocks having more effect on the conditional volatility of the freight market as the model predicts larger shocks in a good market and smaller shocks in a bad market due to the convexity of the supply curve.



Graph 2.a: Short-run equilibrium model (Stopford (2009), 165)

2.2 Empirical Studies of the Dry Bulk Freight and its Volatility

Given the importance of shipping risk in the dry bulk freight market, it is of great interest for market participants to properly measure it and make reliable forecast. From 1970s researchers started to conduct the empirical analyses of freight rate. From 1990s attention started to shift to the risks behind shipping investments. In addition, advancement of new econometric techniques allowed researchers to capture risks more properly. Volatility has become a common way to measure risks in shipping finance and several studies have been performed to explore the volatility in the dry bulk freight market.

2.2.1 Development of the Dry Bulk Market Modelling

Adland and Strandenes (2007) saw primarily two schools of modeling the freight market. One attempted to capture the supply and demand fundamentals for equilibrium prices (Hawdon, 1978; Beenstock and Vergottis, 1989; Hale and Vanags, 1989; Beenstock and Vergottis 1989, 1993). Their work highlighted the application of structural models in the dry bulk market. They built their model on top of assumptions of rational expectations of freight rates and market efficiency. The freight rates predicted by the model were the expectations of all the market participants. Market efficiency then ensured that ship prices would be adjusted by arbitrageurs to new information known to the market. The other one used univariate stochastic model (Kavussanos, 1996; Kavussanos and Nomikos, 1999; Kavussanos, and Alizadeh, 2002; Adland and Cullinane, 2005). The first school is limited by the difficulty of data collection as the number of variables increases, weak econometric relationships, and its deterministic nature. Hence, we do not consider the structural models.

The second school relies on the assumption that all the information is embedded in the current price. Glen (2006) observed that the reduced form autoregressive model had become the popular choice for empirical research while traditional structural modeling had gone out of fashion. Stationarity testing and co-integration examination have become the launching pad for new research that focused on statistical properties of data. In particular, new statistical models that relaxed the restriction of constant variance have made modeling the time-varying volatility of the dry bulk freight rate increasingly popular. A wide variety of studies have emerged to explore seasonality, term structure, stationarity, forecasting ability of financial derivatives, and conditional heteroscedasticity in the dry bulk market. Consequently, we focus on the reduced form autoregressive models in this study.

2.2.2 Attempts to Capture Non-Stationary Dynamics in the Dry Bulk Market

Stationarity is an essential requirement for generating reliable statistical inferences. Early studies attempted to explore the statistical properties related to stationarity.

Glen and Rogers (1997) examined weekly series of Capesize indices of different key trading routes published from 1989 to 1996 that recorded both spot and time-charter dry bulk freight rates. They found the levels to be all nonstationary under both Augmented Dickey-Fuller test and Phillips-Perron test but their first differences to be stationary. Cointegration between each route was identified and attributed to common external drivers such as industrial production, world trade, seaborne cargo movements, and bunker prices. Tvedt (2003) reviewed prior works on the stationarity of drybulk freight rates and second-hand vessels (Kavussanos, 1996; Glen and Rogers, 1997; Glen 1997; Kavussanos, 1997) which all pointed to a random walk process. He argued that a transformation of indices and freight rates into Japanese yen denomination could yield a different result as Asia accounted for a majority of activities in drybulk commodities trading and shipbuilding. After the transformation of data from 1980s to 1999, the indices and freight rates did become stationary. The BFI index was downward mean reverting potentially implying the dynamic where high freight rates induced an increased new building activities and vessel utilization while low freight rates encouraged vessel lay-up and scrapping.

Motivated by the nonstationarity and deterministic seasonal pattern of macroeconomic variables (Osborn, 1990; Canova and Hansen, 1995), Kavussanos and Alizadeh (2001) used monthly data to search for systematic seasonal patterns in freight rate fluctuations within a year between different group of vessels (Capesize, Panamax and Handysize), different contract durations (spot, 1-year and 3-year time charters), and different market conditions (peaks and troughs). They concluded ARIMA and VAR models were most appropriate to model the series and found deterministic seasonality showing freight rates rose in March and April and dropped in June and July. Freight rates of larger vessels fluctuated more than smaller vessels. Longer contracts had smaller seasonal fluctuations than shorter contracts; seasonal fluctuations were sharper when the market picked up than when it was going down.

Conclusively, the level of dry bulk freight is non-stationary and might exhibit seasonality. Hence, in our research we expect the data to have similar properties, and we may need to transform it to conduct our analyses.

2.2.3 Spot-Forward Relationship and Risk-Premium in the Dry Bulk Freight Market

As in many financial markets, there are spot and forward markets for dry bulk freight. Our study focuses on the spot market. However, in order to get the full picture of the dry bulk market, we examine several important studies which explored the relationship between spot and forward prices.

The theoretical foundation of the relationship between spot and forward prices in the dry bulk freight market was provided by Adland and Cullinane (2005). It has two unique features that make it difficult to establish relationships between spot and forward contract with traditional approach. First, the non-storability character makes the usual cash-and-carry strategy inapplicable. Second, the non-tradability character makes constructing replicating portfolios very difficult. Risk of spot market volatility and liquidity risk could contribute to both a positive and negative risk premium as both shipowners and charterers are risk-averse against future spot freight movements. Without further restrictions of their risk preference and bargaining power, it is difficult to tell the influence. Unemployment risk usually motivates shipowners to offer a lower forward freight rate compared with expected future spot rate to make sure the vessels are chartered. Default risk had the opposite effect as it motivated shipowners to demand a higher forward freight rate to account for the possibility that charterers may walk away from a long-term contract. The risk of transport shortage encourages charterers to pay a higher forward freight rate to ensure their ability to transport future cargoes. Technological/legislative risk prompts charterers to pay a lower forward freight rate to compensate increased costs of trading older vessels. They concluded that the net risk premium should in most cases be negative and time-varying depending on the market conditions.

Kavussanos and Alizadeh's (2002) study on term structure of the dry bulk market began from elaborating on the duration of freight contract. Freight rate of shorter-duration or spot contracts was thought to depend on current supply and demand (Stopford ,2009) while freight rate of longer-duration period contracts was believed to depend on expectations of future short-duration freight rates from rational market participants. This was in line with the expectations hypothesis covered by classic financial economics literature (Campbell and Shiller, 1987, 1991). In reviewing the studies done by Hale and Vanags (1989) and Veenstra (1999) on this topic, they considered the studies inconclusive due to insufficient sample size and inappropriate model formulation. Using the tests proposed by Campbell and Shiller (1987, 1991), monthly data from 1980 to 1997 of contracts matured in one year and three years in different vessel group (Handysize, Panamax, and Capesize), they found negative time-varying risk-premia through EGARCH-M specifications. Defying traditional belief of expectation hypothesis, they provided four arguments for explanation: higher fluctuations in the spot market, unemployment risk, vessel relocation costs, uncertainty over voyage costs.

2.2.4 Conditional Heteroscedasticity Models of Volatility for the Dry Bulk Freight Market

The spot market has continued to be the center of empirical research in the dry bulk freight market. A number of studies have shown that returns were stationary, but exhibit volatility clustering. With the development of GARCH-type models in the 90s, it became possible to model heteroscedastic behavior of volatility. Economic intuitions behind risk properties became clear and quantifiable. Our research follows their methodology with improved data and covers more vessel types.

Pioneer of GARCH-modelling for the dry bulk freight market, Kavussanos (1996), attepmted to capture the time-varying dynamics of volatilities with monthly data of spot freight index of different vessel groups (Handysize, Panamax, Capesize) from

1973 to 1992. He found the GARCH process to be stationary and volatilities in the spot and time-charter freight market to behave differently. Handysize vessels were found to have lower volatilities than Panamax and Capesize vessels and Panamax vessels were found to have lower volatilities than Capesize vessels. He attributed it to the capability of smaller vessels to serve more markets and cargoes that made the demand for them less volatile. A major issue of this study was low data frequency, which was not suitable for GARCH specification.

In the new millennium, with data of higher frequency, Marlow et. al. (2008) used daily dry bulk freight rate index BCI, BPI, and JEHSI published by Baltic Exchange and JE Hyde Shipping Index respectively from 1 March 1999 to 23 December 2005 to study the characteristics of the volatility of Capesize, Panamax, and Handysize type of vessels. GARCH (1,1) models were fitted to the daily return of each index and showed that shock from the previous period had more effect on the current volatility of the smallest vessel Handysize as larger capsize and Panamax. On the other hand, past shocks for Handysize were less persistent. They argued it had to do with the higher flexibility of Handysize that could be diverted more easily to more profitable routes, so the memory of volatility was not as long which was very similar to Kavussanos's (1996) speculation. The GARCH processes of all three series were found non-stationary meaning shocks tended to strengthen.

They suspected that the more complex market conditions of greater changes after 2003 made a simple GARCH (1,1) process unable to capture all the market characteristics. The sample was then divided into two parts from 1 January 2003 and EGARCH (1,1) models were fitted. Volatility of Panamax vessels in both the first and second sample periods were found to be asymmetric and higher on negative shocks. Volatility of Capesize vessels in the first sample period and volatility of Handysize vessels in the second sample period were found to be asymmetric and higher on positive shocks. The magnitude of the shocks had an asymmetric impact on all the series. Larger shocks had higher impact on the conditional volatility, comparatively to smaller shocks. The explanations revolved around vessel availability, changing aggregate demand of commodities, and operators' expectations.

In the following years, researchers started to expand standard GARCH models, based on theoretical foundations of dry bulk. Usually, they added exogenous variables into conditional mean and/or conditional variance equations.

In an attempt to have a more profound understanding of the volatility in the dry bulk freight market, Alizadeh and Nomikos (2011) investigated the relationship between volatility and the term structure. Their theoretical basis came from commodities market as a backwardation (spot price is higher than forward prices) market indicates a temporal urge for the buyer to get hold of the commodity hence paying a higher price when the supply is inelastic. If the dry bulk market had followed the same logic, they expected to find higher volatility in a backwardation market. Using weekly observations from 1992 to 2007, they found higher volatility in the spot contract than 1-year and 3-year time charters contract, and by using an EGARCH-X specification, they found shocks to be persistent and have sign effects, as market participants strengthened the possibility of a downturn by their reaction to bad news. Most importantly, they found much higher volatility in backwardation market and the rate increased as the degree of backwardation increased. This confirmed the theory that the freight rate was highly sensitive when the supply is tight, but when there was excess supply in the market to absorb shocks, the volatility would not change a lot.

Xu et al. (2011) used a two-step model to analyze the relationship between fleet size and volatility of spot and time-charter freight rate of Capesize and Panamax with monthly data from 1973 to 2010. They first generated one-step ahead conditional volatilities by using a GARCH model and had it regressed against the changes of fleet size, freight rates, industrial production, and bunker price. They found nonstationarity in variance under the GARCH process and confirmed previous results in the literature that volatility of both spot and time-charter drybulk freight rates is time-varying and clustering (Kavussanos, 1996; Kavussanos, 2003; Adland and Cullinane, 2005). In addition, fleet size is found to positively affect the volatility in particular the volatility of spot Capesize freight rate, which echoed Kavussanos's (1996) finding.

Drobetz et al. (2012) studied the effect of macroeconomic variables and asymmetric effects on the volatility of Capesize and Panamax freight market. Referring to the supply and demand model proposed by Stopford (2009), they expected to observe a positive asymmetric effect of freight rate changes, meaning that positive shocks have a larger impact on the conditional volatility than negative shocks of the same magnitude. Common indices on the market as proxy for world stock market, oil prices, wheat prices, metal prices, commodity prices, TED spread, and term spread were selected based on the same supply and demand model. Daily returns of indices BCI and BPI from March 1999 to October 2011 were fitted to an EGARCH (1,1) model. They did not find asymmetric effect on the sign, but found it on the size of shocks, which was different from the result of Marlow et. al. (2008). Then EGARCH-X models were fitted and slope of the yield curve, world stock market and wheat prices were found to be significant in the conditional variance equation.

2.3 Primary Takeaways from the Literature Review

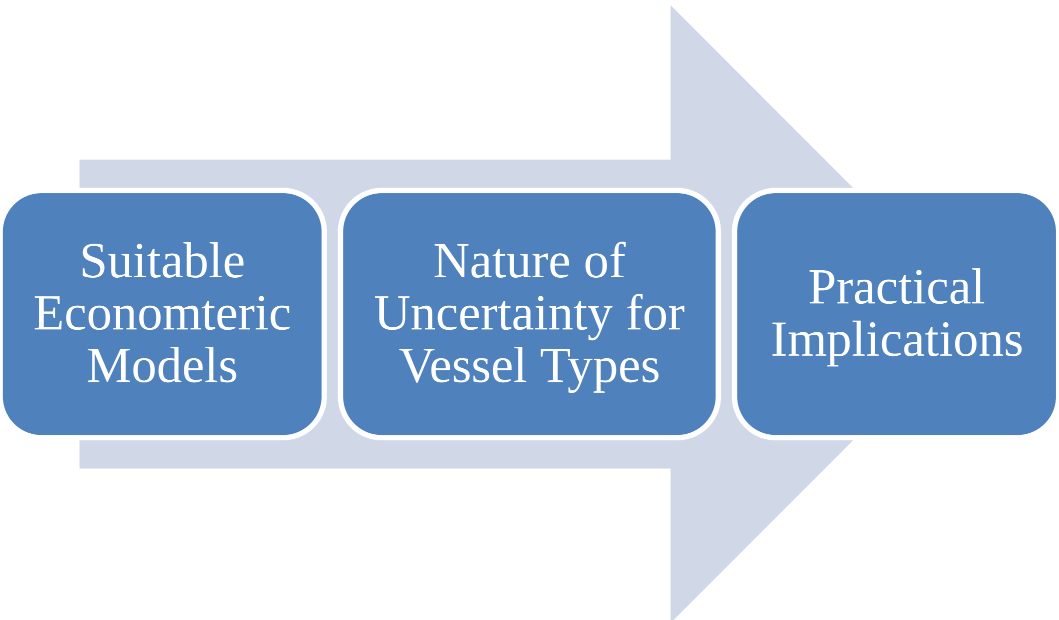
Through GARCH specifications a lot of risk properties in the dry bulk freight market have been discovered and explained. On the other hand, some results are contradictory, which might be due to the inconsistent quality of data. Although efforts have been made to expand the standard GARCH models to better capture the dynamics in the dry bulk market, it is still inconclusive what the best-fitting specification is. Consequently, we decide to put the emphasis of our study on using high quality data, univariate specifications and conducting a comprehensive coverage of all major types of vessels.

3 Research Questions

Previous studies have shed light on ways to model volatility in the dry bulk freight market. Each of them answers specific questions about the nature of risks in the dry bulk freight market. With new data, we aim to conduct a more comprehensive study and answer following questions to generate insights for risk management practice.

1. What is the nature of volatility in the dry bulk freight spot market and how can it be modelled?
2. How accurate are our model forecasts?
3. What are the economic implications for market participants?

By answering the above questions, we hope to achieve the objectives presented in Graph 3.a.



Graph 3.a: Thesis Objectives

4 Methodology

When modeling long-run relationship between variables, it is important to test for non-stationarity. For non-stationary series, shocks do not diminish over time and the standard assumptions for asymptotic analysis do not hold (Brooks, 2014). Fuller (1976) and Dickey and Fuller (1979, 1981) proposed the following Dickey-Fuller (DF) test with the null hypothesis: $H_0: \varphi = 1$ and the alternative hypothesis $H_a: \varphi < 1$ to test the presence of a unit root in the following model:

$$\Delta y_t = \alpha + (\varphi - 1) y_{t-1} + \varepsilon_t$$

The problem is that, under the null hypothesis y_{t-1} is non-stationary so φ does not have a t distribution under a large sample size. They solved the problem by coming up with the distribution of Dickey-Fuller critical values. If the t statistic of $\hat{\varphi}$ is less than the Dickey-Fuller critical values, the null hypothesis that there is a unit root is rejected. To apply the same test for higher order processes, more lags could be included to perform an augmented Dickey-Fuller (ADF) to correct for serial correlations in the error term of the auxiliary regression model. The null hypothesis: $H_0: \varphi = 1$ and alternative hypothesis $H_a: \varphi < 1$ stay the same. The model that includes a drift (α_0) and a trend (T) is specified as follow:

$$\Delta y_t = \alpha_0 + \alpha_1 T + (\varphi - 1) y_{t-1} + \sum_{i=1}^{h} \beta_i \Delta y_{t-i} + \varepsilon_t$$

Phillips and Perron (1988) proposed an alternative test (Phillips-Perron test) to the ADF test. The biggest difference is that lags selection is no longer needed as the auxiliary regression model is simply the same as the one in the (DF) test. They address the serial correlations by transforming the t statistic nonparametrically and the resulted statistic is compared with the Dickey-Fuller critical values.

To capture the dynamics in the data generating process of a random variable, structural models require identification of variables based on underlying theories. By contrast, univariate time series models rely on its past values and past errors that are empirical relevant to the observed samples. In modeling time series data, there are two major lines of models. The first line is Autoregressive (AR) models that let the current observation of a random variable depend on the past observations. The second line is Moving Average (MA) models that have not only shocks in current period but also shocks in the previous period to affect the current observation. AMRA type of models combine the two and make the form compact to keep the number of parameters small. An ARMA (p,q) with p orders of autoregressive terms and q orders of moving average terms could be specified as follows:

$$\mathrm{y_t} = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

Box and Jenkins (1976) proposed a three-step method to correctly fit an ARMA model to time series data in a systematic way. The first step is identification with the goal to first have the series stationary by possibly differencing the series and second find the correct order of AR and MA terms by assessing the autocorrelation and partial autocorrelation plots. The partial autocorrelation function, or PACF measures the correlation between an observation T periods ago and the current observation, after controlling for observations at intermediate lags. The second step is the estimation of parameters. Common estimation techniques include least squares, non-linear least squares, and maximum likelihood. The third step is model checking. In additional to the specified model, a larger model could be fitted and check the significance of the new coefficients. If they are found to be insignificant, the larger model should not be chosen. If the model is correctly specified, the residuals of the model should look like a random drawing from a white noise process. This could be checked by residual plots and statistical tests such as Ljung–Box test and Durbin–Watson test to see if any linear dependence between the residuals is present.

Traditional statistic models such as classical linear regression model require the data series to have constant variance or the estimated parameters would be inefficient. However, it is uncertain if the variance of financial time series is constant. The presence of volatility clustering in some financial time series, meaning large (small) changes tend to be followed by large (small) changes (Mandelbrot, 1963), indicates that the market is more volatile in some periods of time than others (Brooks, 2014). Engle (1982) developed the class of Autoregressive Conditional Heteroscedasticity (ARCH) model which models risks by allowing the conditional variance (h_t) of the time series to depend on the previous values of squared error ($\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-p}$). The conditional mean of y_t is determined by $x_t\,\beta$, a linear combination of lagged variables (x_t) that could take almost any forms and is included in the information set φ_{t-1} with a vector of parameters β. An ARCH (p) model can be written as:

$$y_t\,|\varphi_{t-1} \sim N(x_t\,\beta, h_t)$$
$$h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \dots + \alpha_p\epsilon_{t-p}^2$$
$$\epsilon_t = y_t - x_t\,\beta$$

The model has the desirable econometric application where the previous forecast errors are used to predict the next forecast variance. Most importantly, if the observed time-varying volatility or volatility clustering could be explained by an ARCH process, the researcher could continue to operate on the assumption of unconditional stationarity.

To detect the presence of ARCH effects in the data, Engle (1982) developed an LM test for ARCH effects. The null hypothesis assumes that the series of residuals has no conditional heteroscedasticity against the alternative hypothesis of series being subject to ARCH process. The ARCH model has the following specification:

$$u_t^2 = \alpha_0 + \alpha_1 u_{2t-1} + \dots + \alpha_p u_{2t-p} + \epsilon_t$$

Where at least one of $\alpha_i \neq 0$. The test statistic is a Lagrange multiplier statistic TR^2, where T is the sample size, R^2 is the coefficient of determination for the fitting ARCH(p) model with p lags. Under the null hypothesis, the asymptotic distribution of the test statistic is chi-square with p degrees of freedom.

One question to consider when applying ARCH models is the number of lagged errors. To capture the dependence in the conditional variance, the number of lagged errors can be very large which requires more coefficients to be estimated. As the squared errors are always positive, when the coefficient is negative the whole term is rendered negative. Problems quickly arise when there is one large shock accompanied by one negative coefficients which could make the conditional variance negative. By definition, conditional variance h_t should always be positive. A negative conditional variance is meaningless. Aware of this limitation and relatively arbitrary selection of lag structure in ARCH models, Bollerslev (1986) extended ARCH model in a way similar to extending AR process to ARMA process of time series data. The result is a generalized ARCH (GARCH) model. A GARCH (p,q) model can be written as:

$$y_t \,|\varphi_{t-1} \sim N(x_t\,\beta, h_t)$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}$$

$$\epsilon_t = y_t - x_t\,\beta$$

Letting past conditional variance enter the adaptive learning mechanism, GARCH models essentially have the one-period-ahead conditional variance determined by a weighted average of its long-term average value (α_0), past errors ($\sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2$), and past conditional variances ($\sum_{i=1}^{p} \beta_i h_{t-i}$). As a GARCH (1,1) model can be proven to be a restricted infinite order ARCH model, it is parsimonious and more unlikely to violate the non-negativity constraints for the conditional variance. The conditional variance equation of a GARCH (1,1) model can be written as:

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta h_{t-1}$$

Simple economic interpretation can be drawn by comparing the coefficients in this specification. If the past error coefficient α_1 is large compared with past conditional variance coefficient β, the volatility reacts very quickly to recent shocks and is spikier. If the past conditional variance coefficient β is large compared with the past error coefficient α_1, past shocks have a persistent effect on volatility and dies our slowly. As long as $\alpha_1 + \beta < 1$, the unconditional variance of ϵ_t can be derived and shown as:

$$\mathrm{Var}(\epsilon_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta)}$$

If $\alpha_1 + \beta \geq 1$, the unconditional variance is undefined and the conditional variance would not converge to its long-term (Brooks, 2014)

Following its discovery, GARCH models have been widely used in financial science and various extensions have been developed in order to fit GARCH model for different markets and conditions. One of them, GARCH in Mean (GARCH-M) has a particular interest for financial markets in general and for dry bulk market in particular. It was first developed, based on the ARCH-M model by Engle, Lilien and Robins (1987), which was specified in the following way:

$$y_{t=} \beta + \delta h_t + \epsilon_t$$
$$h_t = \gamma + a \sum w_t \epsilon_{t-t}^2$$
$$\epsilon_t \sim N(0, h_t^2)$$

Adding the conditional standard deviation term to the mean equation make the parameters in both mean and variance equation must be estimated jointly to be asymptotically efficient. This model has an intuitive economic logic: it allows to model for risk-return trade-off and gives a possibility to analyze the nature of risk premium for better market understanding. GARCH-M is identical to a usual GARCH model, presented before, with an addition of any conditional variance term (variance, standard deviation, log of variance etc) into the conditional mean equation. For example, in Brooks (2014) it is specified as:

$$\gamma_t = \mu + \delta\sigma_{t-1} + u_t, u_{t,} \sim (0, \sigma_t^2)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2$$

When δ is positive (negative), it implies that conditional return increases (decreases), when risk, represented by conditional variance, increases.

