Baltic Dry Index:
Modeling and predicting long term relationships with macroeconomic, financial, demand and supply factors.

Xenofon Efthimiou

A Thesis submitted

to the Secretariat of the MSc in

of the Athens University of Economics and Business

as partial fulfillment of the Requirements for the

Master's Degree

In Financial Mathematics (Μαθηματικά της Αγοράς και Παραγωγής)

Athens

{01,October, 2020}

Map.aueb.gr

Contents

Abstract

This thesis is a study of the dry bulk freight rates, particularly the BDI as its index, and the effect of supply and demand factors in the long run in order to be able to make estimations regarding future values of the market. Therefore, we used the BDI as in independent variable which was calculated quarterly and added into the model various independent variables, according to economic theory and literature, in order to investigate how these will affect the BDI using a multivariate time series model (VECM) and the Johansen approach. We examined the significance of each variable and the qualitative and quantitative effect and the impact that they might have in the freight market. The results of this thesis are for the most part in line with earlier papers and studies; however some relationships might differ because of the different and more recent study period that goes up to 2014.

It was proven that, there is in fact a long run relationship between the BDI and some key independent factors like the fleet size development or the World Trade Volume (which is considered to measure the demand for shipping services and it is a major demand indicator in shipping). Both of these factors and others had a significant effect in the long run but found no significant relationship in the short run meaning that the dependent variable, the BDI index, is indeed affected by the examined factors along with their past values.

A literature review of the dry bulk and the freight market was examined and more importantly, the models that were employed for this investigation. Additionally, an overview of the dry bulk market is described and the specific factors that are unique in this shipping sector and might have an important role in the model as exogenous factors.

The concluding chapter includes information regarding the results and the impact this study can have in the shipping market for the decision makers involved in the industry.

KEYWORDS: dry bulk, time series, BDI, VECM, long-run, demand, supply, model,

1. Introduction

The world is constantly changing which can be reflected in the macro-economic factors that many economists and various industries have studied. The Baltic Dry Index is a variable that can be characterized as an economic indicator and it is the most important index in the dry bulk market for shipping. Our main goal here is to seek if there exists a long-term relationship between those factors, and if this is proven, then many people and the players from the shipping as well as other industries can follow the results and identify future trends in order to capitalize on them. Many researchers have studied this topic, using the BDI as one factor against various others, either macro-economic or supply driven.

The BDI is seen by some as the temperature of global economy and trade, therefore high values can indicate growth. Looking at the past, BDI crashed in 2008 during the world financial crisis and it was one of the first indicators that showed this trend and what was about to happen, as stated by Dalibor Gogic, principal analyst at HIS Maritime and Trade. However, the reasons this happened had various causes which include a slower growth rate of China than anticipated and even a large order-book increasing supply of shipping services and creating overcapacity in the market. As for 2014 (the last year on our thesis) and currently in 2016, BDI is at an all-time low where there are very low profit margins for ship-owners. Shipping is not an attractive industry and therefore vessels at low prices are creating opportunities for those who either want to enter the industry or expand their fleet. As in every economic cycle and especially "under perfect competition rules" that are applied in bulk shipping, those who will catch up on the boom of the market will first get the most profit out of it. So, despite BDI being at its lowest, the general economic theory and research that has been done in the industry, has shown that a big decline will in the long-term lead to a reduction in demand-supply gap and the market will start to flourish again. In order for someone to benefit by this cycle, they will need to take the appropriate actions, such as ordering new ships for the ship-owners way before the change actually happens in the market. When things don't go in their favor, then exiting the market as fast as possible, by either scrapping their vessels or selling them, could be the most profitable solution.

The shipping industry is a capital-intensive industry where the stakeholders (ship-owners and charterers mostly) need to be able to take long term decisions and predict the dynamics of the market one or two years in the future. This is due to the fact that a "boom" in the industry is followed by orders of new buildings (so that owners can profit by the extra demand). These orders will not come into circulation before at least 2 years, so if the current situation is about to change, ship-owners face the risk of "oversupply", meaning that they have more ships than needed by the current demand. Especially in the bulk carrier market, which is one of the closest markets to what economists refer to as "perfect competition", potential profits might be immense in periods of economic growth, but in periods of market decline bad decisions may lead to unsustainable conditions for a lot of businesses. So a model which would predict long term relationships and dynamics using easy to find data variables for anyone, like the ones we include in our model, could be vital for a ship-owner in order to avoid potential pitfalls in the future. Our thesis is therefore important for forecasting and identifying market opportunities. If a relationship exists, then someone can make more sophisticated and accurate estimations for future market values and dynamics and eventually make better decisions. Parties like, ship-owners, charterers and the shipping industry in general would appreciate the knowledge and will definitely show a great interest in a research such as this.

Based on past literature and research, there has not been a similar analysis examining the specific factors that we are going to analyze further in our thesis. There have been papers that examine the relationship of the BDI with various other factors, some are examined by us as well, but the majority of research did not involve the specific factors we will present here. The aim objective is the examine the relationship, if it exists, between the BDI and macro-economic factors, as well as some basic supply factors from the shipping industry and determine their impact, eventually making forecasting possible. An initial hypothesis would be that the BDI, as an indicator of global economy, should have a positive and strong correlation with many macro-economic factors. Higher demand for products and economic growth would lead to a higher demand for transportation and the shipping services. By contrast, an increase in the supply factors could drag the BDI down as is the common economic theory.

Regarding the order of information in the thesis, we will firstly explain some basic concepts regarding the shipping industry and then we will comment on the data we have and the hypothetical impact that we expect in our model. Then employing a co-integration model following the Johansen approach we will try and represent this relationship, if it exist, and then make some comments regarding the results. Finally, we will attempt to forecast future values based on the model and evaluate the accuracy of predictions by comparing with actual values.

2. Literature Review

Throughout the years there have been many attempts to examine and investigate various models regarding the shipping market. One of the very first attempts was made by Tinbergen (1959) who examined the effect of supply and demand factors on freight rates. He used the volume of seaborne trade as a measure of demand and the size of the world fleet and the price of bunkers to measure supply. It was observed that an increase in the fleet size would cause the freight rates to drop while the increase on bunker prices would lead to a decrease of supply because of the slow steaming of vessels, leading to less available tonnage in the market. Although, limited in data and less sophisticated in the estimation of freight rates this is the basis in our thesis which we will try to improve upon.

Like other industries, the dry bulk market is characterized by the supply and demand for shipping services according to Strandenes 1984, Beenstock and Vergottis 1993 and Tsolakis 2005. Shipping is a derived demand and depends on the commodity markets, world economic activity and trade volume (Stopford 2009). Supply on the other hand, depends on the fleet size and the available tonnage, new-building orders, bunker prices and the scrapping activity and prices that can reduce the supply of fleet. Furthermore, studies by Hawdon (1978) and Strandenes(1984) have shown that freight rates are determined by the growth in industrial production and consequently the seaborne trade, commodity trade volume, oil prices, the available ship tonnage and factors like ship-building and scrapping activity.

More recently, studies examined the interaction of freight rates with past values of in multivariate time series models as well as their volatility. Kavussanos and Alizadeh (2001) investigated seasonality in the dry bulk market and the impact of factors like the trade volume and industrial production.

Following co-integration methods and testing, Kavussanos, Visvikis and Menachof (2004) examined the unbiasedness hypothesis in the freight forward markets. It was found that FFA prices for up to one or two months to maturity are indeed reliable predictors on the realized spot rates for various routes. This subject deviates from the purpose of this thesis but the techniques that were employed, Co-integration modeling following the Johansen approach, will also be used here.

Tsolakis (2005) has performed an extensive econometric analysis of bulk shipping market and its implications for investment strategies and financial decision-making where he analyzed the four shipping markets and the effects on different types of ships. Additionally, the term structure relationship of time charter and spot charter rates was investigated and found that the time charter market has greater forecasting ability than the spot one. Econometric analysis and co-integration techniques of new-building and second-hand vessel prices were also examined with respect to ship-building costs and freight rates but also with respect to each other (new-building and second-hand). Finally, demolition prices were found to be primarily driven by market conditions and expectations. He concludes with some financial results regarding investments, strategies and portfolio theory. This is an analysis on various combinations regarding shipping variables and many important conclusions can be extracted. However, the analysis does not integrate macro-economic factors outside the shipping industry and this is what we will try to add by focusing exclusively on the dry bulk sector.

An analysis, close to the one we are about to present, was done by Petter Hvalgard Bakke and Daniel Reinsborg on 2012 for a Master Thesis regarding the empirical impacts of macroeconomic risk factors on tanker shipping equities. Using a portfolio of stocks of 18 listed tanker shipping companies as the dependent variable and different macroeconomic factors as the independent ones, they examined their effect on the portfolio returns and it was found that there was a positive and significant relationship with the WTI oil price, VLCC spot rates, tanker fleet size, the MSCI world index and the exchange rate of USD. Building on this approach and using the BDI as our dependent variable instead we will have a similar analysis.