By squaring the error terms, the specification of GARCH model assumes that both positive and negative shocks have a symmetric effect on the volatility. This is a limiting factor for the practical application of GARCH as it is observed that in financial series bad news causes volatility to rise more than good news does. Plenty of research done in the past has confirmed the phenomenon in both the equity and bond market (Black, 1976; Nelson, 1991; Engle and Ng, 1993; Koutmos and Booth, 1995; Bekaert and Wu, 2000; De Goeij and Marquering, 2006). In the equity market, it is commonly theorized that the "Leverage Effect" causes the volatility to increase more on bad news because of the increased leverage when equity value drops that makes investors consider their investment riskier. Nelson (1991) found the exponential GARCH model (EGARCH) that solves the issue the above issue and non-negativity constraints imposed in the GARCH model. One specification of the conditional variance equation of EGARCH (1,1) model assuming conditionally normal errors is as follow:

$$\ln(\mathrm{h}) = \omega + \beta \ln(\mathrm{h}_{t-1}) + \gamma \frac{\epsilon_{t-1}}{\sqrt{\mathrm{h}_{t-1}}} + \alpha \left[\frac{|\epsilon_{t-1}|}{\sqrt{\mathrm{h}_{t-1}}} - \sqrt{\frac{2}{\pi}} \right]$$

This model is no longer subject to violations of non-negativity from negative parameters because h remains positive when the logarithm of h is negative. If γ is significant and not equal to zero, the asymmetric effect is present. When γ is negative, negative shocks have a stronger effect on the volatility. When γ is positive, positive shocks have a stronger effect on the volatility.

A GARCH process has a requirement of conditional normality to have consistent parameters via log-likelihood maximization. However, this assumption is often violated when analyzing financial data. Hence, the estimated coefficient would be inconsistent. To deal with this problem Bollershev and Wooldridge (1992) developed Quasi-Maximum Likelihood Estimation method for GARCH processes. It includes robust inference procedures which enable estimation of asymptotic standard errors, which are valid and consistent under nonnormality. Therefore, we use QMLE in case the assumption of normality is violated.

Goodness of fit for the mean equation is crucially important for the GARCH estimation process, since its shocks directly impact the conditional volatility estimation through the ARCH term. R^2 is traditionally used a scaled goodness of fit statistic. The goodness of fit statistic is given by the ratio of the explained sum of squares to the total sum of squares (Brooks 2014):

$$R^2 = \frac{ESS}{TSS}$$

R^2 is a measure of what part of the dependent variable is explained by the specified model and lays between 0 and 100%. However, its measure has a significant drawback: it never falls with the addition of new explanatory variables, as any added information would not decrease the explained sum of squares. Therefore, we use Adjusted $R^{2,}$ which is defined as:

$$\widetilde{\mathrm{R}}^2 = 1 - \left[\frac{T-1}{T-k}\,(1 - R^2)\right]$$

Where T is the total number of observations and k is the total number of variables. Adjusted R^2 would only increase when the value of added information by a new variable is higher than offsetting amount of lost degrees of freedom. Hence, by using Adjusted R^2 it is possible to determine parsimonious mean equation.

Pagan and Schwert (1990) used news impact curve to present how new information is incorporated into conditional volatility estimates. X-axis takes the one-period lagged shock and Y-axis takes the estimated conditional variance of the next period. The standard GARCH model always produces a symmetric news impact curve that centered around $\epsilon_{t-1} = 0$ because positive shocks and negative shocks of the same magnitude have the same effect on volatility. Larger shocks produce more volatility at a rate proportional to the squared size of the shock. As a result, standard GARCH models run the risk of underpredicting the volatility after bad news and overpredicting the volatility after good news. In addition, it also underpredicts volatility when larger shocks cause volatility to rise more than its quadratic functions allows. Engle and Ng (1993) proposed a diagnostic test on the sign bias and size bias to address the above issues. The normalized squared residuals are regressed against dummy variables of sign bias and size bias and the lagged. If those residuals are found significant, the original variance model is deemed misspecified. The regression model of the test takes the following form:

$$\hat{v}_t{}^2 = \phi_0 + \phi_1 S_{t-1}^- + \phi_2 S_{t-1}^- v_{t-1} + \phi_2 S_{t-1}^+ v_{t-1} + e_t$$

S_{t-1}^- takes 1 if the shock is negative otherwise it takes 0. S_{t-1}^+ takes 1 if the shock is positive otherwise it takes 0. Significance of $\phi_1$would imply sign bias. Significance of ϕ_2 or ϕ_3 would imply size bias. Joint test statistic TR^2 asymptotically follows chi-square distribution where R^2 is the squared multiple correlation of the regression model and T is the number of observations.

When selecting the correct model to fit the data, three information criteria AIC, SBIC, and HQIC proposed by Akaike's (1974), Schwarz's (1978), and Hannan and Quinn (1979) respectively are often used to compare the fit between models. All three information criteria encompass residual sum of squares (RSS), and penalty terms for the loss of degrees of freedom from adding more parameters.

$$\mathrm{AIC} = \ln\left(\hat{\sigma}^2\right) + \frac{2\mathrm{k}}{\mathrm{T}}$$

$$\mathrm{SBIC} = \ln\left(\hat{\sigma}^2\right) + \frac{\mathrm{k}}{\mathrm{T}} \ln(\mathrm{T})$$

$$\mathrm{HQIC} = \ln\left(\hat{\sigma}^2\right) + \frac{2\mathrm{k}}{\mathrm{T}} \ln(\ln(\mathrm{T}))$$

T is the total number of observations. $\hat{\sigma}^2$ is the residual sum of squares divided by T. k is the total number of parameters estimated in the model. The lower the information criteria are, the lower the variance is unexplained by the model, indicating a model of better quality.

Another important diagnostic is the presence of serial correlation in the residuals of the estimated models. In case it is strong and significant, the estimated coefficients would be inefficient. Durbin and Watson (1951) developed a test statistic for serial correlation of first order in the residuals:

$$DW = \frac{\sum_{t=2}^{T}(\widehat{u_t} - \widehat{u_{t-1}})^2}{\sum_{t=2}^{T}\widehat{u_t}^2}$$

DW statistic ϵ [0;4]. When DW equals 2, there is no evidence of serial correlation in the residuals, when it is less (more) than 2, it indicates positive (negative) autocorrelation.

5 Introduction of the Indices

Our research depends on the indices published daily by the Baltic Exchange Ltd (Baltic Exchange) as the proxy for the dry bulk spot freight level. It is therefore essential to have a better understanding of their origin, rationale, and computation.

Headquartered in London, Baltic Exchange is a membership organization with around 600 members in the global shipping market subscribing for access to their freight market information. Its roots date back to 18th century being a collection of "coffeehouses" where merchants exchanged information about cargoes and ships. In the following decades a reliable marketplace for ship and cargoes started to take shape and a centralized exchange was formed. As shipping contracts are private transactions between only the involved partiers without reporting requirements, it is very difficult to gauge the market level from outside. Sitting in the middle of the marketplace and next to a network of shipbrokers, Baltic Exchange introduced the first index Baltic Freight Index in 1985 (replaced by Baltic Dry Index (BDI) in 1999) in an effort to provide the freight market a transparent benchmark of prevailing freight rate that cuts negotiation hassles and lubricate transactions.

To ensure the accuracy and accountability of the published indices, the governing body Baltic Exchange Council oversees the process of indices determination by following the Principles for Financial Benchmarks of International Organization of Securities Commissions. The representatives, including a chairman, are nominated by the members, vetted by the incumbent Baltic Exchange Council, and approved by the board of Baltic Exchange. Each representative represents a segment of the shipping market with at least one dry bulk, wet bulk, and a shipping derivatives broker. In most cases, staff of the Baltic Exchange are forbidden to make direct shipping investment to avoid any conflict of interest.

The indices are calculated by aggregating prevailing market spot freight rate of major shipping routes on different weighting to account for geographic balance. Different types of ships carry different cargoes and trade on different routes. The freight rate is reported in either voyage charter basis (reported as $/mt from one port to another) or time charter basis (reported as $/day) depending on the route. When selecting routes, a steady and significant volume of trade, more transparent and competitive fixtures, and more standard terms are the main criteria. A group of panelists are appointed by Baltic Exchange to submit daily appraisal of prevailing market freight rate. Each panelist is a professional shipbroker which is a member of Baltic Exchange, actively engages in the freight market, does not trade as a principal, and has no sole dependence on any principals for its business.

The five indices used in our study are Baltic Dry Index (BDI), Baltic Capesize Index (BCI), Baltic Panamax Index (BPI), Baltic Supramax Index (BSI) and Baltic Handysize Index (BHSI). Each of them except BDI captures the spot freight level of one individual type of dry bulk carriers. Detailed conditions are set for each type of vessel as the basis of what a standard fixture should be. In reality, day-to-day fixtures in the dry bulk market have so many variables in terms of vessels, routes, and contract terms that a deviation from those set conditions is almost always expected. Therefore, the panelists are asked to exercise their professional judgment when interpreting market information and make appropriate premium/discount adjustment to arrive at the final reported freight rate. One thing to note is that the spot freight rate of each route is multiplied by a multiplier before going into the index. The multiplier serves two purposes. First, the weighting of the route is taken into account. Second, when a new route is added it smoothens the index to avoid a dramatic change of the level of the index. By extracting information from Guide to Market Benchmarks Version 3.1 published in November 2016 by Baltic Exchange, we examine the route composition and its weight in the following section from the largest type of vessel: Capesize to the smallest: Handysize. We also conducted a phone interview with staff at Baltic Exchange for more detailed questions and interpretations. BDI is introduced after the other four indices as it includes elements in all four indices.

5.1 Baltic Capesize Index

Vessel Description (the *routes):

• 180,000mt dwt on 18.2m SSW draft
• Max age 10 yrs
• LOA 290m, beam 45m, TPC 121
• 198,000cbm grain
• 14 knots laden/15 knots ballast on 62mt fuel oil (380cst), no diesel at sea

C2 (5%) Tubarao to Rotterdam. 160,000lt iron ore, 10% more or less in owner's option, free in and out. Laydays/cancelling 20/35 days from index date. 6 days, Sundays + holidays included all purposes. 6 hrs turn time at loading port, 6 hrs turn time at discharge port, 0.5% in lieu of weighing. Freight based on long tons. Age max 18 yrs. 3.75% total commission.
C3 (15%) Tubarao to Qingdao. 160,000mt or 170,000mt iron ore, 10% more or less in owner's option, free in and out. Laydays/cancelling 20/35 days from index date. Scale load/30,000mt Sundays + holidays included discharge. 6 hrs turn time at loading port, 24 hrs turn time at discharge port. Age max 18 yrs. 3.75% total commission.
C4 (5%) Richards Bay to Rotterdam. 150,000mt coal, 10% more or less in owner's option, free in and out, trimmed. Laydays/cancelling 25/40 days from index date. Scale load/25,000mt Sundays + holidays included discharge. 18 hrs turn time at loading port, 12 hrs turn time at discharge port. Age max 15 yrs. 3.75% total commission.
C5 (15%) West Australia to Qingdao. 160,000mt or 170,000mt iron ore, 10% more or less in owner's option, free in and out. Laydays/cancelling 10/20 days from index date. Scale load/30,000mt Sundays + holidays included discharge. 6 hrs turn time at loading port, 24 hrs turn time at discharge port. Age max 18 yrs. 3.75% total commission.
C7 (5%) Bolivar to Rotterdam. 150,000mt coal, 10% more or less in owner's option, free in and out, trimmed. Laydays/cancelling 20/35 days from index date. 50,000mt Sundays + holidays included load, 25,000mt Sundays + holidays included discharge. 12 hrs turn time at loading port, 12 hrs turn time at discharge port. Age max 15 yrs. 3.75% total commission.
C8_14* (5%) Delivery Gibraltar-Hamburg range, laydays/cancelling 3/10 days from index date, transatlantic round voyage, redelivery Gibraltar-Hamburg range, duration 30-45 days. 5% total commission.

C9_14* (7.5%) Delivery Amsterdam-Rotterdam-Antwerp range or passing Passero, laydays/cancelling 3/10 days from index date, redelivery China-Japan range, duration about 65 days. 5% total commission.
C10_14* (15%) Delivery China-Japan range, laydays/cancelling 3/10 days from index date, redelivery China-Japan range, duration 30-40 days. 5% total commission.
C14_14* (15%) Delivery Qingdao spot or retroactive up to a maximum 15 days after sailing from Qingdao, round voyage via Brazil, redelivery China-Japan range, duration 80-90 days. 5% total commission.
C15 (5%) Richards Bay to Fangcheng. 160,000mt coal, 10% more or less in owner's option, free in and out, trimmed, scale load / 30,000mt Sundays + holidays included discharge. 18 hrs turn time at loading port, 24 hrs turn time at discharge port. Laydays/cancelling 25/35 days from index date. Age max 15 yrs. 5% total commission.
C16* (7.5%) Delivery North China-South Japan range, 3-10 days from index date for a trip via Australia or Indonesia or US west coast or South Africa or Brazil, redelivery UK-Cont-Med within Skaw-Passero range, duration to be adjusted to 65 days. 5% total commission.

Table 5.a: Route Composition of Baltic Capesize Index

5.2 Baltic Panamax Index

Vessel Description:

- 74,000mt dwt on 13.95m SSW draft
- Max age 12 yrs
- LOA 225m, beam 32.2m
- 89,000 cbm grain
- 14 knots on 32mt fuel oil (380cst) laden/28mt fuel oil (380cst) ballast, no diesel at sea.

P1A_03 (25%) Delivery Skaw-Gibraltar range, loading 15-20 days from index date, transatlantic round voyage, including east coast South America, redelivery Skaw-Gibraltar range, duration 45-60 days. Cargo basis grain, ore, coal or similar bulk harmless cargo. 3.75% total commission.
P2A _03 (25%) Delivery Skaw-Gibraltar range, loading 15-20 days from index date, for a trip via east coast South America, US Gulf or US east coast to Asia, redelivery Taiwan-Japan range, duration 60-65 days. Cargo basis grain, ore, coal or similar bulk harmless cargo. 3.75% total commission.
P3A _03 (25%) Delivery Japan-South Korea range, loading 15-20 days from index date, transpacific round voyage, either via Australia or Pacific (not including short rounds such as Vostochny to Japan), redelivery Japan-South Korea range, duration 35-50 days. Cargo basis grain, ore, coal or similar bulk harmless cargo. 3.75% total commission.
P4 _03 (25%) Delivery Japan-South Korea range, loading 15-20 days from index date, for a trip via US west coast- British Columbia range or Australia, redelivery Skaw-Passero range, duration 50-60 days. Cargo basis grain, petroleum coke, coal or similar bulk harmless cargo. 3.75% total commission.

Table 5.b: Route Composition of Baltic Panamax Index

5.3 Baltic Supramax Index

Vessel Description:

• 52,454mt dwt on 12.02m SSW draft
• Max age 15 yrs
• LOA 189.99m, beam 32.26m
• 67,756cbm grain, 65,600cbm bale
• 5 holds, 5 hatches
• 4 x 30mt cranes with 12cbm grabs
• 14 knots laden/14.5 knots ballast on 30mt fuel oil (380cst), no diesel at sea

S1A (12.5%) Delivery Antwerp-Skaw range, laydays/cancelling 5/10 days from index date, redelivery Singapore-Japan range (including China), duration 60-65 days. 5% total commission.
S1B (12.5%) Delivery passing Canakkale, laydays/cancelling 5/10 days from index date, redelivery Singapore-Japan range (including China), duration 50-55 days. 5% total commission.
S2 (25%) Delivery South Korea-Japan range, laydays/cancelling 5/10 days from index date, for an Australian or transpacific round voyage, redelivery South Korea-Japan range, duration 35-40 days. 5% total commission.
S3 (25%) Delivery South Korea-Japan range, laydays/cancelling 5/10 days from index date, redelivery Gibraltar-Skaw range, duration 60-65 days. 5% total commission.
S4A (12.5%) Delivery US Gulf, laydays/cancelling 5/10 days from index date, redelivery Skaw-Passero range, duration about 30 days. 5% total commission.
S4B (12.5%) Delivery Skaw-Passero range, laydays/cancelling 5/10 days from index date, redelivery US Gulf, duration about 30 days. 5% total commission.

Table 5.c: Route Composition of Baltic Supramax Index

5.4 Baltic Handysize Index

Vessel Description:

- 28,000mt dwt on 9.78m SSW draft
- Max age 15 yrs
- LOA 169m, beam 27m
- 37,523cbm grain, 35,762cbm bale
- 5 holds, 5 hatches
- 4 x 30mt cranes
- 14 knots average laden/ballast on 22mt fuel oil (380cst), no diesel at sea

HS1 (12.5%) Delivery Skaw-Passero range, laydays/cancelling 5/10 days from index date, redelivery Recalada-Rio de Janeiro range, duration 35-45 days. 5% total commission.
HS2 (12.5%) Delivery Skaw-Passero range, laydays/cancelling 5/10 days from index date, redelivery Boston-Galveston range. Duration 35-45 days. 5% total commission.
HS3 (12.5%) Delivery Recalada-Rio de Janeiro range, laydays/cancelling 5/10 days from index date, redelivery Skaw-Passero range, duration 35-45 days. 5% total commission.
HS4 (12.5%) Delivery US Gulf, laydays/cancelling 5/10 days from index date, for a trip via US Gulf or north coast South America, redelivery Skaw-Passero range, duration 35-45 days. 5% total commission.
HS5 (25%) Delivery South East Asia, laydays/cancelling 5/10 days from index date, trip via Australia, redelivery Singapore–Japan range including China, duration 25-30 days. 5% total commission.
HS6 (25%) Delivery South Korea-Japan range, laydays/cancelling 5/10 days from index date, trip via North Pacific, redelivery Singapore-Japan range including China, duration 40-45 days. 5% total commission.

Table 5.d: Route Composition of Baltic Handysize Index

5.5 Baltic Dry Index

Until 30 June 2009 BDI used to be calculated by the equal weighted average of the BCI, BPI, BHSI and the BSI index (multiplied by a multiplier). The multiplier is a number constructed by the Exchange to ensure that the index level stays consistent through time. Suggested by the members, Baltic Exchange modified its methodology for BDI through the following formula:

BDI = ((Capesize Time Charter Average Rate + Panamax Time Charter Average Rate + Supramax Time Charter Average Rate + Handysize Time Charter Average Rate)/ 4) * 0.110345333

0.110345333 is the multiplier which was last updated on 6 May 2014. The time charter average rate is different from the index level. The time charter average rate is reported in $/day and serves as an intuitive indicator for market participants who need to calculate revenues/costs of their fixture and investment decisions. It only includes routes reported in time charter basis. The reported freight rate is multiplied by its route weighting before going into the time charter average rate. For Capesize, the voyage charter router C2, C3, C4, C5, C7, C15 do not contribute to the Capesize Time Charter Average Rate. Only C8_14, C9_14, C10_14, C14_14, and C16_14 are used with weightings of 25%, 12.5%, 25%, 25%, and 12.5% respectively. For the other three types of vessels, all the routes are included since they are all time charter routes.

6 Data

Analyzing daily volatility requires a high number of observations to derive robust volatility estimates. We used the data from the Baltic Exchange to analyze five indices separately and compare the results.

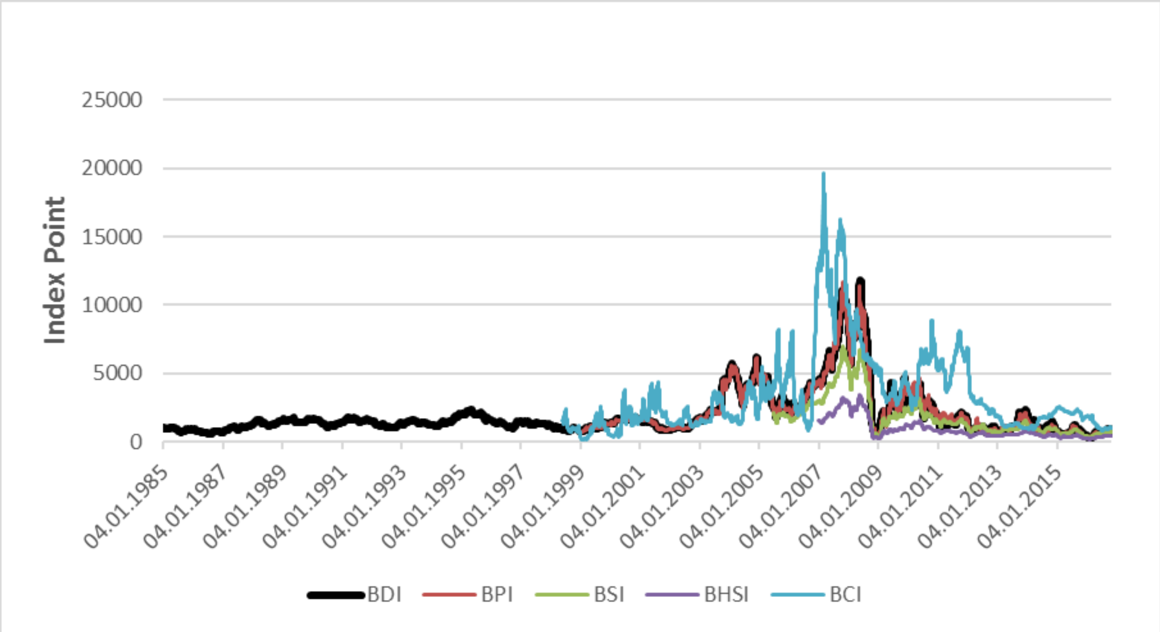
6.1 Raw Data Analysis

The scope of our research includes four major vessel types: Capesize, Panamax, Supramax, Handysize. Daily observations of Baltic Capesize Index (BCI), Baltic Panamax Index (BPI), Baltic Supramax Index (BSI), Baltic Handysize Index (BHSI), and Baltic Dry Index (BDI) are analyzed in the following sections. Table 6.a shows the sample period of each index.

Index	BCI	BPI	BSI	BHSI	BDI
From	1999.03.01	1998.12.31	2005.07.01	2007.01.02	1985.01.04
To	2016.11.03	2016.11.03	2016.11.03	2016.11.03	2016.11.03
Observations	4614	4457	2836	2462	7994

Table 6.a: Sample Period

Graphs 6.a presents the daily index level in all segments of dry bulk market.



Graph 6.a. Daily index level in the dry bulk market

The rates peaked before the start of financial crisis in 2008 and then plummeted, exampled by the 95% drop of BDI. Since then the price has moderately recovered but has not returned to the its previous level.