Finally, a more recent paper by Melike E. Bildirici, Fazıl Kayıkçı and Işıl Şahin Onat regarding the BDI as a major economic policy indicator, examine the relationship between the BDI and economic growth in the United States. It was found that BDI and GDP are co-integrated for the USA by employing a VAR model. This recent analysis is especially important for our research because it was one of the few that included the effects of the recent crisis regime and helped us observe the interaction of the BDI with other macro-economic indicators during an unpredictable period of constant turmoil.

Big part of the research focuses on using the BDI as a predictor of other variables and don't focus on the opposite thought: What drives the BDI itself?

a) Trade, income and the Baltic Dry Index from the European Economic Review(2013), uses the BDI in order to answer questions regarding income of low level countries.

b) New Evidence on the Information and Predictive Content of the Baltic Dry Index (International Journal of Financial Studies,(2013) by Apergis and Payne, uses the BDI as an independent variable to test its predictive power regarding the state of the economy.

c) Recent papers such as "Can the Baltic Dry Index predict foreign exchange rates?" by Han, Liyan, Wan, Li, Xu, Yang (university of Beihang, Beijing , China) research the predictive power that BDI can have on market variables such as fx rates conclude that "… BDI provides statistically significant long-run predictability of currency returns."

The majority of the papers that do try to research the long term factors of the BDI don't use multivariate analysis to capture long term dynamics of the index, but rather alternative ones, that expand from traditional autoregressive techniques to newest Deep Neural Network techniques. Such examples are:

a) "Analysis of Baltic Dry Index based on GARCH Model" (2020) from the International Journal of Social Science and Education Research by Yuliang Zhang and Gang Zhao, focus on short term effects of economic variables.

b) "Time–frequency analysis of the Baltic Dry Index", (2017) focuses on capturing frequency/periodicity variations of the BDI and attempts to supplement the discussion of the BDI cyclical behaviour by highlighting its evolving structure through time.

c) "DERN: Deep Ensemble Learning Model for Short- and Long-Term Prediction of Baltic Dry Index", explores Deep Neural Network techniques to model short and long term relationships of the BDI but they don't take into account economic factors and focus solely on the predictive power of DNNs with other variables but do not explore a wide range of endogenous variables from different areas (macroeconomic, supply, demand).

Finally, part of the research is only targeting a subset of geographical locations (e.g. only China Market, BRICS countries etc.). This is demonstrated in research papers like the following:

a) "The Dynamic Relationships Between The Baltic Dry Index And The BRICS Stock Markets: A Wavelet Analysis" focuses on a specific group of countries only.These five countries, known collectively as the "BRICS" (Brazil, Russia, India, China and South Africa), form an important economic block that is the major scope of this paper, but a more global and more scalable approach is not considered.

b) “The Effect of Baltic Dry Index, Gold, Oil and USA Trade Balance on Dow Jones Sustainability Index World”, by Grigoris Giannarakis1, Christos Lemonakis2 , Asterios Sormas , Christos Georganakis (2017), from the International Journal of Economics and Financial Issues. This study aims to investigate the effect of economic leading indicator of Baltic dry index (BDI) on stock returns of socially responsible stock index.

In conclusion, many papers have (and still do) investigated the BDI from different perspectives but most of them ignore a holistic approach that considers variables from the whole spectrum of economic theory (macroeconomic, supply-demand, economic policy, maritime specific). Our paper will be based on the assumptions and results from others and produce a more updated and specific research regarding the dry bulk sector with data the reach up to 2014q4.

3. Market and Data Description

3.1 Baltic Exchange and the BDI

“Our word our bond” – The Baltic Exchange

The Baltic Exchange is the world’s only independent source of maritime market information for trading and settlement of physical and derivative shipping contracts (www.balticexchange.com). The London-based Baltic Exchange uses information gathered by the maritime community, more than 550 members such as the shipbrokers, owners, charterers and other traders in order to produce frequent assessments of the market but most importantly indices. These indices cover freight rates regarding time and voyage charter for various routes and many different types of vessels, from handy-size to capsize bulk carriers and product tankers to ULCCs (Ultra Large Crude Carrier). Founded in 1744, originally a coffee house, now publishes seven daily indices for both the dry and wet voyages and these include the Baltic Dry Index (BDI), Baltic Panamax Index (BPI), Baltic Capesize Index (BCI), Baltic Supramax Index (BSI), Baltic Handysize Index (BHSI), Baltic Dirty Tanker Index (BDTI) and Baltic Clean Tanker Index (BCTI). Additionally, it provides forwards curves, dry cargo fixture list and other news or settlement data.

The Baltic Dry Index is the one that we are focusing on in our paper as our dependent variable in our analysis that follows. It is a general index regarding the dry bulk sector, measured in time-charter hire rates (hire per day) and includes all other four dry cargo routes (BPI, BCI, BSI, and BHSI). It essentially measures the cost to transport different raw materials such as coal, iron ore or grain and it is considered as a leading indicator of future economic growth. Furthermore, the index measures the demand for shipping services (cargo) against the supply of carriers (deadweight capacity).

3.2 Bulk Carriers

Since we will be focusing on the BDI, it is important to have an understanding of the types of vessels that carry the cargo. A bulk carrier is a large single deck ship which carriers unpackaged cargo which is poured, tipped or pumped into the holds of the ship. Most cargoes are raw materials such as iron ore, grain, coal, phosphates, bauxite or other dry cargo classified as general cargo. The Baltic Exchange uses “model” ships when referring to a certain index. So starting with Capesize ships which are the largest of the ships that take part in the BDI; they can usually carry more than 100,000 deadweight tons of cargo and are mainly carrying iron ore or coal on long routes between well-equipped terminals. Panamax ships have a deadweight capacity of 60,000 to 80,000 tons, gearless and mainly transfer coal, iron ore or grains. A Supramax can carry up to 60,000 tons and normally have on-board equipment for loading and unloading and can visit most ports. Like the previous category, Handysize ships are geared and can carry up to 35,000 deadweight tons of various cargoes in bulk.

3.3 The four shipping markets

In order to understand the behavior of freight rates and indices like the BDI we need to examine the four markets that control the shipping industry. These are the freight market, the purchasing market which can be divided into new-building and second hand, and finally the demolition (recycling) market. Understanding these markets will give us an insight into the shipping industry as a whole, help us identify the causes of the shipping cycle and give us the ability for forecasts. It is also important to note the actors that take part into the market which are the ship-owners, the shipyards, the charterers and finally the scrap yards.

The freight market is the main source of income for shipping companies and it is very closely related to the main operation of a company. Shipping companies offer to transport cargo or commodities in exchange for a cost which is called the freight. This agreement between the shipper and the ship-owner can be of two different types. We have the voyage charter where the ship is chartered for a specific voyages/voyages under a dollar per ton basis for a fixed amount of cargo. Under a time charter agreement, the ship is hired by the day on a dollar per day basis for a specified amount of time, while the charterer has the ability to manage the ship according to his needs. These are the two basic types of chartering; however there are other types of agreements between the two parties that are more complicated, especially in the liner market, with different cost distribution that we will not get into in our thesis.

In the new-building market, ship-owners acquire new ships that must be built, a process that may take up to 3 years. Purchasers may have different reasons for entering this market instead of buying a second-hand vessel; however prices are still affected by supply and demand just like the second-hand market. The actors are the ship-owners and the ship-yards which produce the ships according to the specifications of each order. New-buildings are important for the market since they lead to an increase in the supply, a major factor for the profitability and maritime economy. As a result, prices are closely correlated with other factors like the second-hand prices or the freight rates.

The third category is the second-hand or the sale and purchase market. Ship-owners buy/sell vessels among each other and vessels simply change hands. Participants in this market may decide to buy/sell a vessel because of its age, their change in trading patterns, speculation or because of financial issues. There is a great deal of players that enter this market with no intention to operate the vessel and are purely interested in asset values called asset players. Same as the previous market, prices are volatile and depend on freight rates, age, inflation and expectations (Stopford 2009).

Finally, once a vessel reaches a certain age it will be demolished. This process has created the demolition market, where the steel components of a ship are sold to the scrap yards. The actors in this case are the ship-owners and the demolition yards “buying” the parts of the ship according to the prevailing market prices for steel or other metals. Prices, or dollars per lightweight tonnage of the ship, are also volatile and correlated with the other markets and that can create difficult decisions (whether to keep operating or scrap) for the ship-owners especially in a distressed market like today.

3.4 Data Collection and Sources

For our data collection process we used the Bloomberg database for most of the macro-economic factors and some shipping indices. However, data regarding the supply side of our analysis was gathered from the Clarkson’s Shipping Intelligence Network, a dedicated database for shipping information. It is also important to note, that the process of variable and data selection has been greatly affected by information and news from shipping news sites and newspapers like TradeWinds. Our sample consists of data beginning in Quarter 4 of 1995 until Quarter 2 of 2014. We based our analysis on a quarterly frequency because of the nature of our data. Most macro economic factors are not published daily, while some of them are only available annually. So a quarterly frequency was going to produce the biggest sample possible given the restrictions.

3.5 Introduction for Variables

Our data set consists of quarterly macro-economic factors like the World trade, the MSCI index or the Exchange rate between EUR and USD. Our dependent variable will be the BDI, so our analysis focuses on the existence of this relation between the BDI and the rest of our time series and the impact that each of the variable has if a relationship is proven. By contrast to other studies or papers on this subject, we focus on macro-economic and financial data plus some factors from the supply side of the equation regarding shipping. Obviously, much research has been done involving other data like commodity prices or indices from other business sectors outside shipping and all of these are important factors which we purposely neglect for the objective of our thesis.