Segment	BCI	BPI	BSI	BHSI	BDI
Mean	3462.95	2398	1837	925	1936
Median	2314	1607	1380	658	1394
Standard Deviation	3125	2105	1454	742	1717
Kurtosis	4.83	3.74	2.13	1.68	9.22
Skewness	2.08	1.91	1.64	1.63	2.85
Minimum	161	282	243	183	290
Maximum	19687	11713	6956	3407	11793
Observations	4614	4457	2836	2462	7994

Table 6.b: Descriptive statistics of the index level

The descriptive statistics of the indices' prices pinpoints two extremely important patterns – positive skewness, which indicates that the dry bulk prices lay right to the mean more often, then left or, in other words, high prices are met more often than low, as compared to the average. This probably indicates a bubble in the shpping prices, which took place before the financial crisis in 2008. Next, we see differences in Kurtosis: BDI, Panamax and Capesize are leptokurtic, while Handysize and Supramax are platykurtic.

Overall, we can clearly see a long-term trending in the price level of the indices, therefore the trends need to be eliminated from the data to analyze volatility structure. Besides, for practical purposes the change of levels is more important for the industry.

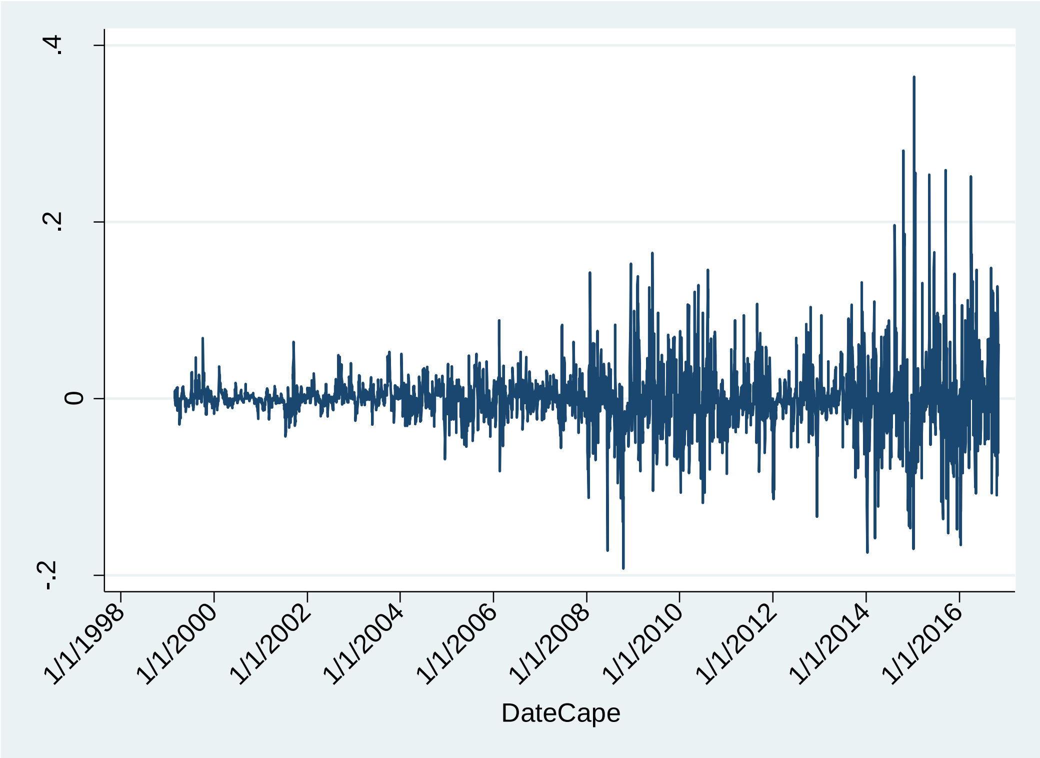
6.2 Preliminary Analysis of the Daily Log Returns

Below we present the summary statistics of daily log returns[1] for BDI, BCI, BPI, BSI, and BHSI.

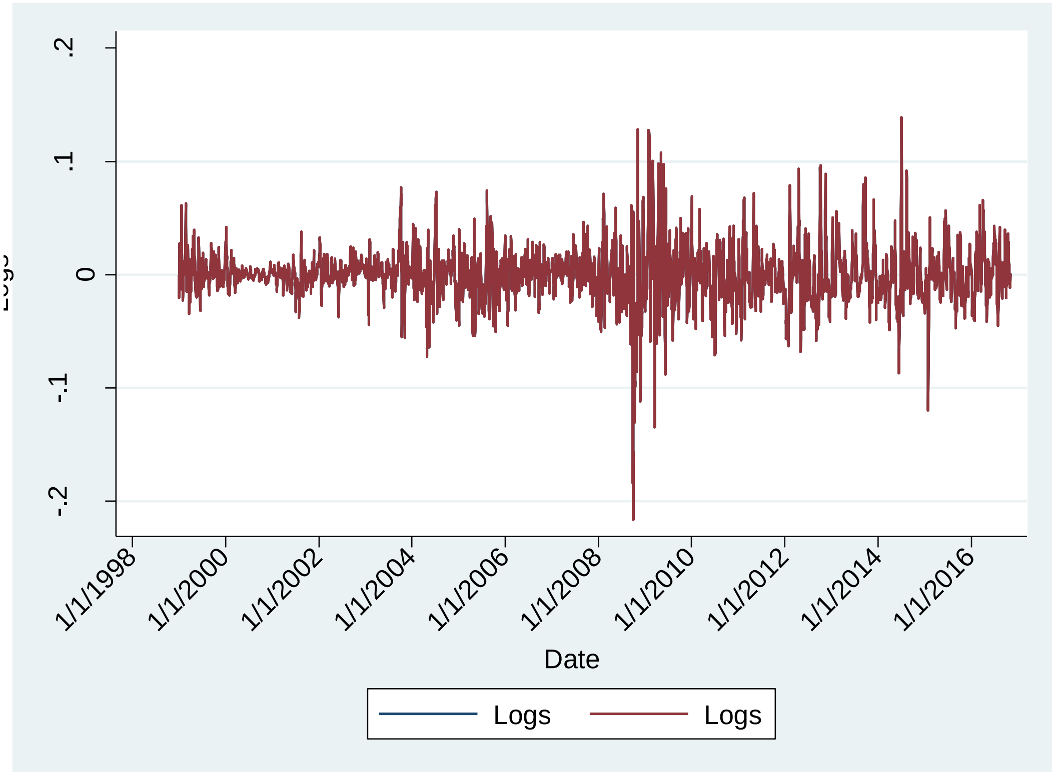
Segment	R_BCI	R_BPI	R_BSI	R_BHSI	R_BDI
Mean	9.41E-05	4.28E-05	-3.76E-04	-5.14E-04	-2.05E-05
Median	0	0	0	0	0
Standard Deviation	0.0340	0.0233	0.0157	0.0130	0.0151
Sample Variance	0.0012	0.0005	0.0002	0.0002	0.0002
Kurtosis	11.82	7.18	18.31	14.01	9.47
Skewness	0.99	-0.08	0.38	-1.14	0.08
Range	0.5566	0.3551	0.3194	0.2358	0.2573
Minimum	-0.1921	-0.2162	-0.1166	-0.1377	-0.1207
Maximum	0.3644	0.1389	0.2028	0.0981	0.1366
J-B Statistic	27533.11	9560.62	39529.90	20576.34	29811.19
Observations	4614	4457	2836	2462	7994

Table 6.c: Descriptive statistics of log returns for the index

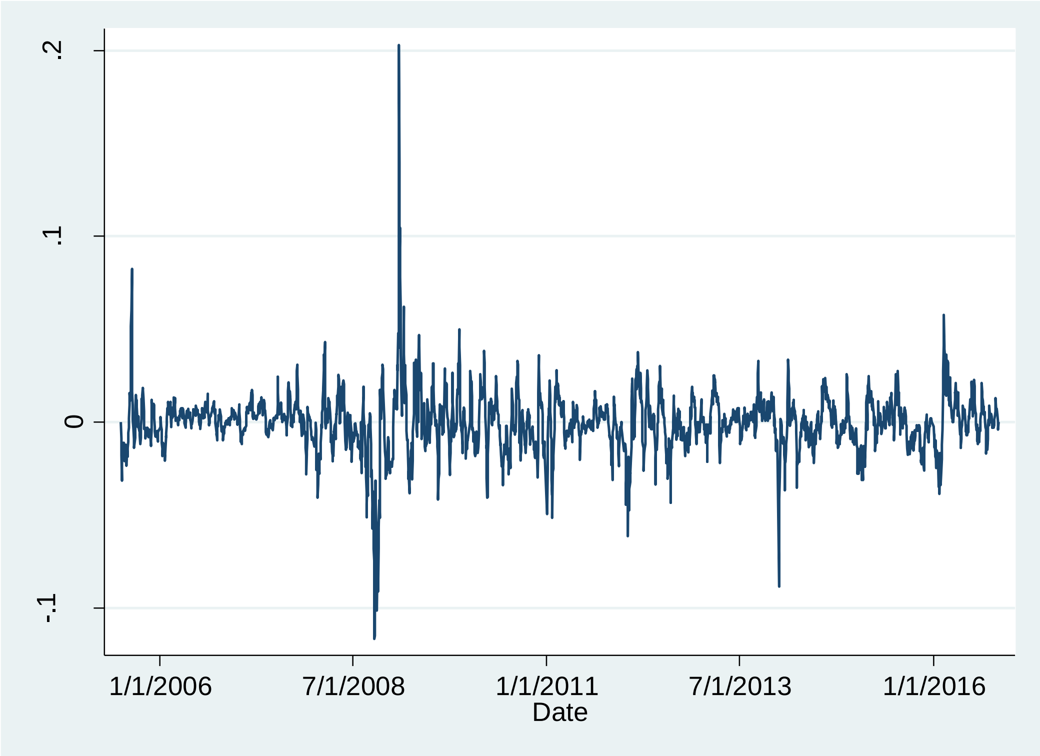
[1] Log returns have been calculated as $Ln(P_{n+1}/P_n)$, where P_n is the index value for the day N



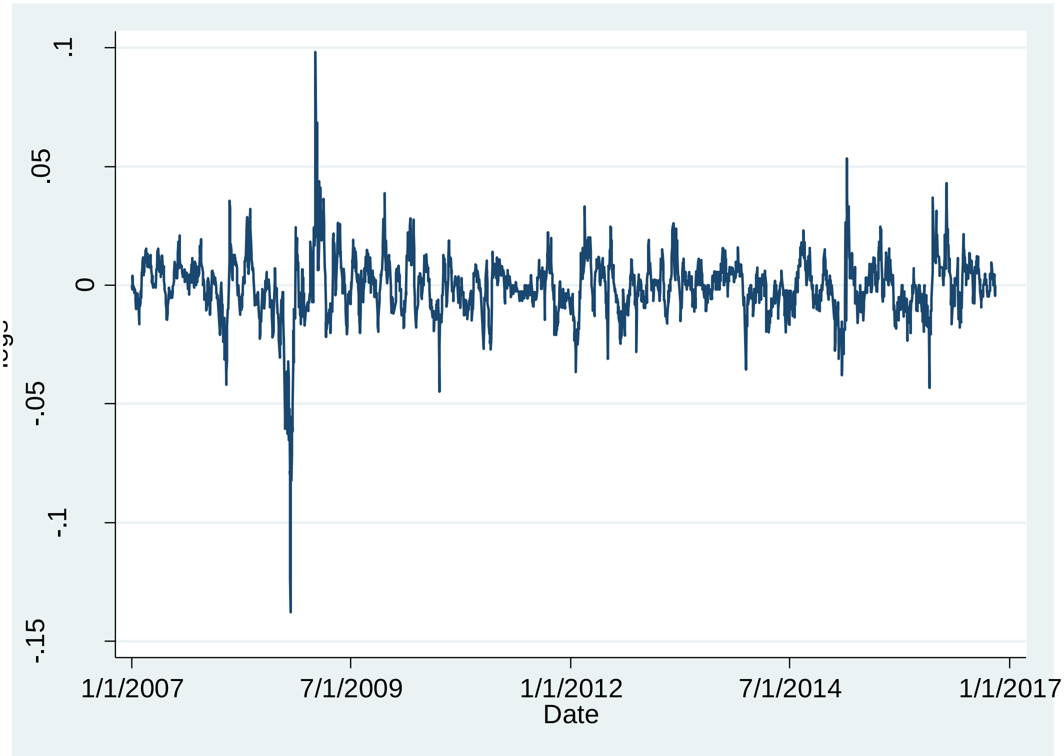
Graph 6.b. Daily log returns of BCI



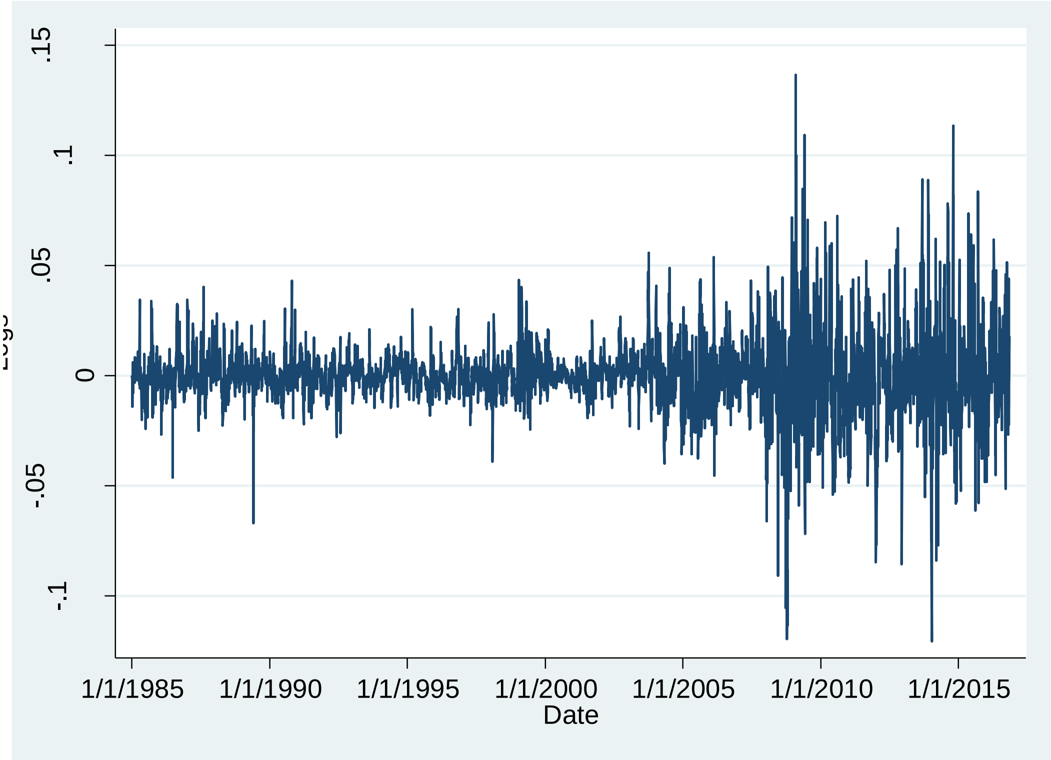
Graph 6.c. Daily log returns of BPI



Graph 6.d. Daily log returns of BSI



Graph 6.e. Daily log returns of BHSI



Graph 6.f. Daily log returns of BDI

From graph 6.b to graph 6.f we observe that daily log returns of all the indices move around 0 and have time-varying volatility. Abnormally high volatilities are observed around the 2008 financial crisis and recently around 2015. We also observe that sometimes bigger changes follow bigger changes, indicating the presence of volatility clustering. Therefore, homoscedastic volatility models are unlikely to explain the dynamics in the dry bulk segment. We observe similar patterns in the graphs of other dry bulk indices.

Next, we analyze the descriptive properties of the daily price changes in all dry bulk segments and the compound BDI index presented in table X. First of all, we notice that both their means and medians are very close to 0. We note that the classical measure of standard deviation increases with the size of the vessels from 1.3% for smallest Handysize ships to 3.4% for Capesize ships. This can likely be due to the limited nature of cargoes and ports for bigger ships. Despite being more volatile, Capesize and Panamax ships earn small positive returns on average, as compared to small negative returns for Handysize and Supramax.. We also observe a difference in maximum and minimum market reactions, especially well-depicted in the statistic of Capesize index with +36.4% daily max gain and -19.2% daily max loss, which signals a potential property of the market: asymmetric reactions to positive and negative news. Yet again, the bigger the vessel class – the greater the difference between the maximum and the minimum price change, hence we yet again have an intention to test for likely asymmetries in the volatility structure.

According to Jarque-Bera statistics, we can reject the hypothesis of normality at 1% significance level for all five indices. Negative skewness for Handysize and Panamax indices daily returns indicates that extremely high losses are more likely than extremely high returns negative losses, while BDI, Capesize and Supramax are positively-skewed, hence for them the opposite is true. In addition, the composite BDI has a skewness closer to 0, showing a potential effect of diversification. All five indices daily log returns are leptokurtic, which indicates they are likelier to have extreme outcomes, as compared to normal distributions. This is quite typical for high-frequency financial data.

We use both Augmented Dickey-Fuller Test and Phillips-Perron Test to examine whether the series are stationary. If both tests yield the same conclusion we could be more certain about the result. The results presented in table 6.d show that for all series the null hypothesis of presence of a unit root is rejected at 5% significance level. We conclude that all series are stationary.

	R_BCI	R_BPI	R_BSI	R_BHSI	R_BDI
ADF t-statistic	-24.15	-19.08	-12.74	-10.66	-27.63
Phillips-Perron t-statistic	-29.82	-12.99	-13.13	-11.64	-30.51
Critical Value (5%)	-2.86	-2.86	-2.86	-2.86	-2.86

Table 6.d: Stationarity test

We use Ljung-Box Q-statistic to test for autocorrelation in the series for order 1, 7 and 30. The results presented in table 6.e reject the null hypothesis that autocorrelations up to lag 1, 7, and 30 equal zero at 1% significance level. We conclude that all series are subject to strong autocorrelation, which is common for high-frequency financial series.

	R_BCI	R_BPI	R_BSI	R_BHSI	R_BDI
Q(1) (p-value)	2036 (0.000)	3300 (0.000)	2249 (0.000)	1941 (0.000)	4861 (0.000)
Q(7) (p-value)	3118 (0.000)	6440 (0.000)	8264 (0.000)	8142 (0.000)	10253 (0.000)
Q(30) (p-value)	3525 (0.000)	6511 (0.000)	9505 (0.000)	10373 (0.000)	11763 (0.000)

Table 6.e: Ljung–Box Q test

7 Empirical Results

In this section, we attempt to find the most suitable model for each series. We apply a number of models based on theories, previous studies, and statistical tests. Then we analyze the result, compare the differences, and derive economic insights in each segment.

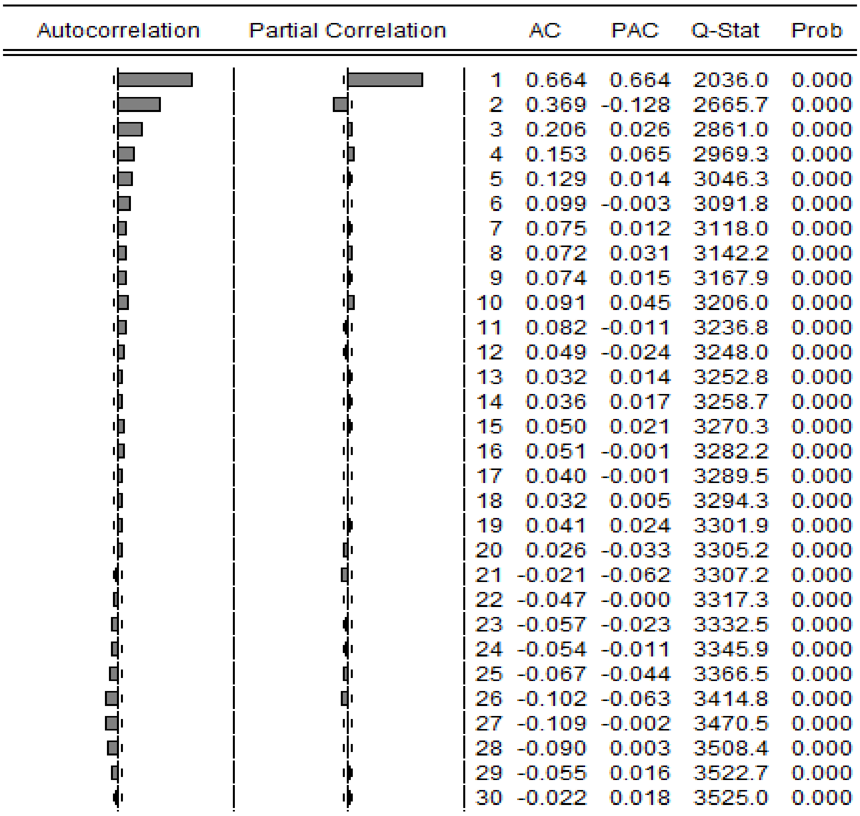
The data of each segment has its own characteristics driven by different economic natures. Therefore, each segment shall be analyzed separately. In the previous studies of dry bulk freight volatilities, several heteroscedastic models, presented in table 7.a, were fitted for different dry bulk segments. In our own estimation process we consider all those models and focus on univariate specifications.

Table 7.a Model specifications in previous dry bulk market studies

Study	Model	Vessel Type
Kavussanos(1996)	GARCH(1,1)	Capesize Panamax Handysize
Kavussanos and Alizadeh (2002)	EGARCH(1,1)-M	
Marlow et al. (2008)	GARCH(1,1) EGARCH(1,1)	
Alizadeh and Nomikos (2011)	EGARCH(1,1)-X	
Xu (2011)	GARCH(1,1)	Capesize Panamax
Drobetz (2012)	EGARCH(1,1) EGARCH(1,1)-X	
Geomelos and E. Xideas (2014)	GARCH(3,3)-M EGARCH(1,3)-M	Capesize
	GARCH(3,3)-M EGARCH(1,6)-M	Panamax
	GARCH(1,3) EGARCH(1,6)-M	Handysize

7.1 Baltic Capesize Index

According to correlogram and Ljung-Box Q statistics, we assume the presence of autoregressive stationary process of 2^{nd} order – ARMA(2:0) model.