3.5.1 BDI – Dependent variable

BDI as mentioned before, the Baltic Dry Index is a general index regarding the dry bulk sector, measured in time-charter hire rates (hire per day) and includes all other four dry cargo routes. It is an index that measures the cost of transporting various materials and reflects the revenue of shipping companies for providing the transportation service. In January 1985, the Baltic Exchange produced the Baltic Freight Index at a daily frequency which later evolved into today's BDI. The importance of the index lies in its ability to measure the demand for shipping services and it is considered one of the most useful economic indicators of economic activity. Over 90% of world trade is carried by the shipping industry (International Chamber of Shipping) which is by far the most cost-effective way of transporting large quantities of goods around the world.

BDI, as an indicator of shipping revenue, is related to the supply and demand equilibrium. The supply of capacity/ships is inelastic, meaning that because of the capital-intensive nature of the business, it can take up to 2 or 3 years for a new ship to enter the market. Therefore, an increase in demand for shipping services can push the index higher much quicker than some other industries and vice versa. Supply cannot adjust at the same pace and demand is not satisfied resulting to a rapid increase in freight rates and the BDI as an index. When supply exceeds demand, we have many ships chasing few cargoes (overcapacity) and that can crash freight rates. However, when we have high demand for shipping services many cargoes chase fewer ships resulting into high freight rates. Therefore, we expect a strong negative relationship of the BDI with the Supply of vessels (Fleet size).

Variable	Obs	Mean	Std. Dev.	Min	Max
BDI	75	2425.347	2020.164	699	9589

Figure 1: BDI Summary statistics

Our sample consists of 75 observations over our examined period. What is most important is that during the shipping boom and the peak of the cycle of 2008, BDI reached a value of 9589 index points, which was about 13 times more than the minimum value, until overcapacity hit and the index fell rapidly into a new all time low. From there, it managed to recover for about 8 months until it started to collapse again following a downwards trends until the end of our sample period.

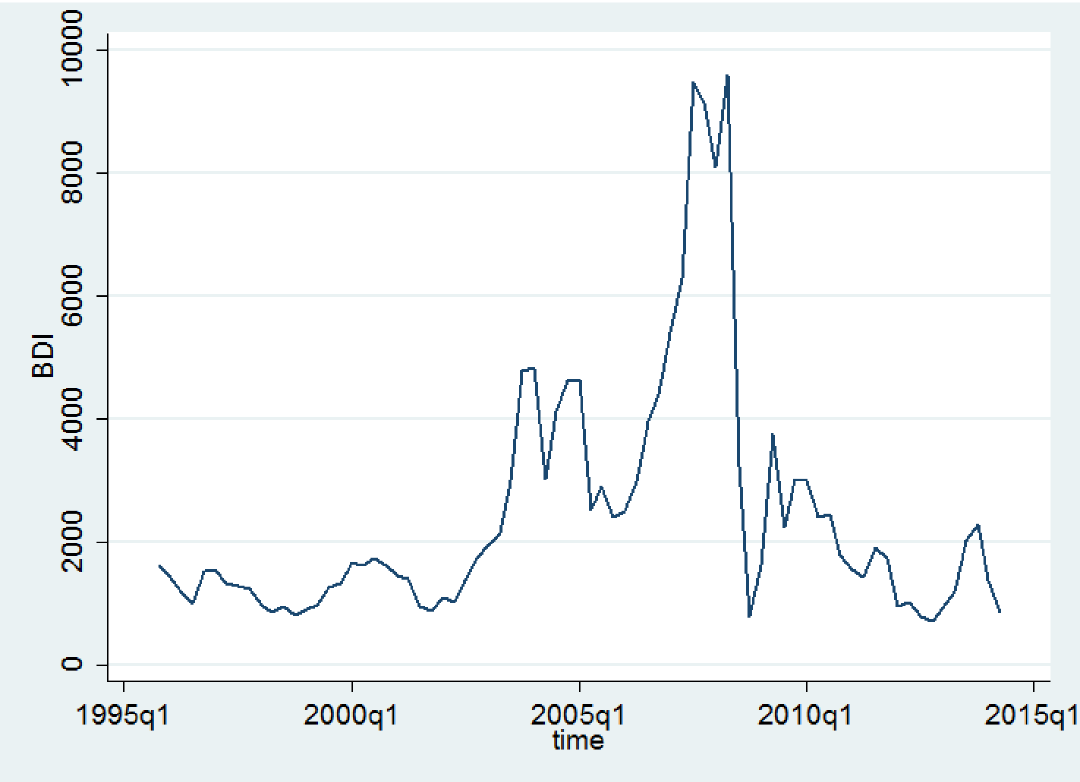


Figure 2: BDI

3.5.2 OECD World Trade Volume

OECD World trade volume, refers to the Volume of goods traded worldwide measured in billion USD and it is produced by the OECD organization .We expect that an increase in the volume of trading goods around the globe will have a positive impact in the demand for shipping services and thus, increase the BDI as well.

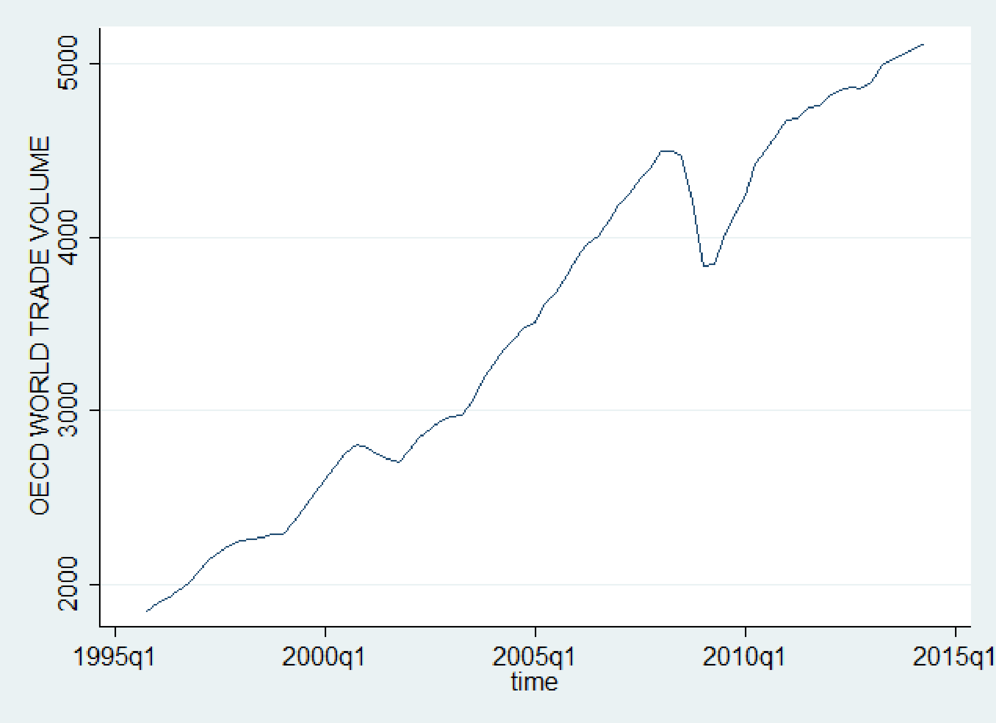


Figure 3: OECD World Trade Volume Index

We expect that the OECD World trade volume Index will be exhibit a positive trend (Apart from the end of the 00's decade post the Lehman's crisis where it started declining for a few months. This is expected since one, if not the biggest, factor driving the world trade volume is the world's population which is increasing continuously post the World War II, even during periods of economic fallouts. Also, China's high growth rate of about 12% in 2010 (National Bureau of statistics of China) played in big role in the increase of the trade volume. This might be an indication later on for its significance in our modeling of the BDI.

3.5.3 VIX

VIX is a popular measure of the implied volatility of S&P 500 index options; the VIX is calculated by the Chicago Board Options Exchange (CBOE). The VIX represents one measure of the market's expectation of stock market volatility over the next 30-day period. In other words, it translates, roughly, to the expected movement in the S&P 500 index over the upcoming 30-day period. It is used as the underlying asset for derivatives. A volatility index would play the same role as the market index for options and futures on the index.

For example, if the VIX is 15, this represents an expected annualized change, with a 68% probability, of less than 15% up or down. One can calculate the expected volatility range for a single month from this figure by dividing the VIX figure of 15 not by 12, but by √12 which would imply a range of +/- 4.33% over the next 30-day period. Similarly, expected volatility for a week would be 15 divided by √52, or +/- 2.08%. or in case of our study the quarterly volatility could be aggregated from 15 divided by $\sqrt{4}$, or +/- 7.5%.

For our modeling, we would like to have at least one global volatility indicator and VIX is the most common used in macroeconomic research. Note that the VIX is a measure of market perceived volatility in either direction, both to the upside and to the downside. So, an increased value of the VIX might indicate better or worse conditions for the global economy. (This will be one of the reasons that we are going to exclude it off our final model among other factors that will be discussed later).

How does VIX affect the BDI? Hypothetically, we expect that during periods of high volatility and uncertainty, according to an interview of Dr. Nicholas Bloom, professor of economics at Stanford University, "…the volatility can lead to uncertainty, and when individuals and businesses are uncertain about the future, they put off big purchases". This will lead to a drop in investment and consequently to a decline in the global markets thus affecting the global demand for shipping services and demand of transportation of goods. A scenario would be that a high value of the VIX could spur a decline to the BDI.

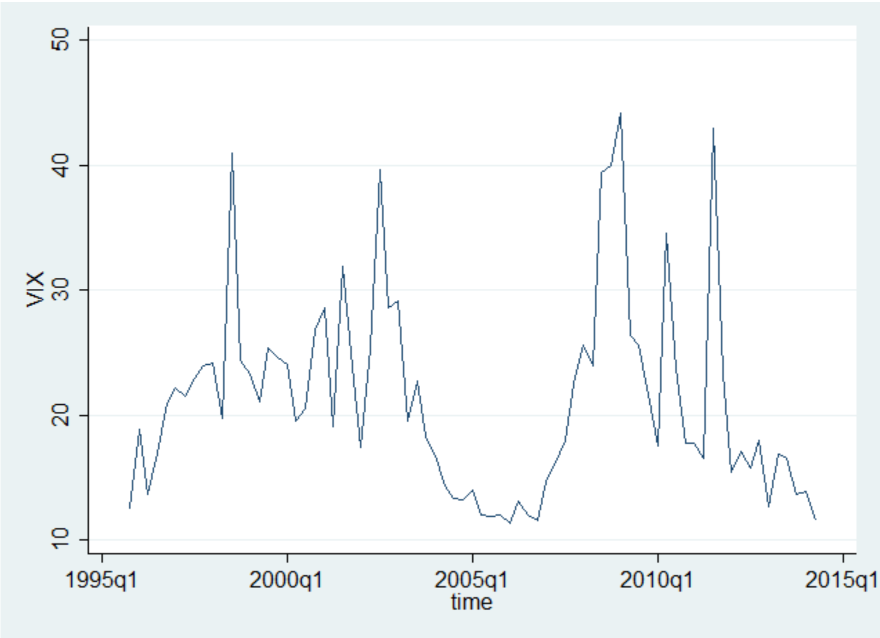


Figure 4: VIX in the last 20 years

From the graph, it is apparent that we are talking about a non-stationary variable. One thing to consider are the frequent spikes upwards that VIX shows periodically. These are signs of spurious nature of the index. When the volatility is high, it will probably not stay at these levels for more than a month, thus the expected drop of the BDI value might not be actually realized in this timeframe. Generally, if the current volatility doesn't last, its effects shouldn't linger either.

3.5.4 US and EU Economic Policy Uncertainty

Policy uncertainty (also called regime uncertainty) is a class of economic risk where the future government policy is uncertain, raising risk premia and leading businesses and individuals to delay spending and investment. Policy uncertainty may refer to uncertainty about monetary or fiscal policy, the tax or regulatory regime, or uncertainty over electoral outcomes that will influence political leadership.

These two indices are used to follow such policy uncertainty in the US and in Europe. Greater policy uncertainty (as in the case of economic uncertainty generated by high volatility in the stock and derivatives markets that we analyzed in the case of the VIX) would lead us to hypothesize that consumers from the USA and the European Union will likely reduce consumption. Since these 2 comprise the biggest percentage of worldwide consumers, we will experience a decline in global demand for products and services thus, dragging the BDI down.

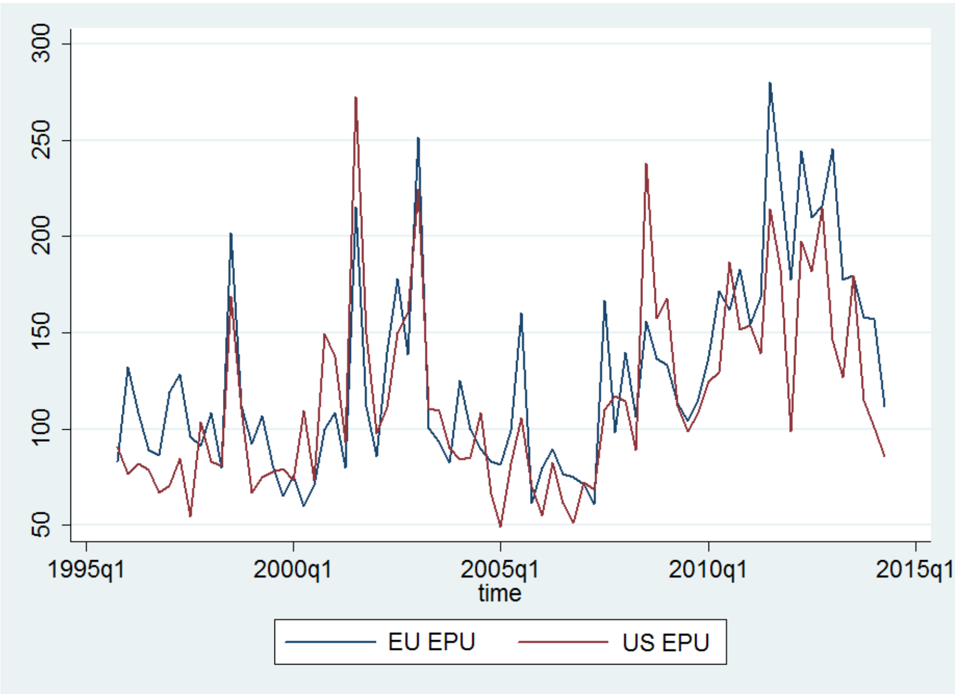


Figure 5: European Policy Uncertainty index

The graph exhibits a general positive correlation between the two indexes, which is logical, because changes in the US policy will most probably affect the EU Policy as well and vice versa. Generally, the European Policy Uncertainty is most of the times higher than that of the US.

3.5.5 MSCI

From Bloomberg: The MSCI World Index is a free float weighted equity index. It was developed with a base value of 100 as of December 31, 1969. The MSCI is a stock market index which includes over 6,000 different stocks from all around the world. It is maintained by MSCI Inc., formerly Morgan Stanley Capital International, and is used as a common benchmark for 'world' or 'global' stock funds. The index includes a collection of stocks of all the developed markets in the world, as defined by MSCI. It is a broad global equity benchmark that represents large and mid-cap equity performance across 23 developed markets countries.

Figure 6: European Policy Uncertainty index

The MSCI World Index has been calculated since 1969, in various forms: without dividends (Price Index), with net or with gross dividends reinvested (Net and Gross Index), in US dollars, Euro and local currencies.

Both Kavussanos and Alizadeh (2002) and Stopford (2009) explained that shipping is clearly influenced by the world economy because of its international presence. Shipping is essentially a facilitator of world trade. Changes in the world economy, measured as the return in global equity markets, is therefore believed to have positive impacts on tanker stock returns. From general macroeconomics, we know that: As the world economy is growing more money will come into circulation leading to increased demand on a world basis. Increased demand influences industrial activity which further leads to an increase in energy consumptions. The increase in energy consumption leads to increased demand for oil which makes the freight rate mechanism adjust for the imbalance between supply and demand by increasing freight rates. Freight rates are the main source of income for the tanker shipping companies and they therefore have a direct impact on the pricing of the companies. Earlier findings by Grammenos & Arkoulis (2002) and Mitter & Jensen (2006) have confirmed this relationship. MSCI is a trusted source for emerging markets equity performance and, consequently a trustworthy macroeconomic index benchmark for the global economy.

```
. summarize MSCI
```

Variable	Obs	Mean	Std. Dev.	Min	Max
MSCI	75	1168.081	253.8123	734.28	1743.42

Figure 7: Descriptive Statistics of MSCI variable

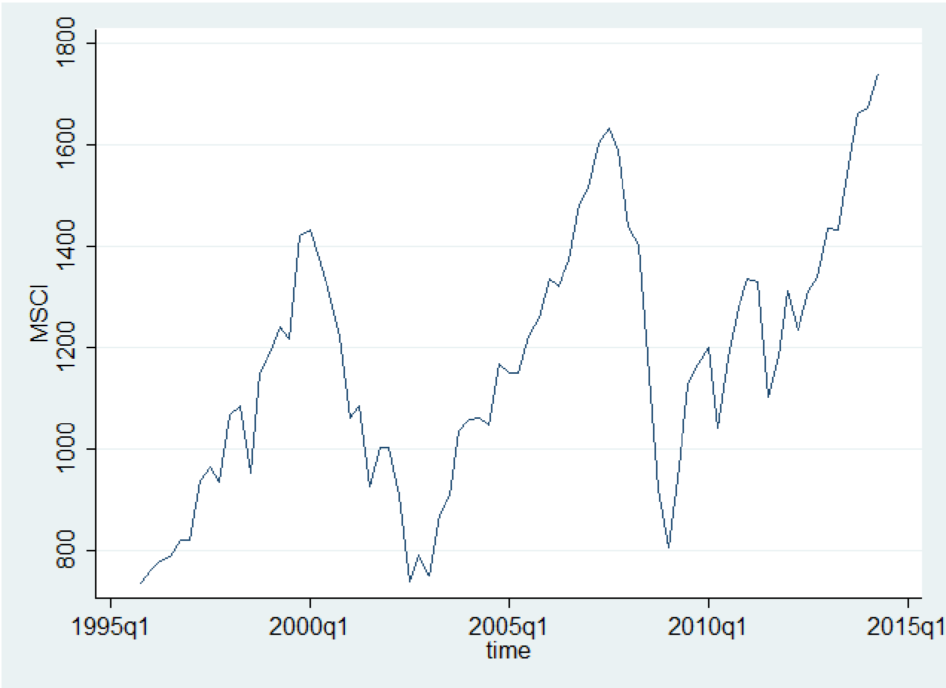


Figure 8: MSCI

We can observe a periodicity in the index from 1995 until the end of 2010. From there and until the present day we observe a positive trend (MSCI has reached each maximum value of 1743.42 in the last 20 years at the end of 2014).