Lag	AC	PAC	Q-Stat	Prob
1	0.664	0.664	2036.0	0.000
2	0.369	-0.128	2665.7	0.000
3	0.206	0.026	2861.0	0.000
4	0.153	0.065	2969.3	0.000
5	0.129	0.014	3046.3	0.000
6	0.099	-0.003	3091.8	0.000
7	0.075	0.012	3118.0	0.000
8	0.072	0.031	3142.2	0.000
9	0.074	0.015	3167.9	0.000
10	0.091	0.045	3206.0	0.000
11	0.082	-0.011	3236.8	0.000
12	0.049	-0.024	3248.0	0.000
13	0.032	0.014	3252.8	0.000
14	0.036	0.017	3258.7	0.000
15	0.050	0.021	3270.3	0.000
16	0.051	-0.001	3282.2	0.000
17	0.040	-0.001	3289.5	0.000
18	0.032	0.005	3294.3	0.000
19	0.041	0.024	3301.9	0.000
20	0.026	-0.033	3305.2	0.000
21	-0.021	-0.062	3307.2	0.000
22	-0.047	-0.000	3317.3	0.000
23	-0.057	-0.023	3332.5	0.000
24	-0.054	-0.011	3345.9	0.000
25	-0.067	-0.044	3366.5	0.000
26	-0.102	-0.063	3414.8	0.000
27	-0.109	-0.002	3470.5	0.000
28	-0.090	0.003	3508.4	0.000
29	-0.055	0.016	3522.7	0.000
30	-0.022	0.018	3525.0	0.000

Graph 7.a: Correlogram for BCI

The Automatic ARIMA Forecasting function provided by EViews suggests an ARMA (2,3) specification. The two estimated models are presented in table 7.b.

Table 7.b: ARMA models specification for BCI

Variable	**AR (2)**	**ARMA (2,3)**
Constant	0.000115	0.000123
	(0.97)	(0.92)
R_BCI_{t-1}	0.75	1.098
	(0)	(0)
R_BCI_{t-2}	-0.13	-0.2
	(0)	(0.0003)
MA_{t-1}		-0.35
		(0)
MA_{t-2}		-0.197
		(0)
MA_{t-3}		-0.1249
		(0)
R^2	0.4505	0.4539
AIC	-4.522	-4.526
BIC	-4.516	-4.516
Durbin-Watson	1.992	1.998

Comparing AR(2) model to ARMA(2;3) model we can see that the latter has a higher R^2 of 45.39%, as compared to 45.05%; both AIC and BIC are slightly lower for ARMA(2;3) as compared to AR(2) . We conclude that despite including slightly more information, ARMA(2;3) loses degrees of freedom, while including 3 extra variables, which complicate the modelling and the interpretation.. Therefore, we opt to go on with a parsimonious AR(2) model. The Durbin-Watson statistic is at 1.992, which indicates a small positive autocorrelation.

To investigate the presence of ARCH-effects we use an ARCH -LM of order 5, according to which we reject the null hypothesis of no heteroscedasticity at 5% level of significance, therefore – we assume the presence of ARCH effects. We see presence of ARCH-effects in first, second and fifth orders of squared residuals' lagged variables.

Table 7.c: ARCH-LM test for BCI

Variable	t-Statistic	p-value
ε_{t-1}^2	22.28	0
ε_{t-2}^2	2.48	0.0132
ε_{t-3}^2	1.56	0.1182
ε_{t-4}^2	0.69	0.4909
ε_{t-5}^2	9.08	0
F-statistic	165.87	0

We start from GARCH(1,1) specification and we find that the standardized residuals appear to be non-normally distributed. The Jarque-Bera statistics strongly rejects the hypothesis of normality (Appendix B). Therefore, robust standard errors are used by selecting Bollershev-Wooldridge Heteroscedasticity consistent covariance matrix, therefore using Quasi Maximum Likelihood.

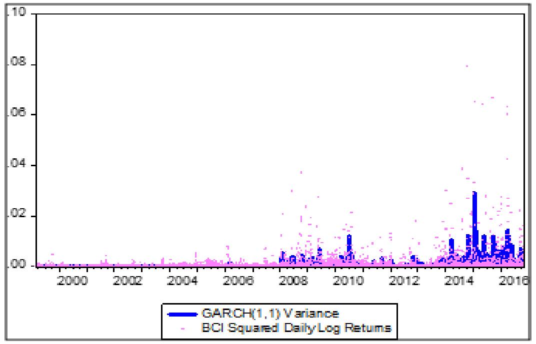
Sign and size bias test, developed by Engle and Ng (1993) joint test statistic, formulated by calculating TR^2, suggests strongly rejecting the null hypothesis of no asymmetric effects. Moreover, both size effects and sign effect appear to be significant at 1% level (Appendix C). Hence, we specify an EGARCH(1,1) model for BCI. EGARCH(1;1) with GED model's asymmetry term was found insignificant, however, the sign and size bias test strongly suggests the presence of asymmetric leverage effects. One of the possible explanation for this phenomenon is higher order of asymmetry in the Capesize returns. our data is a daily data of higher frequency, as opposed to the most of previous studies, we would like to first test for the increased order of ARCH effects, rather than GARCH effects. The reasoning is simple: the shocks are likely to be more important in the short term and the asymmetry is brought by the innovations, hence this is where we seek for the model improvement.

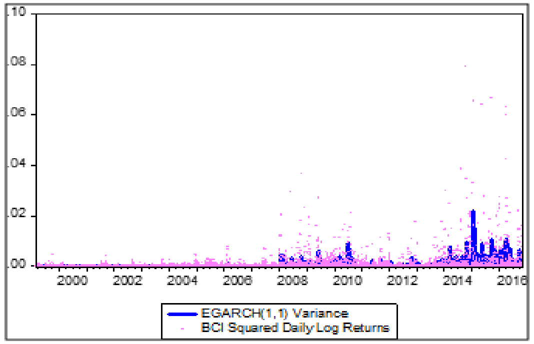
Compliant with the logic of Adland and Cullinane(2005) and with empirical results from N.D. Geomelos and E. Xideas (2014), we simultaneously test for the risk premium in the Capesize spot returns. Hence, we estimate an EGARCH(2;1)-M model. Table 7.d presents the considered models for BCI, their information criteria and diagnostics, followed by the comparison of model forecast and realized volatility (approximated by squared daily log returns) and the QQ-plots of the standardized residuals.

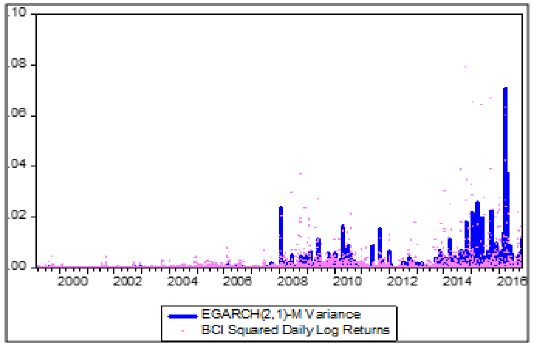
Table 7.d: Results of Estimated GARCH type models for BCI

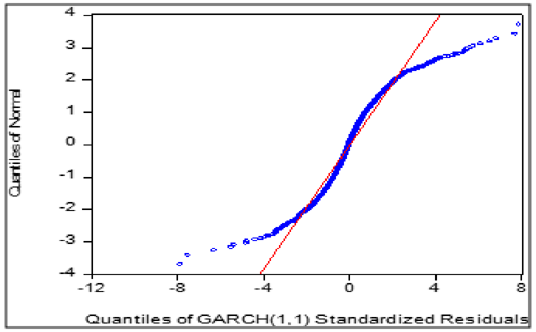
	GARCH(1,1)	EGARCH(1,1)	EGARCH(2,1) -M
Mean Equation	$R\_BCI_t = \underset{(0.40)}{0.00039} + \underset{(0.00)}{0.91}R\_BCI_{t-1} - \underset{(0.00)}{0.14}R\_BCI_{t-2} + \epsilon_t$	$R\_BCI_t = \underset{(0.01)}{0.0014} + \underset{(0.00)}{0.93}R\_BCI_{t-1} - \underset{(0.0)}{0.16}R\_BCI_{t-2} + \epsilon_t$	$R\_BCI_t = \underset{(0.02)}{-0.07}\sigma_t - \underset{(0.03)}{0.0043} + \underset{(0.00)}{0.91}R\_BCI_{t-1} - \underset{(0.00)}{0.15}R\_BCI_{t-2} + \epsilon_t$
Variance Equation	$\sigma_t^2 = \underset{(0.00)}{0.00000069} + \underset{(0.00)}{0.19}\epsilon^2_{t-1} + \underset{(0.00)}{0.85}\sigma^2_{t-1}$	$\log\sigma^2_t = \underset{(0.00)}{-0.32} + \underset{(0.00)}{0.33}\frac{\lvert\epsilon_{t-1}\rvert}{\sigma_{t-1}} + \underset{(0.32)}{0.0025}\frac{\epsilon_{t-1}}{\sigma_{t-1}} + \underset{(0}{0.99}\log\sigma^2_{t-1}$	$\log\sigma^2_t = \underset{(0.00)}{-0.16} + \underset{(0.00)}{0.60}\frac{\lvert\epsilon_{t-1}\rvert}{\sigma_{t-1}} - \underset{(0.00)}{0.41}\frac{\lvert\epsilon_{t-2}\rvert}{\sigma_{t-2}} + \underset{(0.01)}{0.10}\frac{\epsilon_{t-1}}{\sigma_{t-1}} - \underset{(0.01)}{0.10}\frac{\epsilon_{t-2}}{\sigma_{t-2}} + \underset{(0.00)}{0.99}\log\sigma^2_{t-1}$
Adj. R^2	42.65%	42.50%	42.66%
AIC	-5.57	-5.57	-5.63
BIC	-5.56	-5.56	-5.61
Durbin-Watson	2.25	2.28	2.23
ARCH-LM (5)	0.01	0.00	0.00

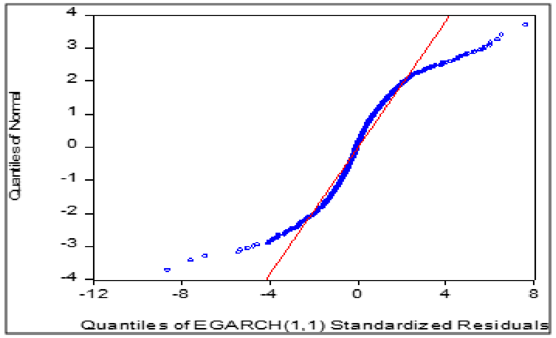
Notes: p-value in the parenthesis; Adj. R^2 is the corrected measure for module accuracy for the variance explained by the model inpus; AIC is the Akaike Information criteria (Akaike, 1974); BIC is the Schwarz information criteria (Schwarz, 1978); Durbin-Watson statistic tests the autocorrelation in the residuals; ARCH-LM test is Engle's test for autoregressive conditional heteroscedasticity.

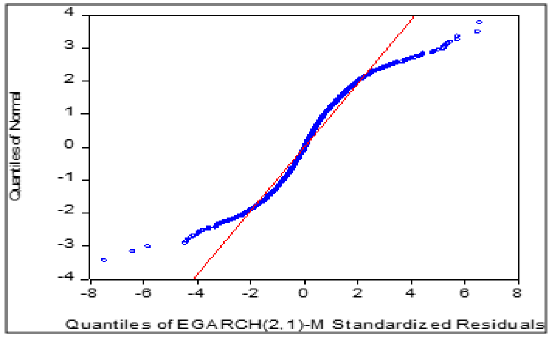






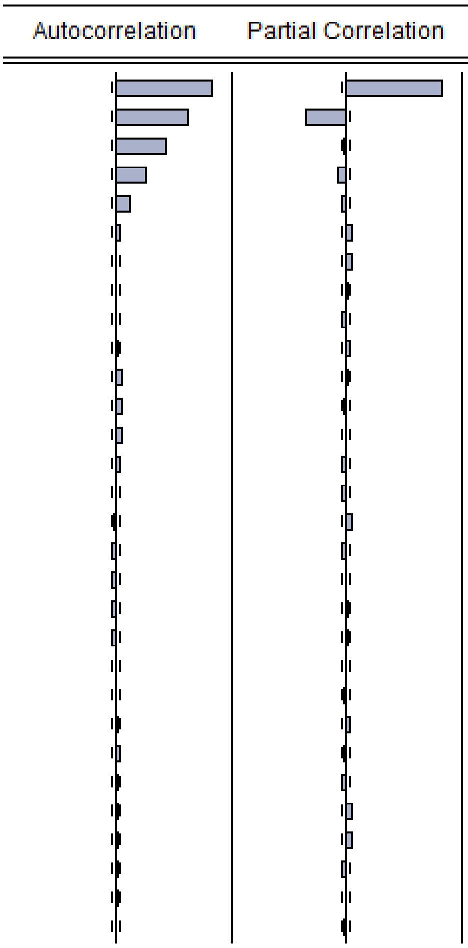






7.2 Baltic Panamax Index

According to correlogram and Ljung-Box Q statistics, we assume the presence of autoregressive stationary process of 2nd order – ARMA(2:0) model.



	AC	PAC	Q-Stat	Prob
1	0.860	0.860	3300.0	0.000
2	0.646	-0.360	5162.9	0.000
3	0.444	-0.009	6043.2	0.000
4	0.269	-0.068	6366.3	0.000
5	0.124	-0.047	6435.1	0.000
6	0.033	0.060	6439.9	0.000
7	-0.001	0.059	6439.9	0.000
8	-0.001	0.009	6439.9	0.000
9	0.007	-0.026	6440.1	0.000
10	0.022	0.027	6442.4	0.000
11	0.042	0.019	6450.4	0.000
12	0.051	-0.023	6462.0	0.000
13	0.047	-0.000	6472.0	0.000
14	0.032	-0.028	6476.5	0.000
15	0.005	-0.035	6476.6	0.000
16	-0.013	0.045	6477.3	0.000
17	-0.029	-0.035	6481.0	0.000
18	-0.040	0.004	6488.1	0.000
19	-0.040	0.014	6495.2	0.000
20	-0.026	0.022	6498.2	0.000
21	-0.008	-0.000	6498.5	0.000
22	0.006	-0.009	6498.6	0.000
23	0.021	0.031	6500.5	0.000
24	0.031	-0.019	6504.8	0.000
25	0.022	-0.046	6506.9	0.000
26	0.012	0.043	6507.6	0.000
27	0.016	0.042	6508.7	0.000
28	0.017	-0.036	6509.9	0.000
29	0.012	0.003	6510.6	0.000
30	0.005	-0.019	6510.7	0.000

Graph 7.b: Correlogram for BPI

The Automatic ARIMA Forecasting function provided by EViews suggests an ARMA (4,4) specification. The two estimated models are presented in table 7.e.

Table 7.e: ARMA models specification for BPI

Variable	AR (2)	ARMA (5,5)
Constant	-0.000057	-0.000000537
	(0.95)	(0.95)
R_BPI_{t-1}	1.17	1.73
	(0)	(0)
R_BPI_{t-2}	-0.36	-1.71
	(0)	(0)
R_BPI_{t-3}		1.81
		(0)
R_BPI_{t-4}		-1.53
		(0)
R_BPI_{t-5}		0.54
		(0)
MA_{t-1}		-0.57
		(0)
MA_{t-2}		0.69
		(0)
MA_{t-3}		-0.75
		(0)
MA_{t-4}		0.36
		(0)
MA_{t-5}		0.148
		(0)
R^2	0.7737	0.7783
AIC	-6.17	-6.18
BIC	-6.16	-6.17
Durbin-Watson	2	1.997

Comparing AR(2) model to ARMA(5;5) model we can see that the latter has a higher R^2 of 77.83%, as compared to 77.37% , both AIC and BIC are slightly lower for ARMA(5;5) as compared to AR(2) . We conclude that despite including slightly more information about log-returns, ARMA(5;5) loses degrees of freedom and does not add much forecasting power, as compared to a simpler AR(2). Therefore, we opt to go on with a parsimonious AR(2) model. The Durbin-Watson Statistic equals 2, indicating no presence of serial autocorrelation.

To investigate the presence of ARCH-effects we use an ARCH -LM of order 5, according to which we reject the null hypothesis of no heteroscedasticity at 5% level of significance, therefore – we assume the presence of ARCH effects. We see presence of ARCH-effects in first 5 orders of squared residuals' lagged variables.

Table 7.f: ARCH-LM test for BPI

Variable	t-Statistic	p-value
ε_{t-1}^2	16.18	0
ε_{t-2}^2	10.52	0
ε_{t-3}^2	-2.5	0.0123
ε_{t-4}^2	5.6	0
ε_{t-5}^2	10.52	0
F-statistic	183.87	0

We start from GARCH(1,1) specification and find that the standardized residuals appear to be non-normally distributed. The Jarque-Bera statistics strongly rejects the hypothesis of normality (Appendix B). Therefore, robust standard errors are used by selecting Bollershev-Wooldridge Heteroscedasticity consistent covariance matrix, therefore using Quasi Maximum Likelihood.

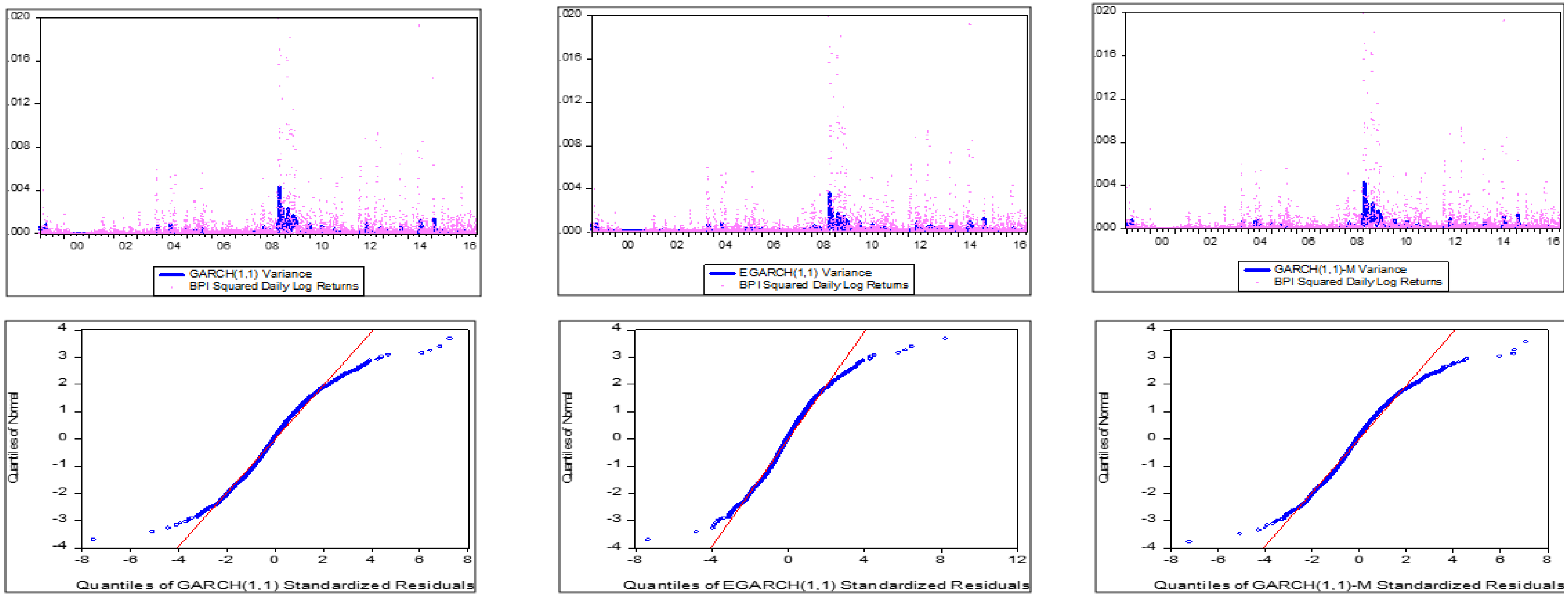
Sign and size bias test, developed by Engle and Ng (1993) joint test statistic, formulated by calculating TR^2, suggests strongly rejecting the null hypothesis of no asymmetric effects. Sign effect appears to be insignificant, indicating absence of asymmetry, while both positive and negative size effects are significant at 1% level (Appendix C). EGARCH(1;1) asymmetry term was found insignificant compliant with the results of the asymmetry test.

Kavussanos and Alizadeh (2002) found a time varying risk premium in the Panamax returns. Hence, we estimate GARCH(1;1)-M model and confirm the presence of positive risk premium. Table 7.g presents the considered models for BPI, their information criteria and diagnostics, followed by the comparison of model forecast and realized volatility (approximated by squared daily log returns) and the QQ-plots of the standardized residuals.

Table 7.g: Results of Estimated GARCH type models for BPI

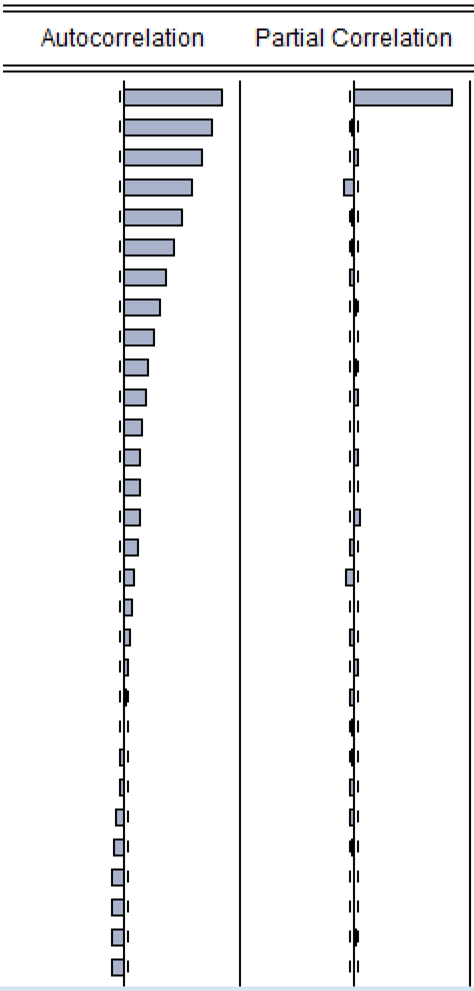
	GARCH(1,1)	EGARCH(1,1)	GARCH-M(1,1)
Mean Equation	$R\_BPI_t = \underset{(0.091)}{0.00089} + \underset{(0.00)}{1.18} R\_BPI_{t-1} - \underset{(0.00)}{0.35} R\_BPI_{t-2} + \epsilon_t$	$R\_BPI_t = \underset{(0.09)}{0.00085} + \underset{(0.00)}{1.17} R\_BPI_{t-1} - \underset{(0.00)}{0.34} R\_BPI_{t-2} + \epsilon_t$	$R\_BPI_t = \underset{(0.11)}{0.05}\sigma_t + \underset{(0.04)}{0.0031} + \underset{(0.00)}{1.18} R\_BPI_{t-1} - \underset{(0.00)}{0.34} R\_BPI_{t-2} + \epsilon_t$
Variance Equation	$\sigma_t^2 = \underset{(0.00)}{0.00000056} + \underset{(0.00)}{0.17}\epsilon^2_{t-1} + \underset{(0.00)}{0.85}\sigma^2_{t-1}$	$\log\sigma^2_t = \underset{(0.00)}{-0.47} + \underset{(0.00)}{0.34}\frac{\lvert\epsilon_{t-1}\rvert}{\sigma_{t-1}} - \underset{(0.95)}{0.0014}\frac{\epsilon_{t-1}}{\sigma_{t-1}} + \underset{(0.00)}{0.98}\log\sigma^2_{t-1}$	$\sigma_t^2 = \underset{(0.00)}{0.00000057} + \underset{(0.00)}{0.17}\epsilon^2_{t-1} + \underset{(0.00)}{0.85}\sigma^2_{t-1}$
Adj. R^2	77.32%	77.34%	77.27%
AIC	-6.70	-6.71	-6.70
BIC	-6.70	-6.70	-6.69
Durbin-Watson	2.02	1.99	2.02
ARCH-LM (5)	0.00	0.00	0.01

Notes: p-value in the parenthesis; Adj. R^2 is the corrected measure for module accuracy for the variance explained by the model inpus; AIC is the Akaike Information criteria (Akaike, 1974); BIC is the Schwarz information criteria (Schwarz, 1978); Durbin-Watson statistic tests the autocorrelation in the residuals; ARCH-LM test is Engle's test for autoregressive conditional heteroscedasticity.