3.5.6 EUR/USD Exchange rate

The EURUSD variable represents the changes in the exchange rate between US dollars and Euro. The relationship between the US dollar rate and other exchange rates is an important factor for shipping companies (again, due to shipping's inherently international nature). The majority of the shipping trade and services around the world is quoted at USD prices. Due to the fact that the euro and U.S. dollar are the world's two largest currencies, representing the world's two largest economic and trading blocs the Euro/USD exchange rate can be considered as the most important exchange rate for the global economy.

The reason for the importance of the US dollar regarding its potential relationship with the BDI is simple. Most shipping companies have their capital costs, voyage costs and revenues in dollars and are often based in non-dollar countries with returns and performance reported in local currencies.

The effects of changes in the value of the dollar can be divided into direct effects and indirect effects. The direct effect of the dollar increasing in value relative to other currencies is an immediate increase in the freight rates. The indirect effect will come as a sequence of macroeconomic connections. An increasing value of the dollar relative to other currencies will lower the demand for goods quoted in dollars, such as oil. This will cause a decrease in the international trade and thus the demand for transportation of oil decreases. Our hypothesis is therefore that dollar appreciation will have both a direct positive impact on the BDI because of the exchange rate effect on freight rates and an indirect negative impact where the appreciation leads to a lower demand for shipping services.

Variable	Obs	Mean	Std. Dev.	Min	Max
EURUSDECHA~E	75	1.222573	.1746641	.849	1.5809

Figure 9: Descriptive statistics for EUR/USD rate

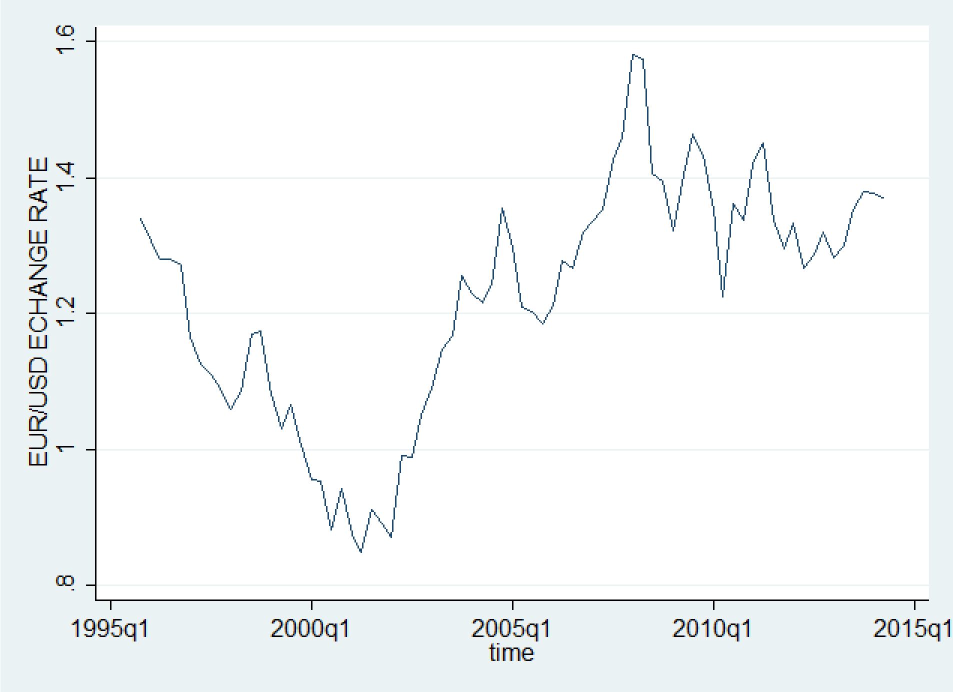


Figure 10: EUR/USD exchange rate in the last 20 years

From the graph, we would expect that our empirical analysis will show that EUR/USD exchange rate would be a non-stationary variable. One thing to note is that despite the crisis of 2008 the exchange rate dropped but nowhere near its minimum value (0.849 at early 2000's). This shows us that big changes of this rate could be attributed changes in global macroeconomic policies (e.g. political changes for the US Federal Reserve or the European Central Bank).

3.5.7 US Federals Funds Rate

In the United States, the federal funds rate is the interest rate at which depository institutions (banks and credit unions) lend reserve balances to other depository institutions overnight, on an uncollateralized basis. The federal funds rate is an important benchmark in financial markets and as, such, we would like to consider it to be a part of our model. The Federal Reserve influences the supply of money in the U.S. economy to make the federal funds effective rate follow the federal funds target rate. Since it has such a big impact in the US economy and for the same reasoning we used before (basic monetary unit of shipping trading and services is the USD), we could argue that it might affect the BDI in a significant way. A hypothesis would be that if the Federal Funds rate is raised, the supply of money in the US economy will decline (borrowing will become more expensive). This will eventually lead to a decline in demand for goods and services quoted in dollars thus affecting the global demand for shipping services and the BDI negatively. Conversely, dropping the interest rates will encourage banks to borrow money and therefore invest more freely. So, this interest rate is used as a regulatory tool to control how freely the U.S. economy operates.

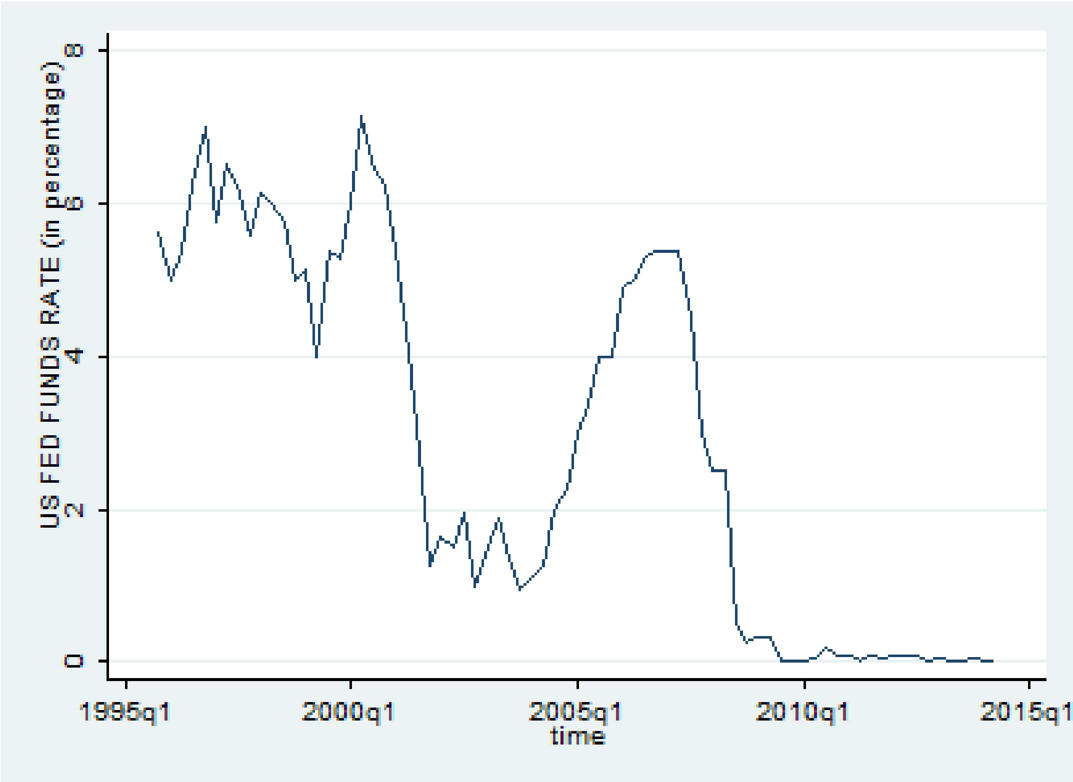


Figure 11: The US Fed Funds Rate through the years

In the last 5 years, the rate has remained almost unchanged in historic low levels. We could argue, for that matter, that for 25% of our sample period this rate would not be useful to for the modelling of BDI. This will be more prevalent later in our analysis and will come up during the trimming of the redundant variables off of our model.

3.5.8 Brent Oil Price

Brent oil is a grade of crude oil and is used as a benchmark in oil pricing. It is the most widely used marker of all types of oil and covers about two-thirds of the oil traded around the world. The decision to select Brent over WTI (Western Texas Intermediate) was made because of its popularity, while most contracts, like Forwards or other derivatives, are based on this type of oil. Brent is the more "international" type of crude oil, produced in the North Sea while the WTI is specifically extracted from the U.S and mainly consumed domestically so it makes more sense to use Brent as a measure. This is measured in dollars per barrel and it is a great indicator of the world prices and economy.

Voyage costs are very important in shipping; therefore bunker expenses closely determine the profitability of a voyage for any shipping company. The price of bunkers is correlated with the crude oil price so it is expected that when oil prices are high, more costs would occur and that would lead to a decrease in profitability. On the other hand, a high oil price is an indicator of a healthy world economy, meaning greater consumption in general and of oil products. So demand for shipping services would increase. However, it would be interesting to analyze the effect of this variable in combination with the rest.

Regarding some descriptive statistics of this time series, we can focus on the values over the examined period where we observe that from around 20$/bbl in 1996, Brent Oil reached a value of 120 $/bbl in 2007 a point a time where the shipping cycle was at its peak. However, due to the effect of a global economic crisis it quickly dropped to around 40 $/bbl in 2009 and managed to stabilize to a price of 110 $/bbl during the 2011 to 2014 period. The mean is about 55.21 $/bbl however, as described before, this is not a good representation the actual average of the time series and it is an issue we will address later in our analysis.

3.5.9 CPI

The Consumer Price Index is a statistical estimate that examines the weighted average prices of a basket of consumer goods and services. Using prices of a sample of items many sub-indices are created and combined to produce the overall index. CPI expresses current prices of goods in terms of prices of previous years to show the effect on purchasing power. This measure is especially useful in calculated the real value of goods or wages, salaries etc and it is one of the most important economic statistics closely observed by economists. CPI is also tied to inflation, since the annual percentage change in the CPI is actually the inflation.

Used as one of the most important macro-economic factors, CPI helps us understand the real economy. Observing the values during our examined period, we can see that CPI has a constant upwards trend, almost linear. Starting at about 50 in 1995 CPI reached a value of 120 in 2015, more than doubling its value. Obviously, the cost of living has increased and that is bad for consumers. On the other hand businesses or the public sector can gain from that because of the fall of real expenses for their operation. Same principals hold for the shipping industry as well.

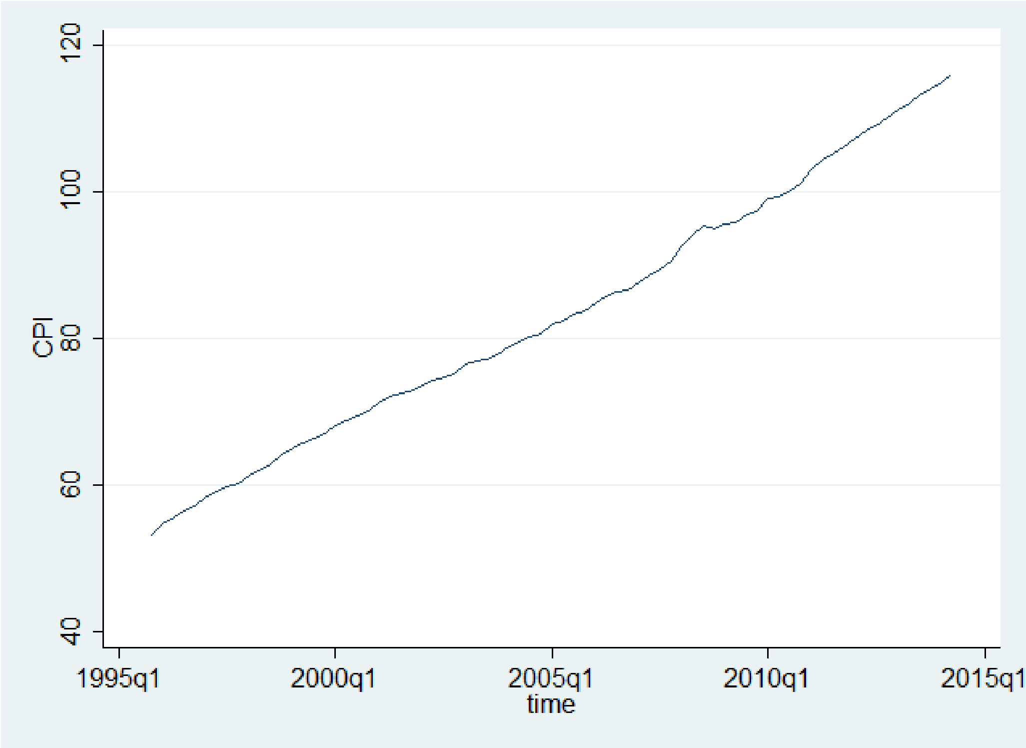


Figure 22: CPI index in the last 20 years

3.5.10 Bulk carrier and Tanker New-building Price Index

This factor is an indicator of the cost for a new-building bulk carrier or tanker vessel. This market, as we mentioned before, is different from the rest of the market because of the time difference between an order and the actual delivery of the vessel. Usually the amount of time required for the completion of a new-building order can be 2 to 3 years. This variable is important for the supply side of the equation is closely related to other factors like the freight rates. As one can assume, in a booming market, more and more ship-owners are willing to invest in an attractive sector with high returns. This would cause the price of new-building vessels to increase because of the high demand and profitability of shipping services. However, due to the time lag of the delivery, when the vessels enter the market, supply of shipping services increases which leads to a decrease in the profitability and freight rates.

Regarding these two indices, in specific, they are produced by the Clarkson's Shipping Intelligence Network. Using January 1988 as a benchmark for new-building prices for both bulk-carriers and tankers (January 1988=100) the indices present the development in prices throughout the years.

From the chart below, we observe two highly correlated variables that move in parallel throughout the examined period. Bulk-carrier prices seem to be, almost always, lower than the tanker ones, and this is mostly attributed to the nature of the tanker vessels, which managed to generate higher profits compared to the dry bulk sector. To be precise, correlation between the two time series is 97.47%.

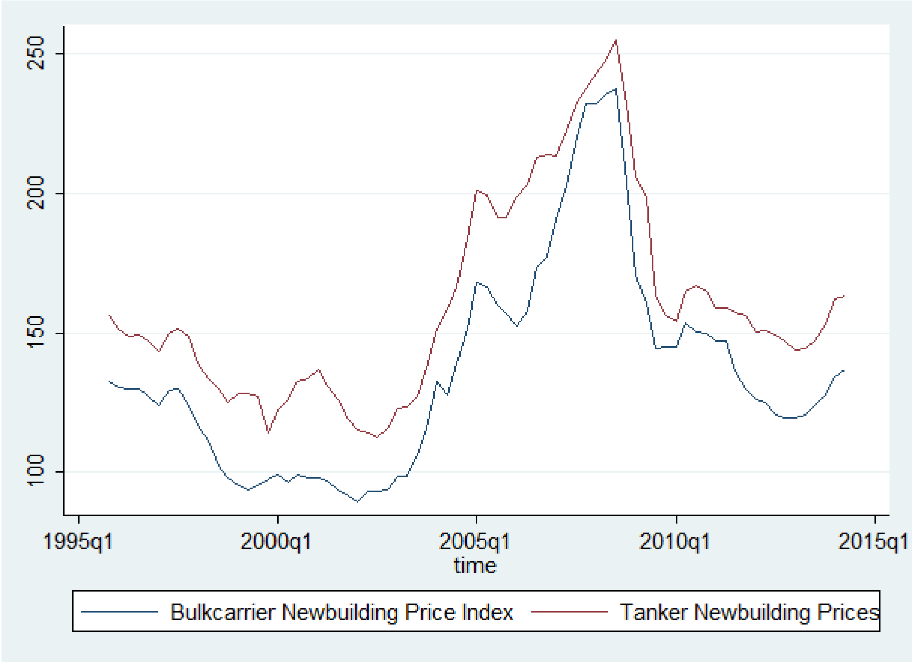


Figure 13: Bulk carrier and Tanker New-building Prices

3.5.11 Bulk carrier and Tanker Second-hand Price Index

Similarly to our previous variables, these two indices present the cost for purchasing a second-hand vessel, either bulk carrier or tanker. Second-hand transactions are very common in shipping, while the amount of vessels changing hands is usually greater than the new-building orders placed each year. The benefit of this type of market is the prompt delivery of the vessel without any delays. This has a huge impact at times where freight rates are high and ship-owners want to get into the business as fast as possible in order to gain from high returns and high freight rates.

The two indices were produced using January 1988 as a benchmark for second-hand prices and at that point the indices had a value of 100. However, it is important to note that during the big shipping boom in 2008 to 2010, Clarkson's Research did not publish benchmark values. It was a period of transition in the S & P markets, characterized by spells of rapidly changing prices levels, low levels of sales activity and a wide spread of price ideas. Therefore, data during the October 2008 to January 2010 was based on rough estimations and should be treated with caution. The indices were continued on January 2010.

Aside from the shipping boom in 2008 where second-hand prices for bulk carriers reached a value of about 495 (peak of our examined period), values for both indices were following the same path indicating strong correlation evidence.