7.3 Baltic Supramax Index

According to the correlogram and Ljung-Box Q statistics, we assume the presence of autoregressive stationary process of 1st order – ARMA(1:0) model.



Lag	AC	PAC	Q-Stat	Prob
1	0.890	0.890	2248.5	0.000
2	0.787	-0.024	4007.9	0.000
3	0.703	0.036	5413.5	0.000
4	0.612	-0.082	6477.2	0.000
5	0.528	-0.013	7270.3	0.000
6	0.454	-0.014	7855.7	0.000
7	0.379	-0.045	8263.9	0.000
8	0.317	0.013	8550.1	0.000
9	0.264	-0.006	8748.6	0.000
10	0.220	0.012	8886.1	0.000
11	0.189	0.032	8988.2	0.000
12	0.164	0.002	9064.8	0.000
13	0.148	0.030	9127.2	0.000
14	0.135	-0.002	9179.3	0.000
15	0.132	0.043	9229.1	0.000
16	0.120	-0.046	9270.5	0.000
17	0.094	-0.075	9295.6	0.000
18	0.073	0.003	9310.9	0.000
19	0.049	-0.035	9317.9	0.000
20	0.032	0.030	9320.9	0.000
21	0.015	-0.025	9321.5	0.000
22	-0.007	-0.020	9321.6	0.000
23	-0.026	-0.011	9323.6	0.000
24	-0.046	-0.025	9329.5	0.000
25	-0.068	-0.027	9342.7	0.000
26	-0.087	-0.021	9364.3	0.000
27	-0.102	-0.007	9394.1	0.000
28	-0.112	0.001	9430.4	0.000
29	-0.115	0.016	9468.2	0.000
30	-0.113	0.006	9505.1	0.000

Graph 7.c: Correlogram for BSI

The Automatic ARIMA Forecasting function provided by EViews suggests an ARMA (4,4) specification. The two estimated models are presented in table 7.h.

Table 7.h: ARMA models specification for BSI

Variable	AR (1)	ARMA (4,4)
Constant	-0.000374	-0.000364
	(0.77)	(0.76)
R_BSI_{t-1}	0.89	0.18
	(0)	(0.12)
R_BSI_{t-2}		0.91
		(0.00)
R_BSI_{t-3}		0.31
		(0.00)
R_BSI_{t-4}		-0.52
		(0.00)
MA_{t-1}		0.735
		(0.00)
MA_{t-2}		-0.29
		(0.10)
MA_{t-3}		-0.52
		(0.10)
MA_{t-4}		0.07
		(0.00)
R^2	0.7919	0.7947
AIC	-7.0351	-7.043
BIC	-7.0288	-7.022
Durbin-Watson	1.96	1.99

Comparing AR(1) model to ARMA(4,4) model we can see that the latter has a higher R^2 of 79.47%, as compared to 79.19% and a slightly lower AIC criterion of -7.043 as compared to -7.0351 in AR(1). However, its BIC is higher, being at -7.022 versus -7.0288 in the AR(1). Hence, we conclude that despite including slightly more information, ARMA(4,4) loses degree of freedom and does not add much more forecasting power compared to an AR(1). Durbin-Watson statistic is around 1.96, close to 2, indicating the absence of strong autocorrelation in the residuals. Therefore, we opt to go on with a parsimonious AR(1) model.

To investigate the presence of ARCH-effects we use an ARCH -LM of order 5, according to which we reject the null hypothesis of no heteroscedasticity at 5% level of significance, therefore – we assume the presence of ARCH effects.

Table 7.i: ARCH-LM test for BSI

Variable	t-Statistic	p-value
ε_{t-1}^2	3.74	0.0002
ε_{t-2}^2	2.29	0.0219
ε_{t-3}^2	2	0.0458
ε_{t-4}^2	0.32	0.7497
ε_{t-5}^2	0.15	0.8805
F-statistic	5.35	0.0001

Particularly, we see presence of ARCH-effects in first 3 orders of squared residuals' lagged variables.

We start from GARCH(1,1) specification and find that the standardized residuals appear to be non-normally distributed. The Jarque-Bera statistics strongly rejects the hypothesis of normality (Appendix B). Therefore, robust standard errors are used by selecting Bollershev-Wooldridge Heteroscedasticity consistent covariance matrix, therefore using Quasi Maximum Likelihood.

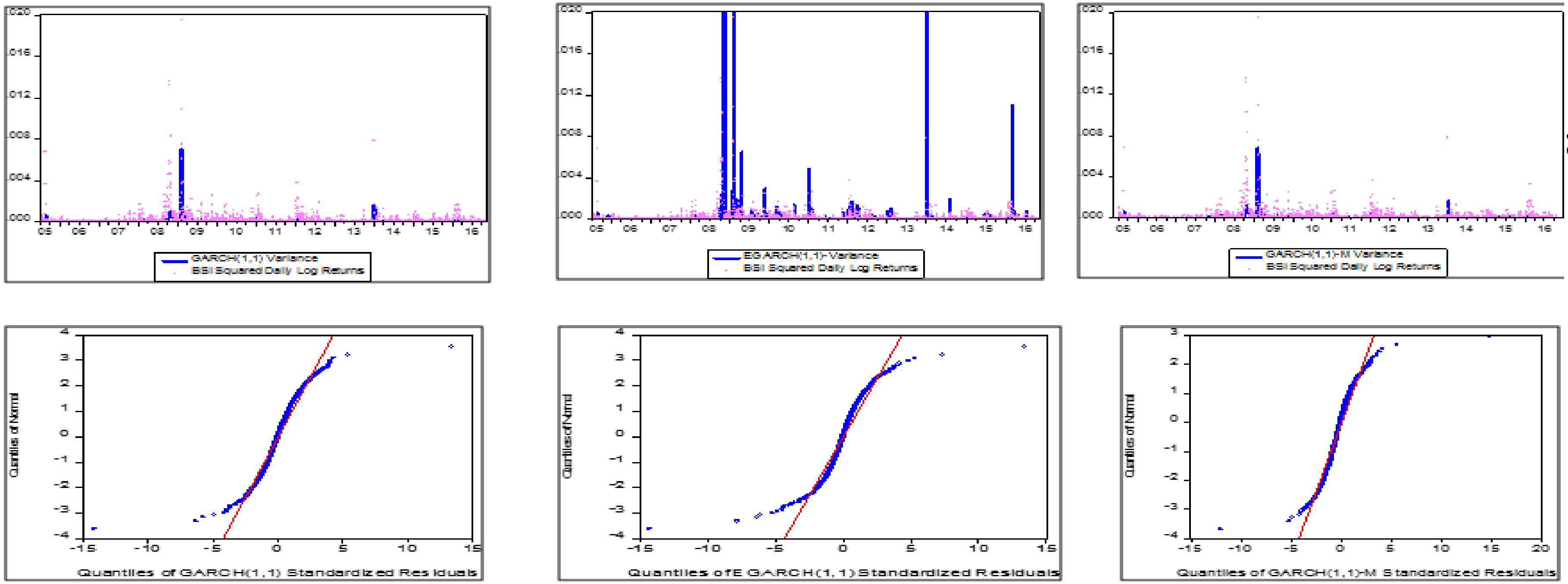
Sign and size bias test, developed by Engle and Ng (1993) joint test statistic, formulated by calculating TR^2, suggests strongly rejecting the null hypothesis of no asymmetric effects. Sign effect is insignificant, but both positive and negative size effects are significant at 1% level (Appendix C). EGARCH(1;1) asymmetry term was found insignificant, confirming the results of sign and size bias test.

Compliant with the logic of Adland and Cullinane(2005), we test for the risk premium in the Supramax spot returns. Hence, we estimate a GARCH(1;1)-M model and confirm the presence of positive risk premium. Table 7.j presents the considered models for BSI, their information criteria and diagnostics, followed by the comparison of model forecast and realized volatility (approximated by squared daily log returns) and the QQ-plots of the standardized residuals.

Table 7.j: Results of Estimated GARCH type models for BSI

	GARCH(1,1)	EGARCH(1,1)	GARCH-M(1,1)
Mean Equation	$R\_BSI_t = 0.0010 + 0.90R\_BSI_{t-1} + \epsilon_t$ (0.33) (0.00)	$R\_BSI_t = 0.00037 + 0.84R\_BSI_{t-1} + \epsilon_t$ (0.66) (0.00)	$R\_BSI_t = 0.17\sigma_t + 0.01 + 0.90R\_BSI_{t-1} + \epsilon_t$ (0.01) (0.00) (0.00)
Variance Equation	$\sigma_t^2 = 0.0000050 + 0.25\epsilon^2_{t-1} + 0.81\sigma^2_{t-1}$ (0.06) (0.00) (0.00)	$\log\sigma^2_t = -9.86 + 1.04\frac{\lvert\epsilon_{t-1}\rvert}{\sigma_{t-1}} + 0.15\frac{\epsilon_{t-1}}{\sigma_{t-1}} + 0.09\log\sigma^2_{t-1}$ (0.00) (0.00) (0.41) (0.37)	$\sigma_t^2 = 0.00000056 + 0.25\epsilon^2_{t-1} + 0.80\sigma^2_{t-1}$ (0.04) (0.00) (0.00)
Adj. R^2	79.18%	78.90%	78.96%
AIC	-7.51	-7.28	-7.52
BIC	-7.5	-7.27	-7.51
Durbin-Watson	1.97	1.83	1.98
ARCH-LM (5)	0.97	0.05	0.98

Notes: p-value in the parenthesis; Adj. R^2 is the corrected measure for module accuracy for the variance explained by the model inpus; AIC is the Akaike Information criteria (Akaike, 1974); BIC is the Schwarz information criteria (Schwarz, 1978); Durbin-Watson statistic tests the autocorrelation in the residuals; ARCH-LM test is Engle's test for autoregressive conditional heteroscedasticity.



7.4 Baltic Handysize Index

According to correlogram and Ljung-Box Q statistics, we assume the presence of autoregressive stationary process of 2nd order – ARMA(2:0) model.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.887	0.887	1940.9	0.000
		2	0.808	0.098	3552.0	0.000
		3	0.740	0.030	4904.0	0.000
		4	0.674	-0.013	6025.3	0.000
		5	0.602	-0.059	6919.6	0.000
		6	0.526	-0.067	7603.0	0.000
		7	0.467	0.023	8142.3	0.000
		8	0.416	0.016	8570.9	0.000
		9	0.379	0.050	8926.7	0.000
		10	0.344	0.008	9219.3	0.000
		11	0.311	-0.006	9458.4	0.000
		12	0.281	-0.008	9654.6	0.000
		13	0.256	-0.000	9817.3	0.000
		14	0.239	0.025	9959.0	0.000
		15	0.217	-0.015	10076.	0.000
		16	0.190	-0.039	10165.	0.000
		17	0.172	0.023	10239.	0.000
		18	0.145	-0.052	10291.	0.000
		19	0.112	-0.052	10322.	0.000
		20	0.086	0.005	10340.	0.000
		21	0.060	-0.017	10349.	0.000
		22	0.038	0.002	10353.	0.000
		23	0.017	-0.006	10354.	0.000
		24	0.004	0.014	10354.	0.000
		25	-0.014	-0.032	10354.	0.000
		26	-0.025	0.009	10356.	0.000
		27	-0.036	-0.015	10359.	0.000
		28	-0.041	0.015	10363.	0.000
		29	-0.042	0.013	10368.	0.000
		30	-0.045	-0.002	10373.	0.000

Graph 7.d: Correlogram for BHSI

The Automatic ARIMA Forecasting function provided by EViews suggests an ARMA (6,5) specification. The two estimated models are presented in table 7.k.

Table 7.k: AR models specifications for BHSI

Variable	**AR (2)**	**ARMA (6,5)**
Constant	-0.000512	-0.000515
	(0.68)	(0.67)
R_BHSI $_{t-1}$	0.8	0.64
	0	(0.48)
R_BHSI $_{t-2}$	0.098	0.37
	0	(0.05)
R_BHSI $_{t-3}$		-0.18
		(0.68)
R_BHSI $_{t-4}$		-0.47
		0
R_BHSI $_{t-5}$		0.79
		(0.1)
R_BHSI $_{t-6}$		-0.3
		(0.57)
MA $_{t-1}$		0.15
		(0.87)
MA $_{t-2}$		-0.19
		(0.73)
MA $_{t-3}$		0.11
		(0.19)
MA $_{t-4}$		0.6
		(0.0002)
MA $_{t-5}$		-0.35
		(0.597)
R^2	0.789	0.7925
AIC	-7.4	-7.4
BIC	-7.39	-7.38
Durbin-Watson	2	1.99

Comparing AR(2) model to ARMA(6;5) model we can see that the latter has a higher R^2 of 79.25%, as compared to 78.9% ; same level of AIC criterion of -7.4, but higher level of BIC at -7.38 as compared to -7.39 in AR(2) . Hence, we conclude that despite including more information, ARMA(6;5) loses degree of freedom and does not add much forecasting power, as compared to a simpler AR(2). Therefore, we opt to go on with a parsimonious AR(2) model. Durbin-Watson statistic equals 2, indicating the absence of strong autocorrelation in the residuals. Therefore, we opt to go on with a parsimonious AR(2) model.

To investigate the presence of ARCH-effects we use an ARCH -LM of order 5, according to which we reject the null hypothesis of no heteroscedasticity at 5% level of significance, therefore – we assume the presence of ARCH effects.

Table 7.L: ARCH-LM test for BHSI

Variable	t-Statistic	p-value
ε_{t-1}^2	9.66	0
ε_{t-2}^2	11.42	0
ε_{t-3}^2	-3.35	0.0008
ε_{t-4}^2	3.78	0.0002
ε_{t-5}^2	4.54	0
F-statistic	79.11	0

Particularly, we see presence of ARCH-effects in first 5 orders of squared residuals' lagged variables being significant at 1% level.

We start from GARCH(1,1) specification and find that the standardized residuals appear to be non-normally distributed. The Jarque-Bera statistics strongly rejects the hypothesis of normality (Appendix B). Therefore, robust standard errors are used by selecting Bollershev-Wooldridge Heteroscedasticity consistent covariance matrix, therefore using Quasi Maximum Likelihood.

Sign and size bias test, developed by Engle and Ng (1993) joint test statistic, formulated by calculating TR^2, suggests strongly rejecting the null hypothesis of no asymmetric effects. Sign effect is significant at 10% level, while both positive and negative size effects are significant at 1% level (Appendix C).

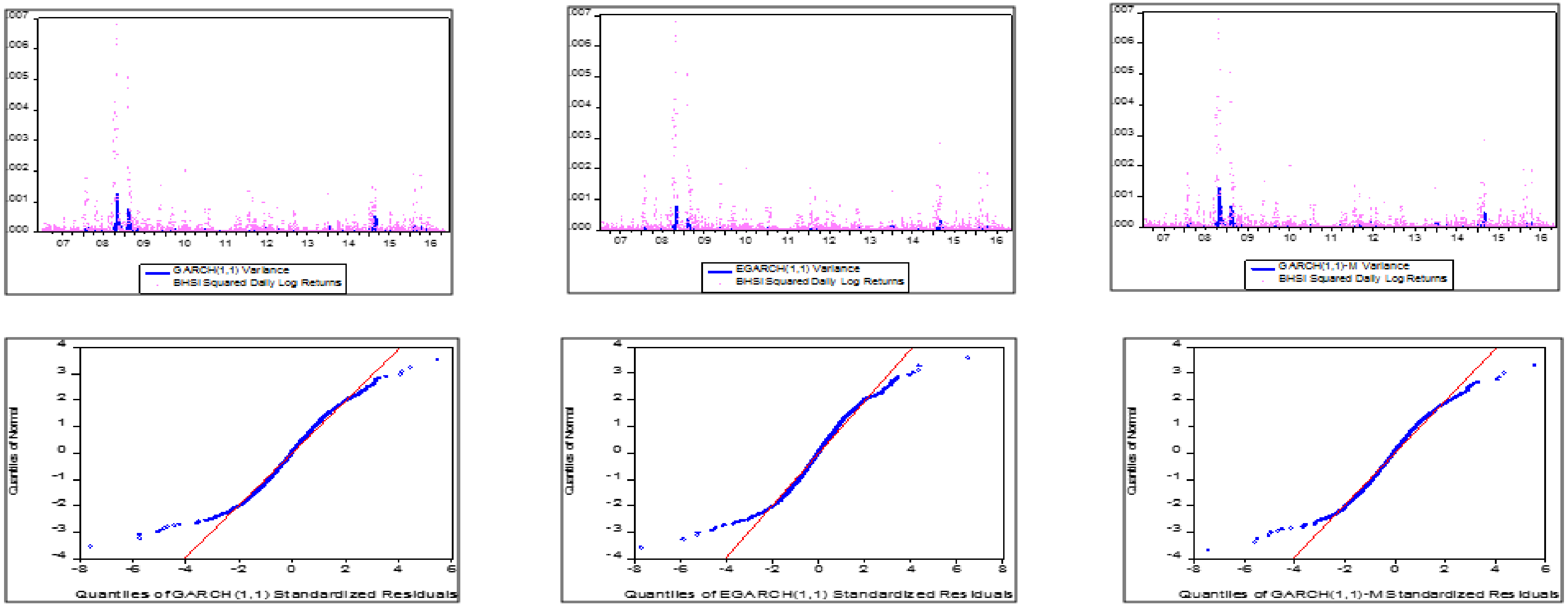
EGARCH(1;1) asymmetry term was found insignificant under the assumption of normal error distribution, however, Lu et al. (2007) suggested using GED distribution for EGARCH class models in dry bulk freight market. Under GED asymmetric effects are insignificant at 5% level, hence we proceed using symmetric models.

Compliant with the logic of Adland and Cullinane(2005) ,we test for the risk premium in the Supramax spot returns. Hence, we estimate GARCH(1;1)-M model and confirm the presence of positive risk premium. Table 7.m presents the considered models for BHSI, their information criteria and diagnostics, followed by the comparison of model forecast and realized volatility (approximated by squared daily log returns) and the QQ-plots of the standardized residuals.

Table 7.m: Results of Estimated GARCH type models for BHSI

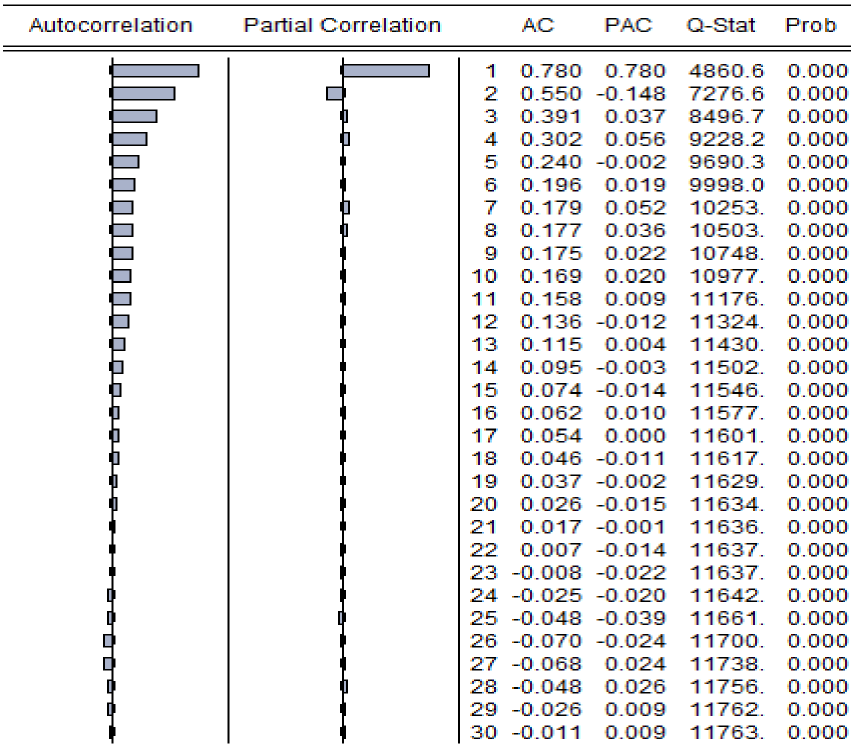
	GARCH(1,1)	EGARCH(1,1)	GARCH-M(1,1)
Mean Equation	$R\_BHI_t = -0.00068 + 0.81R\_BHI_{t-1} + 0.09R\_BHI_{t-2} + \epsilon_t$ (0.41) (0.00) (0.00)	$R\_BHI_t = -0.0012 + 0.80R\_BHI_{t-1} + 0.08R\_BHI_{t-2} + \epsilon_t$ (-0.06) (0.00) (0.00)	$R\_BHI_t = 0.10\sigma_t + 0.0031 + 0.81R\_BHI_{t-1} + 0.09R\_BHI_{t-2} + \epsilon_t$ (0.016) (0.098) (0.00) (0.00)
Variance Equation	$\sigma_t^2 = 0.0000017 + 0.16\epsilon^2_{t-1} + 0.79\sigma^2_{t-1}$ (0.01) (0.00) (0.00)	$\log\sigma^2_t = -0.79 + 0.29\frac{\lvert\epsilon_{t-1}\rvert}{\sigma_{t-1}} - 0.03\frac{\epsilon_{t-1}}{\sigma_{t-1}} + 0.94\log\sigma^2_{t-1}$ (0.00) (0.00) (0.0853) (0.00)	$\sigma_t^2 = 0.0000017 + 0.16\epsilon^2_{t-1} + 0.79\sigma^2_{t-1}$ (0.00) (0.00) (0.00)
Adj. R^2	78.92%	78.90%	79.08%
AIC	-7.75	-7.83	-7.75
BIC	-7.73	-7.82	-7.73
Durbin-Watson	2.02	2.01	2.04
ARCH-LM (5)	0.94	0.65	0.87

Notes: p-value in the parenthesis; Adj. R^2 is the corrected measure for module accuracy for the variance explained by the model inpus; AIC is the Akaike Information criteria (Akaike, 1974); BIC is the Schwarz information criteria (Schwarz, 1978); Durbin-Watson statistic tests the autocorrelation in the residuals; ARCH-LM test is Engle's test for autoregressive conditional heteroscedasticity.



7.5 Baltic Dry Index

According to correlogram and Ljung-Box Q statistics, we assume the presence of autoregressive stationary process of 2nd order – ARMA(2:0) model.



	AC	PAC	Q-Stat	Prob
1	0.780	0.780	4860.6	0.000
2	0.550	-0.148	7276.6	0.000
3	0.391	0.037	8496.7	0.000
4	0.302	0.056	9228.2	0.000
5	0.240	-0.002	9690.3	0.000
6	0.196	0.019	9998.0	0.000
7	0.179	0.052	10253.	0.000
8	0.177	0.036	10503.	0.000
9	0.175	0.022	10748.	0.000
10	0.169	0.020	10977.	0.000
11	0.158	0.009	11176.	0.000
12	0.136	-0.012	11324.	0.000
13	0.115	0.004	11430.	0.000
14	0.095	-0.003	11502.	0.000
15	0.074	-0.014	11546.	0.000
16	0.062	0.010	11577.	0.000
17	0.054	0.000	11601.	0.000
18	0.046	-0.011	11617.	0.000
19	0.037	-0.002	11629.	0.000
20	0.026	-0.015	11634.	0.000
21	0.017	-0.001	11636.	0.000
22	0.007	-0.014	11637.	0.000
23	-0.008	-0.022	11637.	0.000
24	-0.025	-0.020	11642.	0.000
25	-0.048	-0.039	11661.	0.000
26	-0.070	-0.024	11700.	0.000
27	-0.068	0.024	11738.	0.000
28	-0.048	0.026	11756.	0.000
29	-0.026	0.009	11762.	0.000
30	-0.011	0.009	11763.	0.000

Graph 7.e: Correlogram for BDI

Table 7.o: AR models specifications for BDI

Variable	**AR (2)**	**MA (2)**
Constant	-0.0000134	-0.00000171
	(0.9743)	(0.95)
R_BDI_{t-1}	0.9857	
	(0)	
R_BDI_{t-2}	-0.15	
	(0)	
MA_{t-1}		0.85
		(0)
MA_{t-2}		0.427
		(0)
R^2	0.6165	0.574
AIC	-6.5	-6.4
BIC	-6.5	-6.4
Durbin-Watson	1.988	1.77

Comparing AR(2) model to MA(2) model we can see that the latter has a lower R^2 of 57.4%, as compared to 61.65% ,both AIC and BIC are slightly higher for MA(2) as compared to AR(2) . We conclude that AR(2) is superior to MA(2), hence we proceed with manually-selected model. Durbin-Watson statistic is at 1.988, indicating a presence of small positive autocorrelation.

To investigate the presence of ARCH-effects we use an ARCH -LM of order 5, according to which we reject the null hypothesis of no heteroscedasticity at 5% level of significance, therefore – we assume the presence of ARCH effects.

Table 7.p: ARCH-LM test for BDI

Variable	t-Statistic	p-value
ε_{t-1}^2	22.71	0
ε_{t-2}^2	2.29	0.02
ε_{t-3}^2	6.61	0
ε_{t-4}^2	2.75	0.006
ε_{t-5}^2	2.91	0.0036
F-statistic	183.3	0

We see presence of ARCH-effects in first 5 orders of squared residuals' lagged variables.

We start from GARCH(1,1) specification and we find that the standardized residuals appear to be non-normally distributed. The Jarque-Bera statistics strongly rejects the hypothesis of normality (Appendix B). Therefore, robust standard errors are used by selecting Bollershev-Wooldridge Heteroscedasticity consistent covariance matrix, therefore using Quasi Maximum Likelihood.

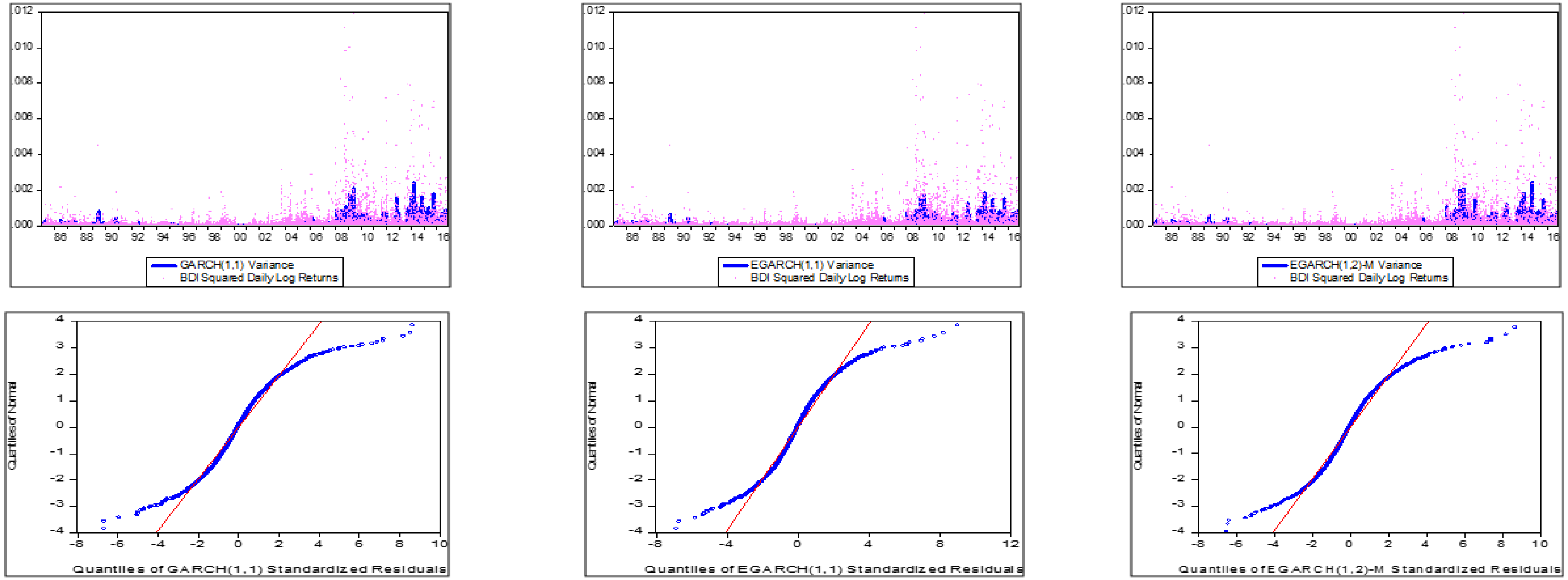
Sign and size bias test, developed by Engle and Ng (1993) joint test statistic, formulated by calculating TR^2, suggests strongly rejecting the null hypothesis of no asymmetric effects. Moreover, both size effects and sign effect appear to be significant at 1% level (Appendix C). Hence, we specify an EGARCH(1,1) model for BDI. EGARCH(1;1) model's asymmetry term was found significant at 10%, indicating presence of leverage effects.

We do not find a significant risk premium in EGARCH(1,1)-M. However, as BDI is a composite index, its risk structure might be more complicated and account for higher order information effects. Geomelos and E. Xideas (2014) used EGARCH(1;3)-M for Capesize segment, so we try increasing the order of GARCH term to capture more complicated risk-premium structure. Risk premium is positive and significant at 10 % for EGARCH(1,2)-M. Table 7.r presents the considered models for BDI, their information criteria and diagnostics, followed by the comparison of model forecast and realized volatility (approximated by squared daily log returns) and the QQ-plots of the standardized residuals.

Table 7.r: Results of Estimated GARCH type models for BDI

	GARCH(1,1)	EGARCH(1,1)	EGARCH(1,2)-M
Mean Equation	$R\_BDI_t = 0.00029 + 0.89R\_BDI_{t-1} - 0.05R\_BDI_{t-2} + \epsilon_t$ (0.28) (0.00) (0.00)	$R\_BCI_t = 0.00040 + 0.89R\_BDI_{t-1} - 0.06R\_BDI_{t-2} + \epsilon_t$ (0.16) (0.00) (0.00)	$R\_BDI_t = 0.05\sigma_t + 0.0016 + 0.89R\_BDI_{t-1} - 0.07R\_BDI_{t-2} + \epsilon_t$ (0.09) (0.05) (0.00) (0.00)
Variance Equation	$\sigma_t^2 = 0.00000027 + 0.15\epsilon^2_{t-1} + 0.87\sigma^2_{t-1}$ (0.00) (0.00) (0.00)	$\log\sigma^2_t = -0.34 + 0.30\frac{\lvert\epsilon_{t-1}\rvert}{\sigma_{t-1}} + 0.0028\frac{\epsilon_{t-1}}{\sigma_{t-1}} + 0.99\log\sigma^2_{t-1}$ (0.00) (0.00) (0.09) (0.00)	$\log\sigma^2_t = -0.42 + 0.39\frac{\lvert\epsilon_{t-1}\rvert}{\sigma_{t-1}} + 0.04\frac{\epsilon_{t-1}}{\sigma_{t-1}} + 0.50\log\sigma^2_{t-1} + 0.49\log\sigma^2_{t-}$ (0.00) (0.00) (0.04) (0.00) (0.00)
Adj. R^2	60.90%	60.96%	60.97%
AIC	-7.30	-7.31	-7.33
BIC	-7.30	-7.31	-7.32
Durbin-Watson	1.95	1.96	1.97
ARCH-LM (5)	0.00	0.00	0.00

Notes: p-value in the parenthesis; Adj. R^2 is the corrected measure for module accuracy for the variance explained by the model inpus; AIC is the Akaike Information criteria (Akaike, 1974); BIC is the Schwarz information criteria (Schwarz, 1978); Durbin-Watson statistic tests the autocorrelation in the residuals; ARCH-LM test is Engle's test for autoregressive conditional heteroscedasticity.



8 Discussion

After estimating all the models, we start the discussion with a thorough examination of the models' coefficients and various statistical measures to arrive at statistical interpretation. Based on a few criteria, for each index, the most suitable model is chosen, which allows us to analyze and compare the conditional volatility across the segments. Lastly, we compare our results with previous studies and derive economic implications.

8.1 Statistical Interpretation

8.1.1 Autoregressive Models for Dry Bulk Freight Returns

With presence of autocorrelation effects in the return series, we have found that AR(1) and AR(2) are the most suitable models to capture the dynamics of the returns. The first autoregressive term is always positive and higher than the second term, indicating the ability of operators to incorporate market information efficiently. The Adjusted R^2 for Capesize is 43%, which is much lower than for Panamax (77%), Supramax (79%) and Handysize (79%). We suspect that the underlying data-generating process for the return of Capesize is more complex than other vessels, because the prices of its carried cargoes, mainly coal and iron ore, are driven by the complicated interplay of global economy. Therefore, a simple autoregressive model is insufficient to adequately model the process.

8.1.2 GARCH Models for Dry Bulk Freight Returns

With appropriate models for the mean, the volatility can be modelled properly. We start with comparing the coefficients of GARCH(1,1) among all indices. The ARCH term for R_BCI, R_BPI, R_BSI, R_BHSI are 0.19; 0.17; 0.25 and 0.16 respectively. Supramax and Capesize have higher response to new market shocks, than Panamax and Handysize. The GARCH terms are 0.85 for R_BCI, 0.84 for R_BPI, 0.81 for BSI and 0.79 for BHSI. The degree of shock persistence decreases with the vessel size. For BDI, the ARCH term equals 0.15 and GARCH term equals 0.87. This shows that shocks have long-lasting memory, rather than being spiky for the dry bulk market. The sum of ARCH and GARCH terms is greater than 1 for all the indices, including the composite BDI, except the Handysize. This means that for BDI, Capesize, Panamax and Supramax the volatility shocks tend to strengthen, while for Handysize they will gradually weaken over time. Hence, GARCH(1,1) process is non-stationary for BDI, Capesize, Panamax and Supramax, while Handysize has unconditional variance of 0.034‰. ARCH-LM test for the GARCH residuals of Capesize and Panamax suggests rejecting the null hypothesis of no ARCH effects up to the order 5, indicating a potential model misspecification. D-W statistics of Capesize (2.25) indicates presence of small negative serial autocorrelation in GARCH residuals.

8.1.3 Risk Premium in the Dry Bulk Freight Market

Coniditonal volatility term, added to mean equations, appears to be significant and helps to explain the return of the dry bulk indices, except Panamax (p-value 0.11). The positive risk premium exists in Supramax (0.17) and Handysize (0.10), while the negative risk premium exists for Capesize (-0.07). Overall, a positive risk premium (0.05) for dry bulk market, can be identified from the result of BDI EGARCH-M specification.

8.1.4 Asymmetric Effects in the Dry Bulk Freight Market

In our EGARCH specifications we did not confirm presence of significant asymmetric effects for Panamax, Supramax and Handysize. In our EGARCH(2,1)-M specification for Capesize shocks from 2 previous periods have a strong, but opposite impact, as the 1st order of ARCH effects has a positive impact of 0.6 on the level of variance, while the 2nd order of ARCH effects has a negative impact of -0.41. The first order asymmetry term (0.1) indicates that positive shocks have higher impact than negative ones on the conditional volatility, while the second order asymmetry term (-0.1) indicates that negative shocks have higher impact than positive ones. The GARCH term (0.996) indicates high volatility persistence. EGARCH(1,2)-M, specified for R_BDI, has positive asymmetric term (0.04), meaning that positive innovations have higher impact that negative ones. The sum of GARCH terms (0.986) indicates high volatility persistence. ARCH-LM test rejects the null hypothesis of no ARCH effects up to order 5 in the residuals of both models, indicating a potential model misspecification. The D-W statistic (2.23) of EGARCH(2,1)-M model indicates the presence of small negative serial autocorrelation.

For the EGARCH models to be reliable, the conditions of stationarity should be examined. EGARCH processes for Capesize and BDI satisfy the necessary stability condition with sum of GARCH terms equal less than 1. However, Wintenberger and Cai (2011) showed that EGARCH models require continuous invertibility, similarly to an AR process, in order to estimate consistent parameters. Otherwise, the specification of EGARCH process cannot be used to reliably forecast volatility models. Winterberger (2012) developed a Stable QML for EGARCH (1,1) process, which solves the problem of stationarity and continuous invertibility. Previous studies of volatility in the Dry Bulk market had been done before the recognition of this EGARCH repercussion or simply have ignored this problem. We recognize a possible misspecification of EGARCH models due to failed sufficient conditions of stationarity, namely: possible non-existence of unconditional volatility or failed conditions of continuous invertibility. Therefore, if used for further research or practical applications, our specifications shall be tested for necessary and sufficient conditions first.

8.2 Measuring the Conditional Volatility in the Dry Bulk Freight Market

Different specifications provide different estimations of conditional volatility. We attempt to choose the most suitable model that can reflect the properties of the actual volatility best.

8.2.1 Model Selection

Following our multiple estimated specifications, we choose the most suitable models, based on these technical reasons:

- Lower information criteria indicate better trade-off between model’s goodness of fit and its simplicity;
- Considering the model’s ability to take into account the asymmetric shock effects;

- Residuals diagnostic test results for autocorrelation and ARCH-effects;
- Model's ability to track the changes in squared daily log returns, as the proxy for realized conditional volatility;

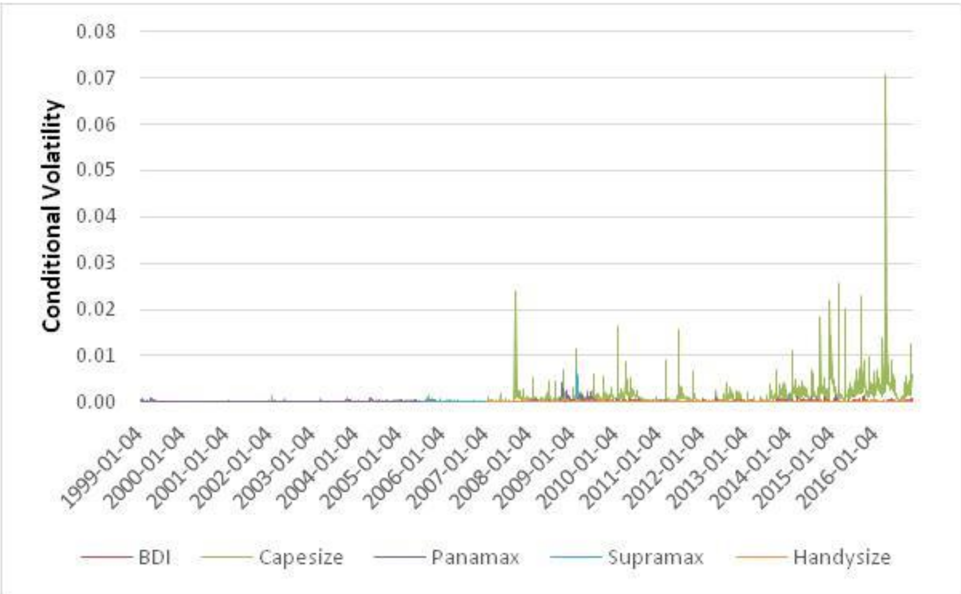
Table 8.a presents the most suitable model for each index.

Table 8.a: Selected Models for the Baltic Dry Indices

Index	Conditional Mean	Conditional Variance
BCI	AR(2)	EGARCH (2;1)-M
BPI	AR(2)	GARCH(1;1)
BSI	AR(1)	GARCH (1;1)-M
BHSI	AR(2)	GARCH (1;1)-M
BDI	AR(2)	EGARCH (1;2)-M

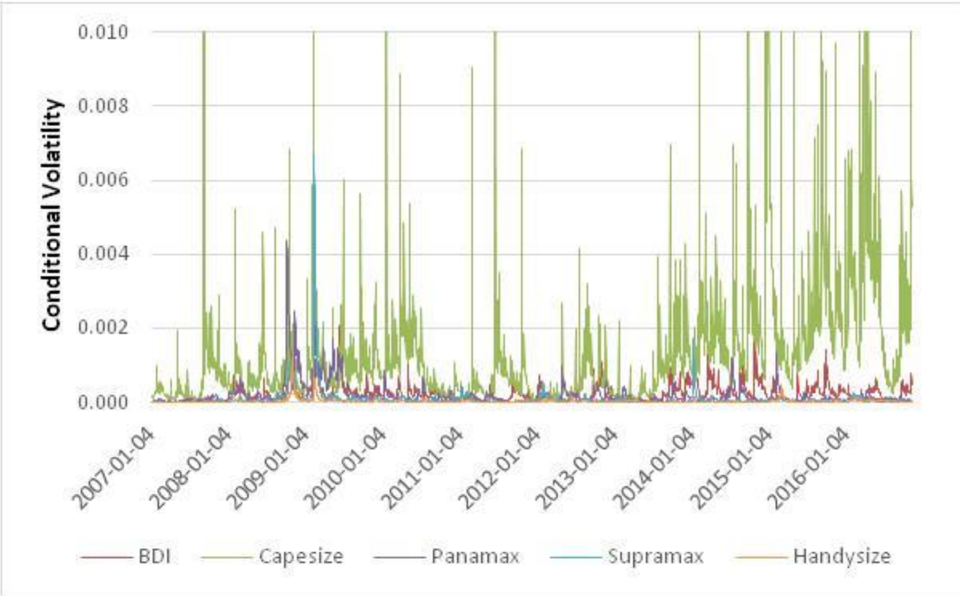
8.2.2 Comparison of Conditional Volatilities in the Dry Bulk Freight Market

Graph 8.a displays the conditional volatility of each segment.



Graph 8.a: Conditional Volatility in the Dry Bulk Market

Before 2007, volatility is low across the segments. From mid-2007, volatility starts to rise and becomes unpredictable. The first huge spike occurred in the end of 2007 and was followed by smaller but more frequent spikes in the next 9 years. In the last three years the volatility of Capesize has been extremely wild with the highest jump in April 2016. To look into the difference between segments, we focus on the period between 2007 and 2016 with graph 8.b.



Graph 8.b: Conditional Volatility during 2007-2016 in the Dry Bulk Market

Capesize has the highest volatility in every period. During the 2008 financial crisis volatility of all the segments radically increased. Before 2015, when the volatility of Capesize deviates from other segments', it usually converges rapidly in the following days. However, after 2015, this pattern does not seem to be present and the difference continues to grow. This is likely to be explained by the shift in operators' behavior due to realized differences in market segments: risk-averse investors try to use smaller vessels to decrease their idiosyncratic risk, while risk-loving investors use Capesize more to get higher potential gains. Likewise, according to Marlow et al. (2008), the degree of concentration of Capesize owners has been rising in recent years which may also explain the increasing volatility. Overall volatility is higher for larger vessels and is lower for smaller vessels, which confirms the conclusions of Kavussanos (1996) and Marlow et al. (2008). Smaller vessels are more versatile and therefore less affected by short-term shocks in demand. Larger vessels are constrained by their technical profile that makes it difficult for them to accommodate different cargoes, routes and ports.

8.3 Estimated versus Realized Volatility in the Dry Bulk Freight Market

One of the missing points in the previous studies is the forecast quality of the estimated models. We decide to analyze how the estimated volatility reflects the realized volatility in the dry bulk freight market, which is approximated through the squared daily log returns. Although one should expect GARCH type models to underestimate market shocks, it is useful to learn the degree of underestimation and the differences between vessel segments. Table 8.b represents the quantitative summary of differences between estimated and realized volatilities.