Variable	Obs	Mean	Std. Dev.	Min	Max
BulkCarrie~d	75	166.9533	95.03806	75.49415	495.0918
TankerSeco~x	75	142.3535	43.86921	97.23196	246.959

Figure 14: Descriptive statistics of newbuilding prices for bulk carriers and tankers

3.5.12 Bulk Carrier Fleet Development

This factor covers the supply side of shipping regarding the dry bulk sector. Measured in million deadweight tons, it is essentially the amount or capacity that ships can carry. It is expected that as new-building orders (order book) increase, more new vessels will enter the market increasing supply. Supply is only decreased when vessels are sent for scrap/demolition. Second-hand transactions have no effect, since ships are only "changing hands" still operating in the same market under the same conditions. So it would make sense to examine the relationship of new-building prices and fleet size, perhaps in another paper.

Fleet development is the most important determinant factor of Supply of shipping services; therefore many interesting results can be extracted when we examine the relationship with other variables like the freight rates or the world trade volume. The fleet growth rate which reflects the difference between shipbuilding and scrapping moves pro-cyclically but it also tends to lag behind rates (Tsolakis, 2016)

From our data, which we gathered through the Clarkson's Shipping Intelligence Network, we observe that there is an upwards trend throughout the examined period and what is most interesting is that, although there was a collapse in freight rates in around 2009 to 2010, the bulk carrier fleet kept increasing. Fleet size reaches its maximum value of 739 million DWT at the end of our sample of Q2 2014.

Variable	Obs	Mean	Std. Dev.	Min	Max
TotalBulkc~e	75	385.7926	147.3522	240.4394	739.461

Figure 15: Bulk carrier fleet development

3.5.13 Tanker Fleet Development

Similarly to the previous variable, Tanker Fleet Development shows the amount of deadweight capacity (in millions) available by tanker vessels. We must note that, only tankers of 10,000 DWT or more were included in our sample. Smaller tankers are usually involved in short sea shipping; therefore bigger vessels are a better indicator and measure when we analyze macro-economic factors.

From the histogram below we observe, that from 1994 to 2010, fleet size for both sectors was almost identical. However, at Q1 of 2010 bulk carriers kept increasing at a faster rate than tankers. The shock was not as evident in the tanker sector as it was in the dry bulk and tankers maintained the same trend, steadily increasing the amount of deadweight capacity available in the market.

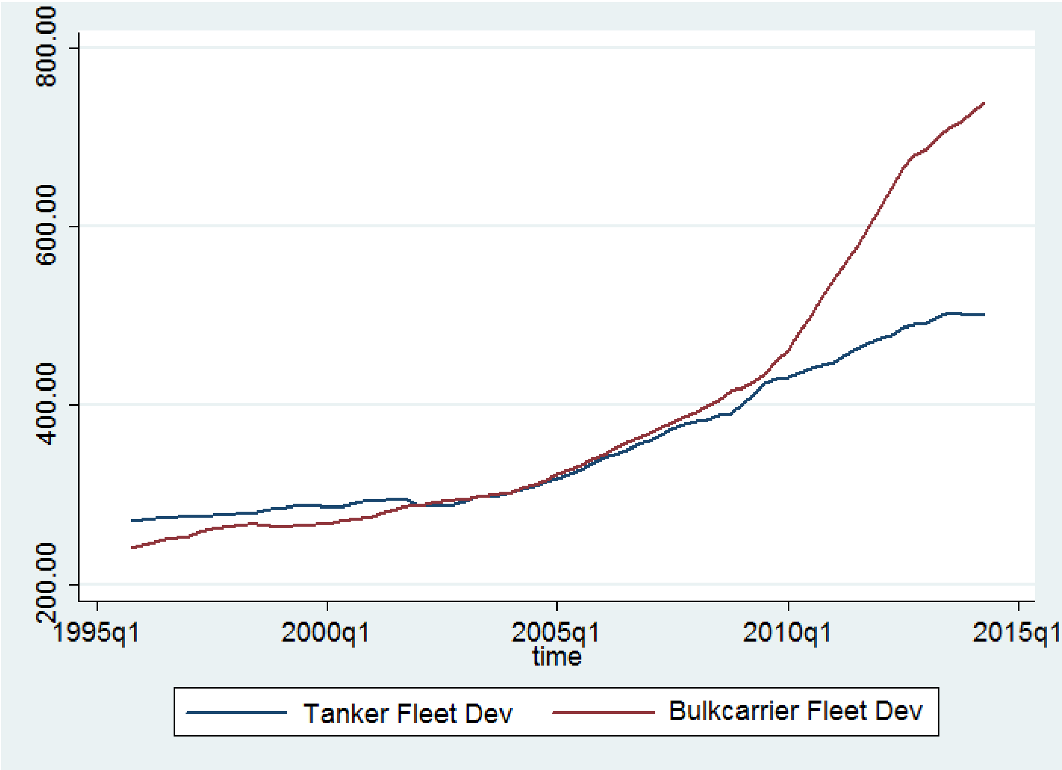


Figure 16: Comparison of Tanker and Bulk carrier fleets

3.5.14 Bulk carrier and Tanker Scrap Price

Continuous progress in technology and increasing costs of ageing ships over 30 or sometimes 20 years old have lead ship-owners to the decision to scrap their vessels. When, second-hand markets are not profitable for an old vessel, the only “buyer” is the demolition marker meaning the scrap yards. This is the final step in a ships life cycle.

This factor involves the profit that a ship-owner can get by scrapping the vessel. Assuming that the ship has reached a certain age and is no longer capable of operation due to technical issues or market conditions, the only profit can come from selling the steel that the ship is made off. Scrap prices are measured in dollars ($) per light-deadweight. Most of scrapping is being done in the Far East, the Indian Subcontinent or Bangladesh. We observe that, there is little difference in prices across these areas while correlation was more than 95%. This holds for both bulk and tanker sector, and therefore the decision to choose one of the areas would not make a difference in our analysis. We eventually, decide on two variables, one of which was the Bulk Scrap Prices based on Far East and the second one was Tanker Scrap Prices also based in the Far East.

Scrap prices are important in comparison with second-hand prices. As mentioned before, ship-owners face decisions between selling and scrapping a vessel when there is no profit in operating it. So the relationship between scrap and second-hand prices is an important one. For example, high scrap prices can generate immediate returns for ship-owners who struggle in the market making. At the same time, scrapping large amounts of deadweight due to the attractive prices in the demolition market will lead to a reduction in supply. The result is excessive demand with many cargoes chasing fewer ships, making freight rates sky rocket. Consequently, whenever a ship is scrapped, the supply side will decrease making the market more attractive for the rest.

As we can see in the following graph, scrap prices for bulk and tankers move in parallel and are almost identical for the most part. This is expected since vessel types make no difference for demolition yards which are basically only interested in the steel or other materials that can come from them.

Figure 17: Comparison of Bulkers' and Tankers' scrap prices

3.6 Hypothetical Impact

Trying to explain the relationship between the dependent variable and the factors mentioned before we need to examine the impact on the BDI, if there exists such an impact, and then its direction. Below is a table for each independent variable and the hypothetical impact we assume it will have on the BDI prior to our analysis.

Variable	**Description**	**Impact**
OECD World Trade Volume	Volume of trading good around the globe	+
VIX	Implied volatility index of S&P 500	-
US and EU Economic Policy Uncertainty	Uncertainty Policy indices	-
MSCI World index	Global Stock Market Index	+
EUR/USD Exchange rate	Exchange rate	+/-
US Federals Funds Rate	Interest rate	+/-
Brent Oil Price	Crude oil price	+/-
CPI	Consumer Price Index	+/-
Bulk carrier and Tanker New-building Price Index	Cost for new-building vessels	+
Bulk carrier and Tanker Second-hand Price Index	Cost for second-hand vessels	+
Bulk Carrier Fleet Development	Fleet size of bulk carriers	-
Tanker Fleet Development	Fleet size of tankers	-
Bulk carrier and Tanker Scrap Price	Scrap prices	+/-

Table 1: Hypothetical impact of the independent variables on the BDI

4. Methodology

4.1 Preparation

In order to achieve the thesis' objectives, it was decided to use quantitative methods and specifically econometric analysis of time series data.

In this chapter, we will focus on primary research which involves the data collection (a process described previously) and then data analysis based on established methodology. Therefore our analysis will be based on data covering the period of Q4 1995 to Q4 2014, meaning we are dealing with time series data. This is a collection of continuous random variables at a quarterly frequency. Based on the method derived in Basic Econometrics (Damodar N. Gujarati, Dawn C. Porter, 2009) a time series variable is denoted as Y_t where Y is the name of the variable and t the point in time. In our case t denotes the specific quarter and year of our time series.

In order to be able to process time series data, distinction between stationary and non-stationary variables is important. A stationary time series is one whose statistical properties such as mean, variance, autocorrelation, etc. are all constant over time and is said to be integrated of order 0 or I(0) (Damodar N. Gujarati, Dawn C. Porter, 2009). Therefore when a time series is stationary we can obtain meaningful statistics such as the mean or correlations with other variables and these are useful descriptions of future behavior. However, many of the real world data are not like that and said to be non-stationary and exhibit trends, cycles or random walks. Based on what we said before, a non-stationary time series does not have a constant mean, variance or auto covariance structure.