	BCI	BPI	BSI	BHSI	BDI
Overestimation %	58%	37%	41%	37%	50%
Average Overestimation	0.00062	0.0001	0.000049	0.000019	0.000066
Underestimation %	42%	63%	59%	63%	50%
Average Underestimation	0.0018	0.00069	0.00032	0.000224	0.00034

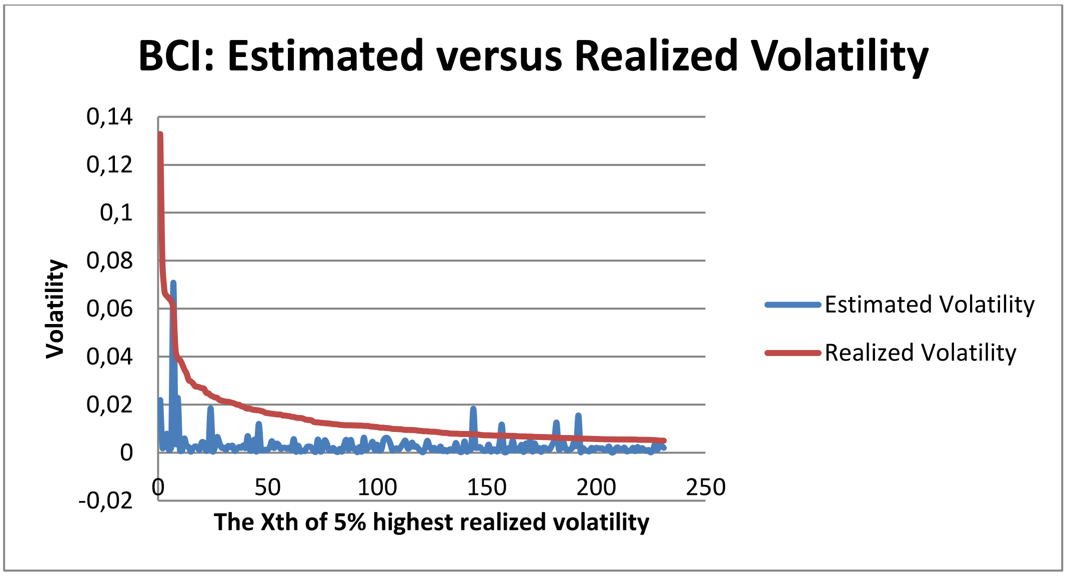
Table 8.b: Differences between Estimated and Realized Volatility

In 58% of the days in the sample period, our model overestimates the realized volatility for Capesize index, but for all other segments our models generally underestimate the realized volatility. Moreover, we notice that the degree of underestimation is substantially higher for all the indices, including BDI. In addition, from tables 7.d, 7.g, 7.j, 7.m and 7.r, we observe the constant underestimation of realized volatility in cases when it is very high. The models seem to be unable to provide reliable forecasts when realized volatility suddenly increases. As extreme cases often cause extreme losses, it is of great interest to know how our models perform when the volatility is highest. Therefore, we perform a more detailed analysis in the cases of 5% and 1% of highest realized volatility. Table 8.c shows the underestimation of realized volatility across the indices:

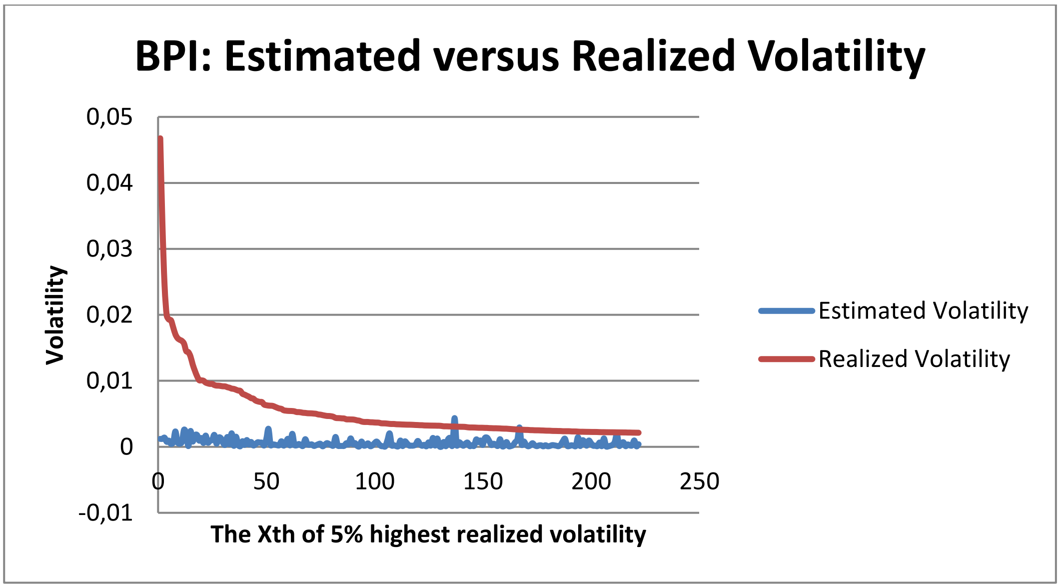
	BCI	BPI	BSI	BHSI	BDI
Average Underestimation					
5% extreme cases	0.011	0.0047	0.0022	0.0017	0.0022
1% extreme cases	0.027	0.0119	0.0048	0.0047	0.0055

Table 8.c: Underestimation of Realized Volatility in Extreme Cases

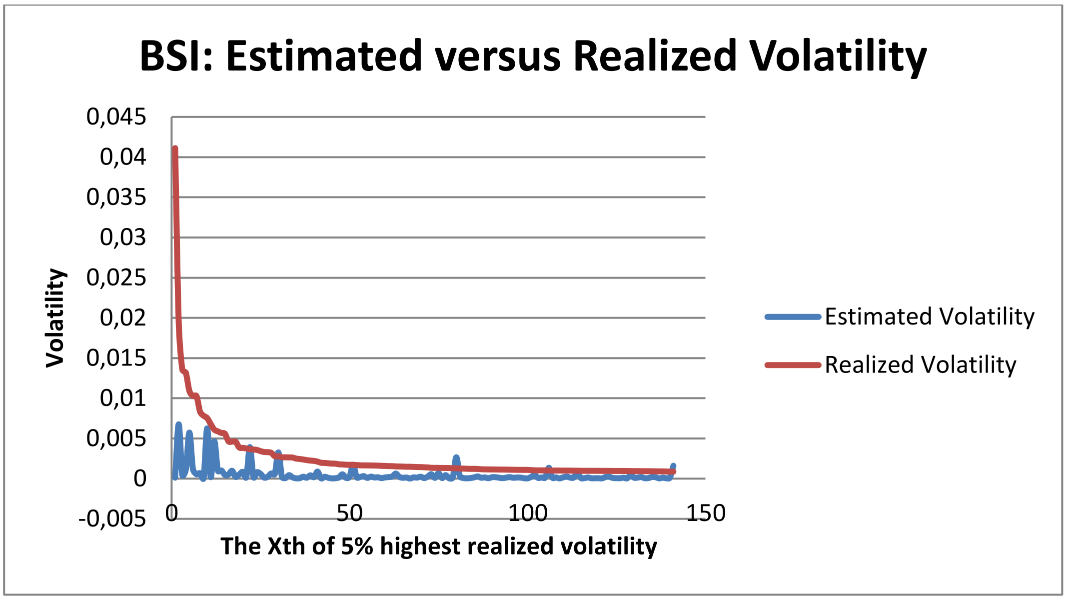
We observe a strong pattern: when realized volatility increases, the degree of underestimation of all the models also increases. Furthermore, the degree of underestimation systematically increases with the size of the vessel. The degree of underestimation is the most severe for Capesize. Graphs 8.c - 8.g display the top 5% of realized volatility and their corresponding values of estimated volatility for all the indices:



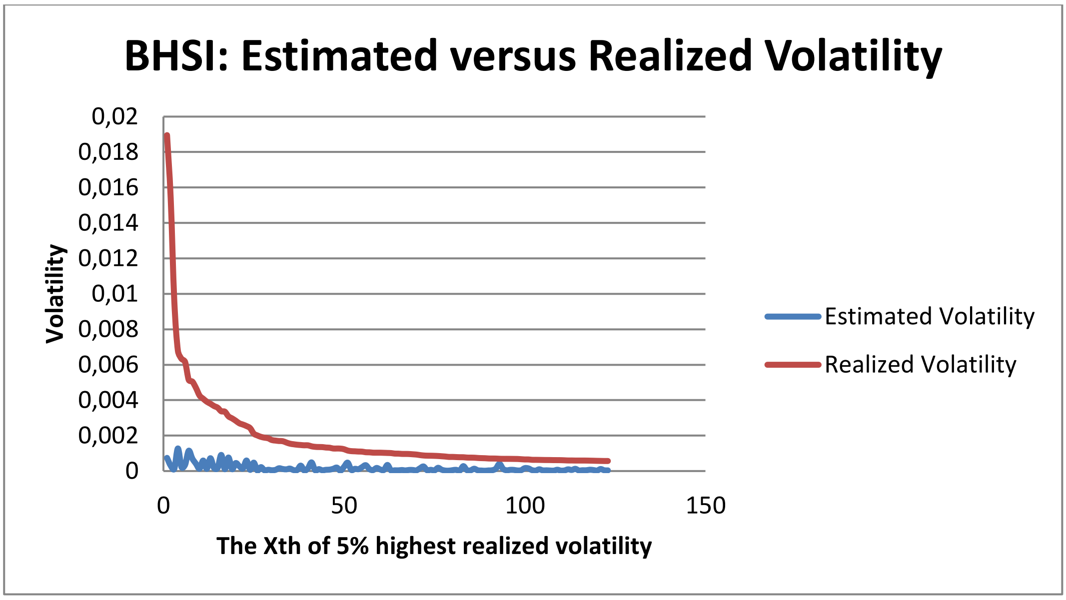
Graph 8.c: BCI: Estimated versus Realized Volatility



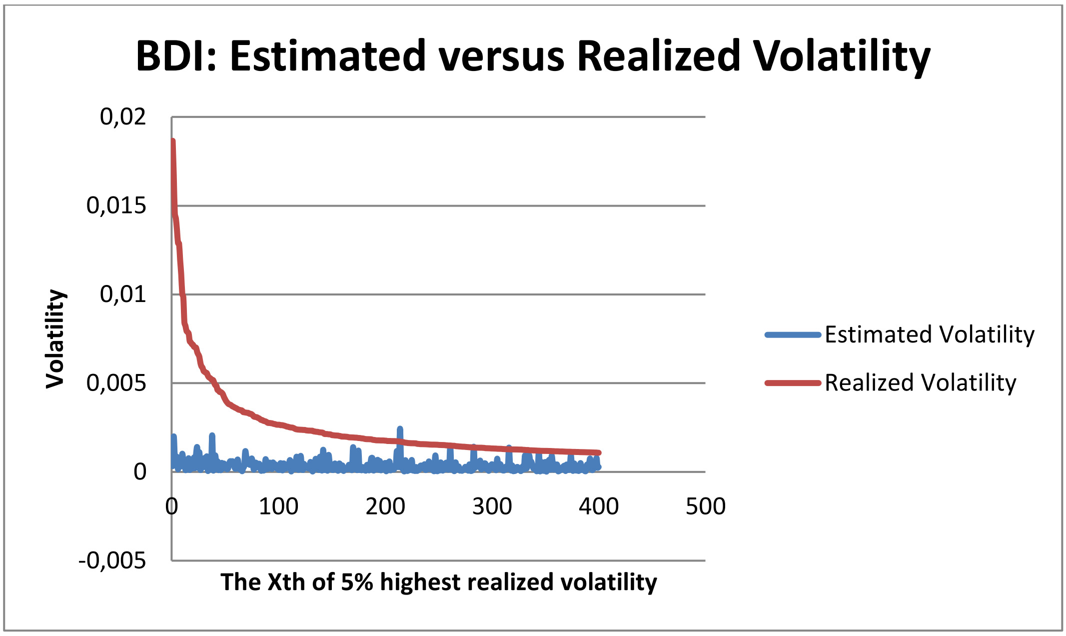
Graph 8.d: BPI: Estimated versus Realized Volatility



Graph 8.e: BSI: Estimated versus Realized Volatility



Graph 8.f: BHSI: Estimated versus Realized Volatility



Graph 8.g BDI: Estimated versus Realized Volatility

From the graphs, it is clear that our models consistently underestimate volatility in the 5% of most extreme cases. The degree of underestimation continuously grows, as the realized volatility peaks. Although all the models seem to perform satisfactorily between 5% and 1% of extreme cases, they cannot reliably forecast the volatility in the 1% of the most uncertain outcomes. We conclude that our models have a serious limitation: they are not able to capture and reliably predict extreme market outcomes. Therefore, for risk-management purposes, our models can provide a floor of expected uncertainty, but not its ceiling. For bigger vessels, this floor is likely to be further away from the realized volatility.

8.4 Economic Interpretation

8.4.1 Impact of Vessel Versatility

Our GARCH (1,1) specifications confirm the findings of Marlow et al. (2008), Xu et al. (2011), and Geomelos (2014) that for Capesize, Panamax, Supramax, and BDI, the GARCH process is non-stationary. But we find that for Handysize the GARCH process is stationary. We attribute this less complicated dynamics of the smallest vessels to a more stable demand and supply relationship that can be modeled and forecast more easily. The ARCH term of Handysize is the smallest which contradicts with Kavussanos (1996), Marlow et al. (2008) but confirms with

Xu (2011) and Drobetz (2012) where they found smaller vessels (Panamax) have lower ARCH term than larger vessels (Capesize). We believe there could two explanations: first, the market for Handysize is more competitive because of the large number of vessels that compete in the same market and market information is incorporated very efficiently. The immediate impact from market shocks for the next period is not as big; second, the market may have incorporated new information after the 2008 financial crisis. Therefore, models in earlier papers may not be adequate for today's market. Volatility persistence is high across the segments which indicates the market has a long memory for risks. This can be explained by the economic nature of dry bulk market. In the short-term, the risk factors, relevant for the market, cannot be influenced or eliminated by the market players, because of the global market structure and its reliance on economic conditions. By the same logic, smaller vessels have lower shock persistence, because they are more versatile and can adjust their positions sooner to eliminate the risk. This is shown in our findings: the smaller the vessel size, the smaller the GARCH terms.

8.4.2 Higher Volatility on Positive News and Short-term Optimism

From our EGARCH specifications, we find that Capesize volatility has contrarian dynamics, where big increase in volatility is likely to be followed by a lower volatility. Opposite signs of two sign asymmetry terms indicate that the market is subject to short-term optimism: the first day the news come out the market tends to overestimate good news, but underestimate bad ones, but realizes its misevaluation next day and corrects for irrational behavior. Overall, positive innovations have higher (asymmetric) impact on the volatility of the dry bulk market (represented by BDI), as compared to negative innovations. This empirical result does not seem to be consistent with the empirical result of Marlow et al. (2008) and Drobetz et al. (2012), who studied Capesize, Panamax and Handysize market. However, it provides support for the Drobetz's theoretical assumptions of positive asymmetric effects, implied from the short-term supply and demand model of Stopford (2009). In the short-run, new vessels cannot be added to the market to react to the increased demand in a tight market. This is why volatility is impacted more by positive shocks.

8.4.3 Higher Risk Leads to Higher Return, except Capesize

Positive risk premium is found in Supramax, Handysize, and the overall market (BDI). This could be explained by the adjustment of risk-averse operators that seek to exit the spot market by chartering out their ships at a lower rate for a longer period of time. Risk-loving operators who remain in the spot market could enjoy higher spot prices because of lower supply in the spot market. Negative risk premium is found in Capesize. This could be explained by the nature of supply and demand in this segment. Capesize carries more specific cargoes and sail to more specific ports. The vessels cost more capital to be operated and therefore put more pressure on operators to find employment for them. When the spot market gets volatile, operators chase limited cargoes by discounting their prices and their returns drop.

9 Conclusions

Dry bulk freight market has gone more volatile in recent years. From the largest Capesize type of vessels to the smallest Handysize, each segment has its unique trading activities and exhibit different market movements. The increasing volatility not only threats market participants' bottom line, but also posts intriguing challenges for academic research.

This study analyzes the volatility in the dry bulk freight spot market. Indices published by Baltic Exchange from 1985 to 2016 are used as the proxy for the dry bulk freight of vessels in different sizes. In previous studies, many questions were raised about the stationarity of dry bulk freight. We find levels of dry bulk freight to be non-stationary, but their return series to be stationary. Autocorrelation and volatility clustering are observed across all dry bulk segments. This motivates us to model the volatility with AR-GARCH type specifications. We find first and second order autoregressive process very suitable to explain the return dynamics.

We find return series of dry bulk freight market spot prices to exhibit conditional heteroscedasticity. Our results confirm that GARCH class models are suitable for explaining the volatility of the dry bulk freight market. Adding risk premium into the model helps to capture volatility dynamics. When modelling Capesize volatility, higher order models shall be considered due to the increased complexity of its economic nature.

GARCH specification shows that for all types of vessels, shocks are highly persistent. Degree of persistence decreases with the vessel size, reflecting differences of self-memory in each segment. In line with most of recent studies, our GARCH (1,1) processes are non-stationary with the exception of Handysize. This means shocks strengthen over time which contradicts common economic intuition and implies that GARCH (1,1) process is not sufficient to capture the volatility dynamics in the dry bulk market.

Motivated by leverage effect in the equity market and short-term supply and demand model in the dry bulk market, we use EGARCH specification to capture asymmetric effect in volatility. We do not find asymmetric effect in Panamax, Supramax, and Handysize, but find contrarian asymmetric dynamics in Capesize, indicating constant bias of overoptimism among operators. Positive asymmetric effect is found in BDI, indicating higher volatility response in an upward dry bulk freight market.

In finance, risk is often found to explain return. We find positive risk premium in Supramax, Handysize, and BDI. It shows that in those markets, operators who stay in the spot market during the time of high uncertainty are rewarded by higher returns. In Capesize we find negative risk premium which might seem unusual. However, when talking into account the inflexible nature of Capesize vessels, negative risk premium may simply reflect the discounts offered by operators who seek to avoid high costs of unemployment.

Capesize has the highest conditional volatility among all segments followed by Panamax. Before 2015, volatility of Capesize converges rapidly to the level of other indices. After 2015, we observe that this difference increases, indicating that Capesize is getting riskier. It may be explained by a shift in the strategy among operators who have different risk appetite after the turbulent market during the 2008 financial crisis. Risk-loving operators focus on Capesize while risk-averse operators avoid Capesize.

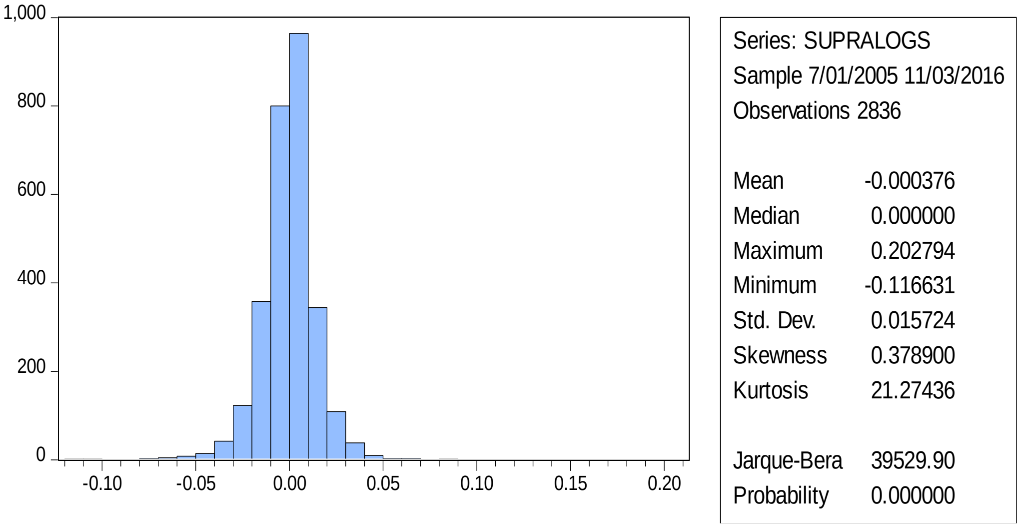
Our study has different implications for participants in the dry bulk freight market. When shipowners add vessels to their fleet, size does matter in terms of risk and return. Investing in Capesize gets continuously riskier compared to other segments. Adding smaller vessels into their portfolio could decrease risks. When the market is going up, operators should expect higher uncertainty and position themselves accordingly. Cargo owners who need to charter Supramax and Handysize vessels should expect to pay a higher freight when the market is riskier while the opposite holds for Capesize.

Overall, this study contributes to understanding of the uncertainty in the dry bulk freight spot market and provides results of various model specifications of each dry bulk segment. There are four main limitations for our study. First, the index could be misrepresentative for the real conditions of a local market, therefore using indices as proxies may not yield the most accurate forecast with any models. Second, for BCI, BPI and BDI we use data that span about 20 years. When their levels and returns are observed across the whole period, the movements appeared to behave differently after the 2008 financial crisis. It is also important to note that changes in technology and regulations over the past twenty years could have shifted the trading fundamentals. As a result, if the underlying data generating process has changed along the period, the results, our models provide, may not be appropriate for future applications. Third, non-stationarity of GARCH models and potential continuous non-invertibility of exponential GARCH models may indicate their misspecification. Finally, in general, our models underestimate the conditional volatility during the periods of extremely high market uncertainty (e.g. average 0.027 underestimation in 1% highest volatility for Capesize). This problem gets more pronounced as the vessel size increases. Therefore, our models may not be an effective tool to help market participants forecast risks in an extremely volatile market.

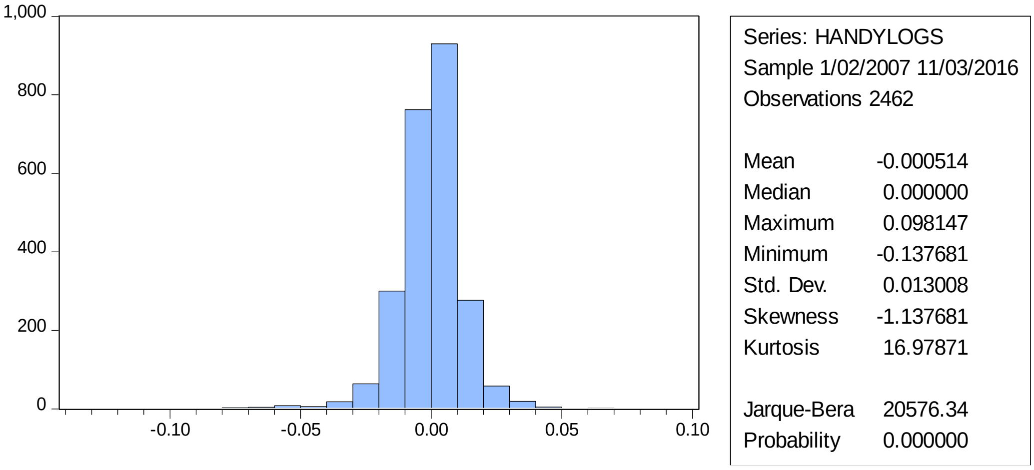
Our suggestions of further research include the following. First, the drivers behind the extremely high volatility in the dry bulk freight spot market should be researched to provide better fundamentals for model specifications. Second, more profound analysis of stationarity properties of GARCH type models shall be conducted, so their results can be relied on. Third, the exogenous variables, relevant to the dry bulk market, may be considered when modelling both the returns and their volatility. Fourth, in practice, the forward market is as important as the spot market. Therefore, similar research shall be done to complete understanding of the dry bulk freight market.

10 Appendices

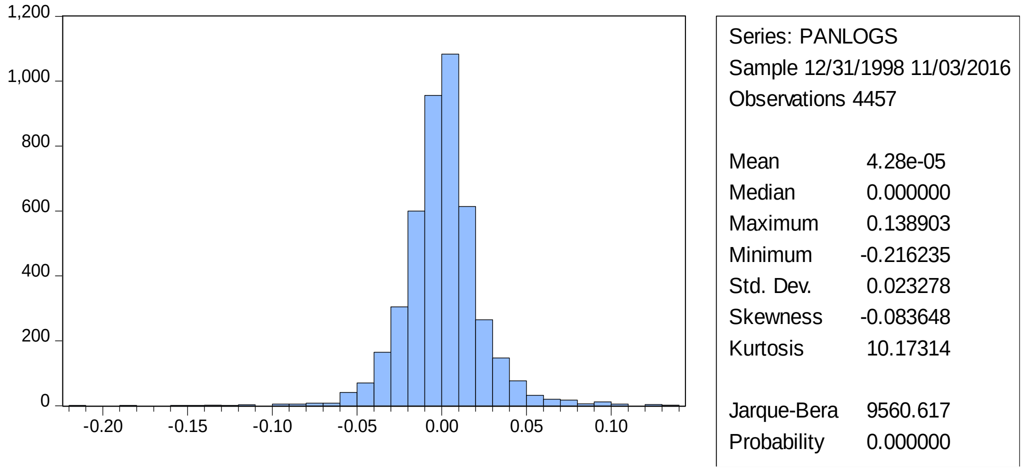
10.1 Appendix A. Histograms of Log Return Distributions



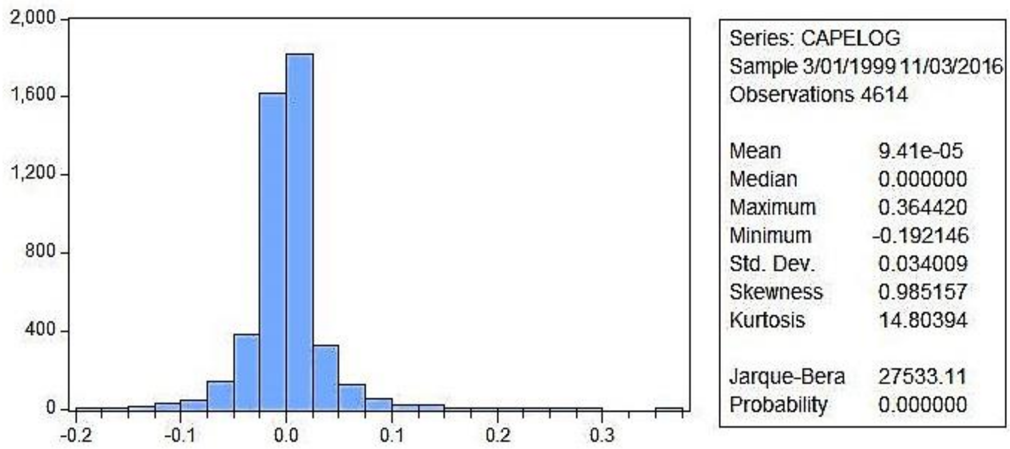
Graph 10.a: Histogram of Supramax Log Return Distributions



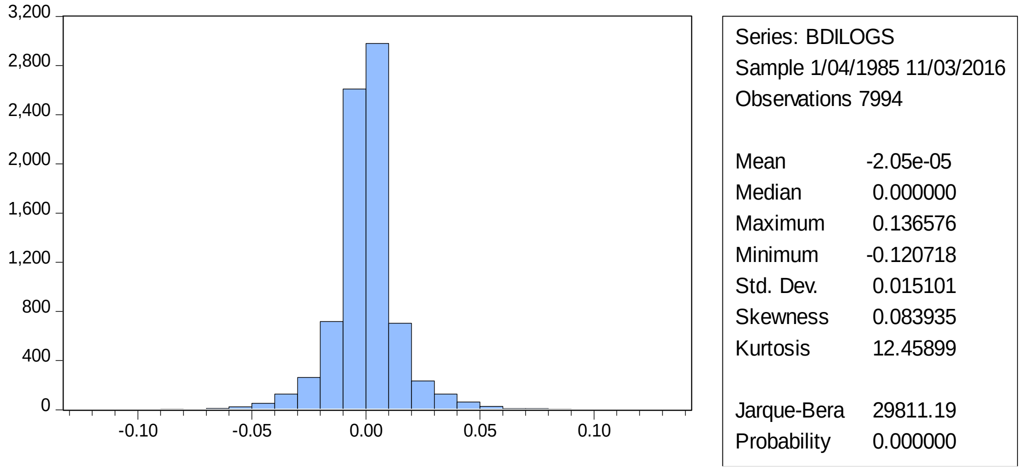
Graph 10.b: Histogram of Handysize Log Return Distributions



Graph 10.c: Histogram of Panamax Log Return Distributions

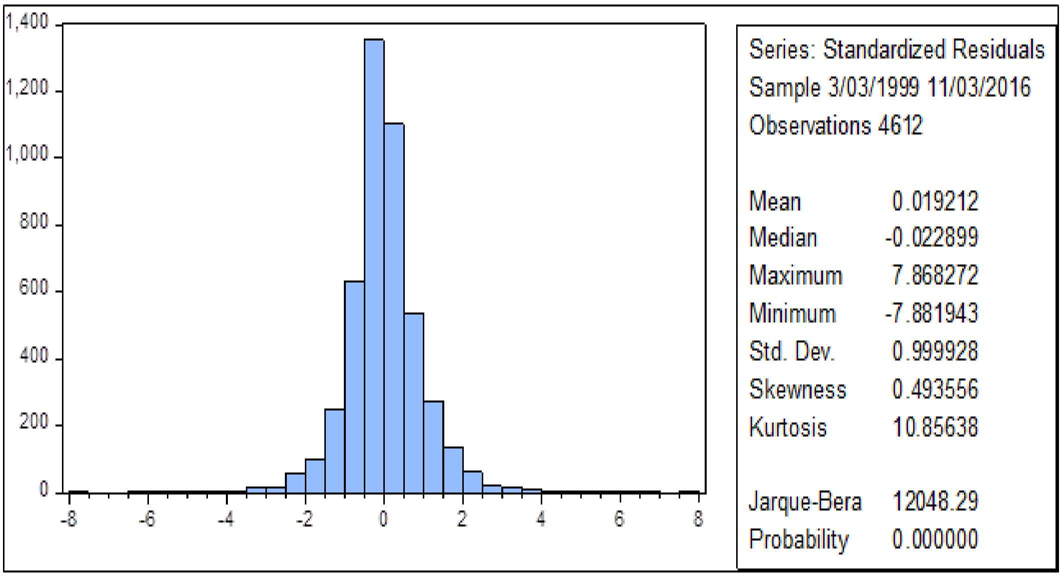


Graph 10.d: Histogram of Capesize Log Return Distributions

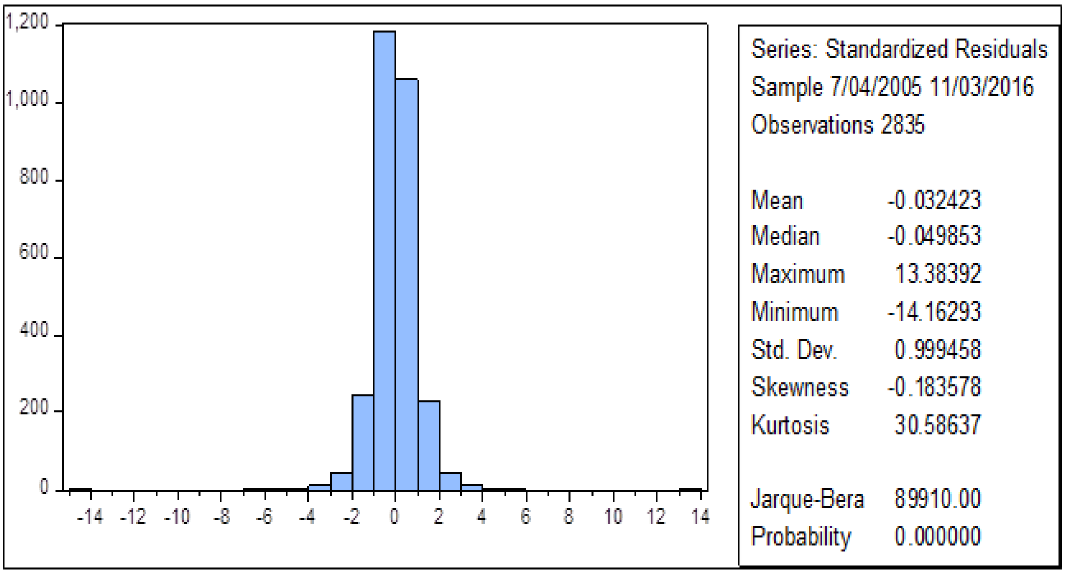


Graph 10.e: Histogram of BDI Log Return Distributions

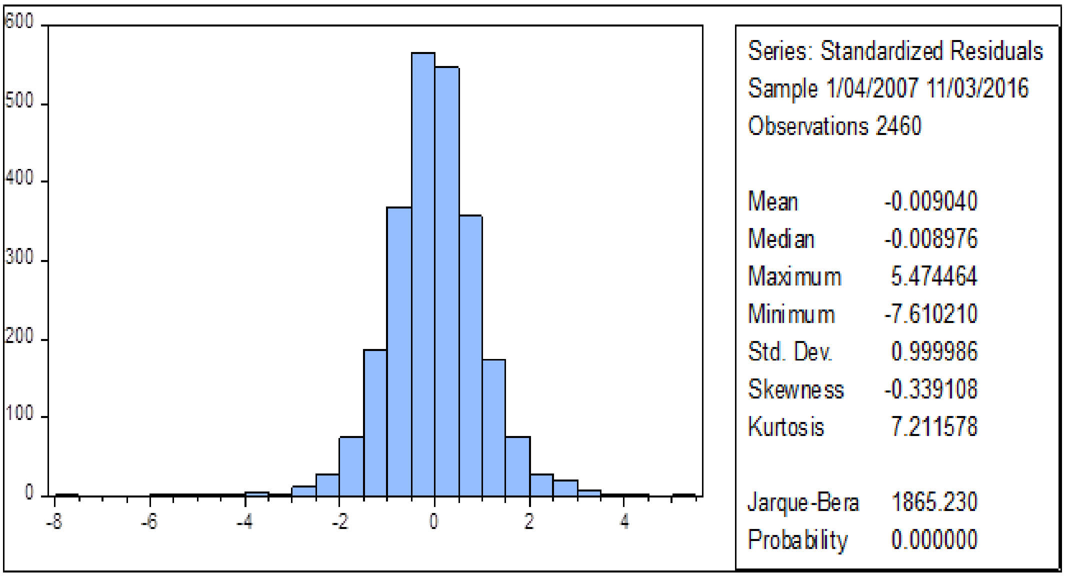
10.2 Appendix B. Standardized Residuals Distribution



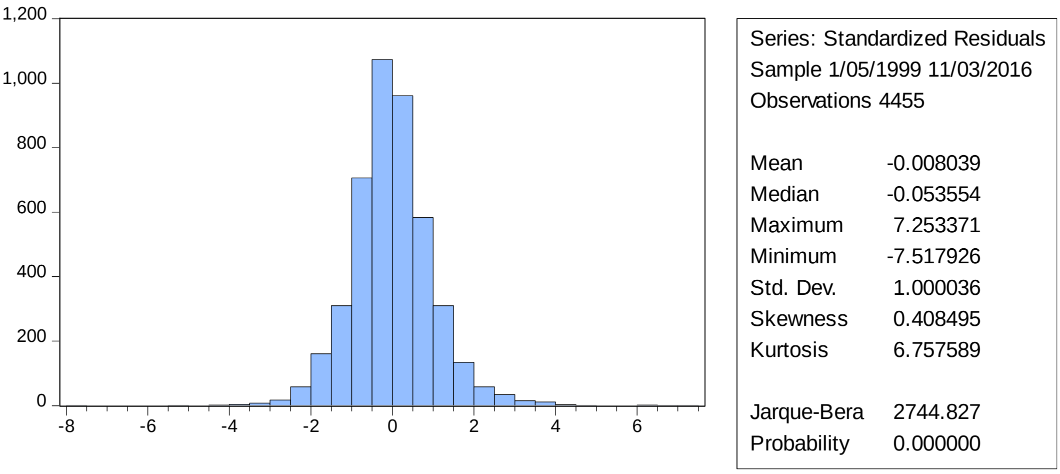
Graph 10.f: Standardized residuals of BCI



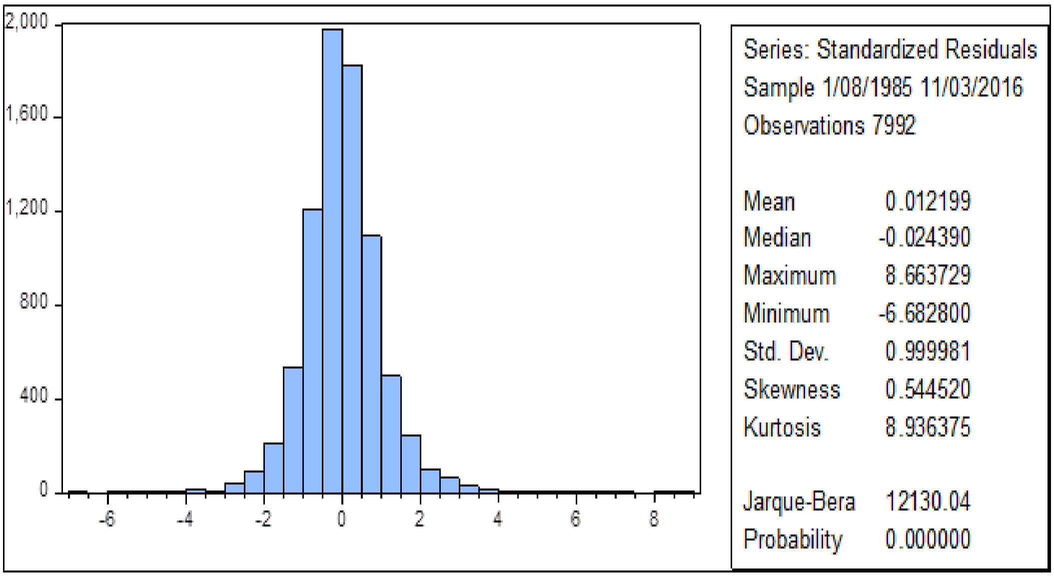
Graph 10.g: Standardized residuals of BSI



Graph 10.h: Standardized residuals of BHSI



Graph 10.i: Standardized residuals of BPI



Graph 10.j: Standardized residuals of BDI

10.3 Appendix C. Sign and Size Bias Test Results

Table 10.a: GARCH(1;1) size and sign bias test for R_BCI

Variable	Coefficient	t-statistic	p-value
C	-0.000174	-2.54	0.011
Sign	0.000287	3.06	0.0022
Negative size bias	-0.033	-12.6	0
Positive size bias	0.066	25.91	0
LM Test	***	710.65	0

Table 10.b: GARCH(1;1) size and sign bias test for R_BSI

Variable	Coefficient	t-statistic	p-value
C	0.000019	1.08	0.28
Sign	-0.0000128	-0.51	0.61
Negative size bias	-0.0093	-3.4	0.0007
Positive size bias	0.009	4.07	0
LM Test	***	28.38	0.00000301

Table 10.c: GARCH(1;1) size and sign bias test for BHSI

Variable	Coefficient	t-statistic	p-value
C	0.00000278	0.47	0.63
Sign	-0.0000139	-1.67	0.0957
Negative size bias	-0.0123	-11.96	0
Positive size bias	0.008226	-11.96	0
LM Test	***	201.77	0

Table 10.d: GARCH(1;1) size and sign bias test for BPI

Variable	Coefficient	t-statistic	p-value
C	-0.00000461	-0.39	0.6975
Sign	0.00000858	0.5186	0.6041
Negative size bias	-0.01473	-13.5	0
Positive size bias	0.02	19.62	0
LM Test	***	510.46	0

Table 10.e: GARCH(1;1) size and sign bias test for BDI

Variable	Coefficient	t-statistic	p-value
C	0.0000165	2.39	0.0168
Sign	-0.0000276	-2.87	0.004
Negative size bias	-0.016	-22.22	0
Positive size bias	0.015	21.51	0
LM Test	***	865.38	0

11 Bibliography

Adland, R., & Cullinane, K. (2005). A time-varying risk premium in the term structure of bulk shipping freight rates. *Journal of Transport Economics and Policy (JTEP)*, *39*(2), 191-208.

Adland, R., & Strandenes, S. P. (2007). A discrete-time stochastic partial equilibrium model of the spot freight market. *Journal of Transport Economics and Policy (JTEP)*, *41*(2), 189-218.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716-723.

Alizadeh, A. H., & Nomikos, N. K. (2011). Dynamics of the term structure and volatility of shipping freight rates. *Journal of Transport Economics and Policy (JTEP)*, *45*(1), 105-128.

Alizadeh, A., & Nomikos, N. (2009). *Shipping derivatives and risk management*. Palgrave Macmillan.

Beenstock, M., & Vergottis, A. (1989). An econometric model of the world tanker market. *Journal of Transport Economics and Policy*, 263-280.

Beenstock, M., & Vergottis, A. (1989). An econometric model of the world market for dry cargo freight and shipping. *Applied Economics*, *21*(3), 339-356.

Beenstock, M., & Vergottis, A. (1993). The interdependence between the dry cargo and tanker markets. *Logistics and Transportation Review*, *29*(1), 3.

Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, *13*(1), 1-42.

Black, F. (1976). Stuedies of stock price volatility changes.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307-327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, 542-547.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric reviews, 11(2), 143-172.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric reviews, 11(2), 143-172.

Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control, revised ed*. Holden-Day.

Brooks, C. (2014). *Introductory econometrics for finance*. Cambridge university press.

Campbell, J. Y., & Shiller, R. J. (1987). Cointegration and tests of present value models. *Journal of political Economy*, *95*(5), 1062-1088.

Campbell, J. Y., & Shiller, R. J. (1991). Yield spreads and interest rate movements: A bird's eye view. *The Review of Economic Studies*, *58*(3), 495-514.

Canova, F., & Hansen, B. E. (1995). Are seasonal patterns constant over time? A test for seasonal stability. *Journal of Business & Economic Statistics*, *13*(3), 237-252.

Cullinane, K. (1995). A portfolio analysis of market investments in dry bulk shipping. *Transportation Research Part B: Methodological*, *29*(3), 181-200.

de Goeij, P., & Marquering, W. (2006). Macroeconomic announcements and asymmetric volatility in bond returns. *Journal of Banking & Finance*, *30*(10), 2659-2680.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), 427-431.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 1057-1072.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 1057-1072.

Drobetz, W., Richter, T., & Wambach, M. (2012). Dynamics of time-varying volatility in the dry bulk and tanker freight markets. *Applied financial economics*, *22*(16), 1367-1384.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The journal of finance*, *48*(5), 1749-1778.

Engle, Robert F., David M. Lilien, and Russell P. Robins. "Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model." Econometrica 55, no. 2 (1987): 391-407. doi:10.2307/1913242.

Fuller, W. A. (1976). *Introduction to statistical time series*. Wiley, New York

Gao, Y., Zhang, C., & Zhang, L. (2012). Comparison of GARCH Models based on Different Distributions. JCP, 7(8), 1967-1973.

Geomelos, N. D., & Xideas, E. (2014). Forecasting spot prices in bulk shipping using multivariate and univariate models. Cogent Economics & Finance, 2(1), 923701.

Glen, D. R. (1997). The market for second-hand ships: Further results on efficiency using cointegration analysis. *Maritime Policy and Management*, *24*(3), 245-260.

Glen, D. R. (2006). The modelling of dry bulk and tanker markets: a survey.*Maritime Policy & Management*, *33*(5), 431-445.

Glen, D. R., & Rogers, P. (1997). Does weight matter? A statistical analysis of the SSY Capesize index. *Maritime Policy and Management*, *24*(4), 351-364.

Gratsos, G. A., Thanopoulou, H. A., & Veenstra, A. (2012). Dry bulk shipping.*The Blackwell Companion to Maritime Economics*, *11*, 187.

Gray, J. W. (1987). Futures and options for shipping. Colchester: Lloyd's of London Press.

Hale, C., & Vanags, A. (1989). Spot and period rates in the dry bulk market: Some tests for the period 1980-1986. *Journal of Transport Economics and Policy*, 281-291.

Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 190-195.

Hawdon, D. (1978). Tanker freight rates in the short and long run. *Applied Economics*, *10*(3), 203-218.

Jing, L., Marlow, P. B., & Hui, W. (2008). An analysis of freight rate volatility in dry bulk shipping markets. *Maritime Policy & Management*, *35*(3), 237-251.

Kavussanos, M. G. (1996). Comparisons of volatility in the dry-cargo ship sector: Spot versus time charters, and smaller versus larger vessels. *Journal of Transport economics and Policy*, 67-82.

Kavussanos, M. G. (1997). The dynamics of time-varying volatilities in different size second-hand ship prices of the dry-cargo sector. *Applied Economics*, *29*(4), 433-443.

Kavussanos, M. G., & Alizadeh-M, A. H. (2001). Seasonality patterns in dry bulk shipping spot and time charter freight rates. *Transportation Research Part E: Logistics and Transportation Review*, *37*(6), 443-467.

Kavussanos, M. G., & Alizadeh-M, A. H. (2002). The expectations hypothesis of the term structure and risk premiums in dry bulk shipping freight markets.*Journal of Transport Economics and Policy (JTEP)*, *36*(2), 267-304.

Kavussanos, M. G., & Nomikos, N. K. (1999). The forward pricing function of the shipping freight futures market. *journal of futures markets*, *19*(3), 353-376.

Koutmos, G., & Booth, G. G. (1995). Asymmetric volatility transmission in international stock markets. *Journal of international Money and Finance*, *14*(6), 747-762.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business 36 (Oct. 1963)*, 394-419

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the Econometric Society, 347-370.

Osborn, D. R. (1990). A survey of seasonality in UK macroeconomic variables. *International Journal of Forecasting*, *6*(3), 327-336.

Osei-Wusu, E. (2011). Relationship Between Return, Volume and Volatility in the Ghana Stock Market (Doctoral dissertation, Thesis. Departmen of Financial and Statistics, Hanken School of Econometrics).

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of econometrics*, *45*(1), 267-290.

Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 335-346.

Randers, J., and U. Göluke. (2007). Forecasting Turning Points in Shipping Freight Rates: Lessons from 30 Years of Practical Effort.

Rossi, E. (2010). Univariate GARCH models: a survey (in Russian). Quantile, (8), 1-67.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461-464.

Stopford, M. (2009). *Maritime Economics 3e*. Routledge.

Tvedt, J. (2003). A new perspective on price dynamics of the dry bulk market. *Maritime Policy & Management*, *30*(3), 221-230.

UNCTAD (2015). *Review of Maritime Transport, 2015* (No. UNCTAD/RMT/2015).

Veenstra, A. W. (1999). The term structure of ocean freight rates. *Maritime Policy & Management*, *26*(3), 279-293.

Wintenberger Olivier. Continuous invertibility and stable QML estimation of the EGARCH(1,1) model. 2012.

Wintenberger, O., & Cai, S. (2011). Parametric inference and forecasting in continuously invertible volatility models. arXiv preprint arXiv:1106.4983.

Xu, J. J., Yip, T. L., & Marlow, P. B. (2011). The dynamics between freight volatility and fleet size growth in dry bulk shipping markets. *Transportation research part E: logistics and transportation review*, *47*(6), 983-991.