It is essential, that variables that are non-stationary, to be treated differently from those that are stationary. This is mainly because of two reasons (Brooks C. Introductory Econometrics for Finance 2014). Firstly, "shocks" to a stationary series will gradually die away, meaning that a shock during time t will have a smaller effect in time $t + 1$ and so on. By contrast shocks in non-stationary data will not have a smaller effect as time goes by. Secondly, the use of non-stationary data can lead to "spurious regressions" and that means that standard regressions techniques could lead to false results and unusually high (or low) R^2.

Non-stationary can be depicted by these models:

The random walk model with a drift $y_t = \mu + y_{t-1} + u_t$, the trend-stationary process $y_t = a + \beta_t + u_t$ and the random walk with a drift generalized $y_t = \mu + \varphi y_{t-1} + u_t$. Based on this last model, for a non-stationary series it is assumed that $\varphi = 1$ where for a stationary one $\varphi < 1$.

We can use the graph of the BDI, our dependent variable, as an example of a non-stationary time series. This variable exhibits the properties of a random walk and does not have a constant mean.

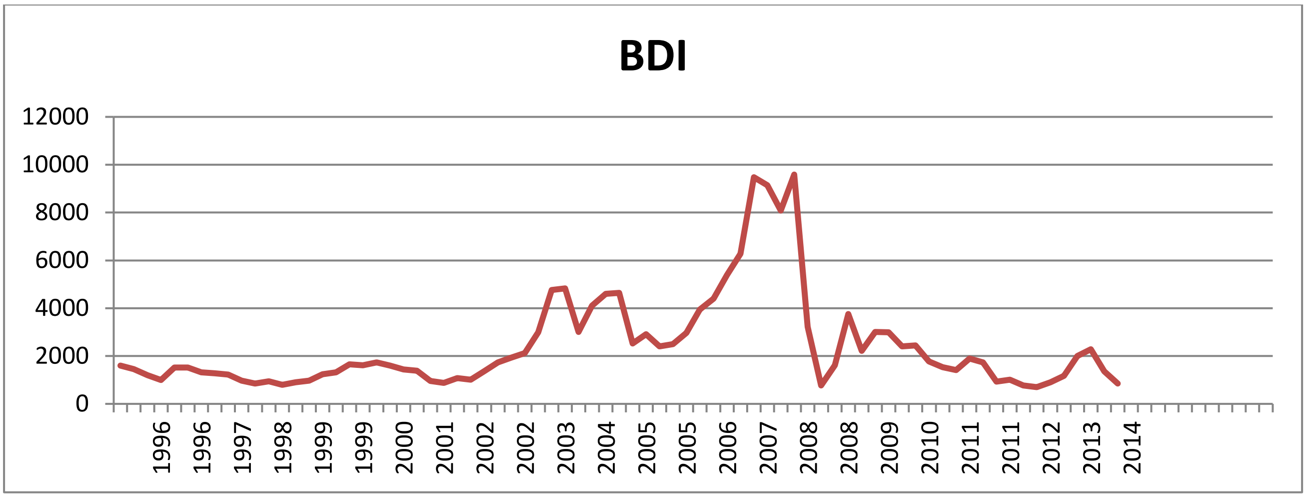


Figure 4.1 BDI Time series chart

Having said that, we need to define our time series/variables based on their order of integration and this can be done by testing for a unit root. Economic theory suggests that there exists a long-run relationship between non-stationary series and if these variables are I (1) or integrated of order 1 (meaning non-stationary and must be differenced once $\Delta Y = Y_t - Y_{t-1}$ to become stationary) then co-integrating techniques can be employed to model this relationship (explanation provided below). Therefore, unit root testing is the first step towards this model where the existence of a unit root implies that a series is non-stationary. For this purpose we will employ two tests, the Augmented Dickey Fuller test and the Phillips-Perron test. The Dickey Fuller Test (Dickey and Fuller 1981) was one of the early ones for unit root testing however we will use it in combination with the PP test since it may fall short in case there is a structural break in the intercept or slope of the regression, something that the Phillips-Perron test can "cover". For higher order models include more than one lags, which is our case here, we will use the Augmented Dickey Fuller test. Starting with an auto-regressive model of the form $y_t = \beta y_{t-1} + \varepsilon_t$ and subtract y_{t-1} from both sides, a new equivalent equation arises $Y_t = \gamma y_{t-1} + \varepsilon_t$, where $\gamma = \beta - 1$. Both tests follow the same procedure using a hypothesis test and examining the null hypothesis that $H_0\ \varphi = 1$ against the one-sided alternative that $H_A\ \varphi < 1$ in the generalized model we mentioned before. These tests can be conducted allowing for an intercept or not, with or without trend. If the pseudo t-stat is lower than the DF critical value (obtained from the test tables) then we reject the null hypothesis of non-stationarity and fail to reject in the opposite case.

We perform both of these tests for our variables in the 5% significance level for lags 0, 1 and 2 with trend and again without trend. The optimal number of lags for these tests is somewhat arbitrary and tends to be as much art as science. A common rule of thumb is to make an initial guess based on the frequency of the data so in our case the common choice is 4 lags based on quarterly data. Then try some alternatives to make sure the results are not highly sensitive to your choice of lag length. The software we will use is STATA which is a data analysis and statistical software and therefore our results and output will be based on that. Regarding the unit root tests and the Augmented Dickey Fuller in specific, STATA provides the following output:

```
. dfuller BDI, trend lags(0)

Dickey-Fuller test for unit root                   Number of obs   =        74

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.315            -4.097            -3.476            -3.166
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.4257
```

Figure 4.2 BDI Time series chart

We use BDI again as an example in order to perform the Augmented Dickey Fuller test for unit root. The previous command shows that we decided to include trend and 0 lags. The result is a test statistic of -2.315 which, according to what we said before about the hypothesis testing procedure, is lower in absolute terms than the critical value at the 5% confidence level (-3.476) and therefore we fail to reject the null hypothesis that there is a unit root and conclude that our variable is non-stationary and at least I (1). We follow the same procedure for lags 1 and 2 and then again removing trend. This is done for all our time series and then depending on the results we conclude whether a variable is stationary or not.

When a variable is found to be non-stationary or I (1), that means that it needs to be differenced once to become stationary or in other words, ΔY_t is stationary. Therefore, we generate the first-difference for all variables that have a unit root according to our previous tests and follow the same procedure for each new generated variable (ADF and Phillips-Perron tests) to examine if they continue to have a unit root. We can create a similar table with the test results and conclude about the stationary of the new variables. Assuming that none of our time series had more than 2 unit roots we do not examine higher orders of integration.

4.2 Co-integration model (VECM) – Johansen approach

Having determined the order of integration of our variables we can move on with examining the existence of a long-run relationship between them. In order to explain this relationship or dynamic co-movement of variables we can use an example. Let's assume that two series are individually non-stationary and integrated of order 1 or I (1). This means that their changes are I (0) and stationary and usually a linear combination of these variables will still be I (1). However, if there exists a linear combination of X and Y such that $Z = X - aY$ is I (0), then X and Y are said to be co-integrated. In general and according to Engle and Granger (1987) the components of a vector Z are said to be co-integrated of order (d, b) , denoted as $Z \sim CI(d-b)$ if all components of Z are I(d) and if there exists at least one vector β (≠0) such that $\beta Z{\sim}I(d-b), b > 0$ where β is called the co-integrating vector. Even though these two-time series are individually non-stationary, there is still a linear combination of them which is I (0) and acts as the equilibrium. Time series may "move together" over time even if they are non-stationary, because of some influences like market conditions, which implies that they are bound by some relationship in the long run. More formally and according to Engle-Granger if a vector of time series is I(d) but there exists a linear combination which is integrated to a lower order, the time series is said to be co-integrated.

Financial theory suggests that in practice many financial variables contain one unit root and that two or more time series are expected to hold some long-run relationship with one another, for example in spot and future prices of commodities. In this example, we have prices for the same asset at different points in time and we therefore expect a strong relationship between them. They will both be affected by the same information/variables and maintain a long-run relationship or in other words an equilibrium relationship and a constant mean that would be returned to frequently.

In order to avoid having variables with similar statistics (mean, variance, correlation etc) we decided to reduce the number of them based on the following procedure. One of our main goals is to remove multicollinearity[1], therefore the list of our independent variable was divided into four categories such that each category is fully represented in the final model and factors within the same category that are highly correlated can be combined and represented by a single one. Below is a table that represents that:

[1] A phenomenon in which two or more independent variables in a multiple regression are highly correlated. This may affect the coefficient estimates and the model in general.

C1	C2	C3	C4
OECD World Trade Vol.	VIX	EURUSD Exchange rate	Bulk Newbuilding price
	US Economic Policy Uncertainty	USFedFundsRate	Tank Newbuilding price
	EU Economic Policy Uncertainty	Brent Oil Price	Bulker secondhand price
	MSCI	CPI	Tank secondhand price
			Bulk Fleet Development
			Tank Fleet Development
			Bulk Scrap price
			Tank Scrap price

Table 1: Categories of variables

Group of variables were developed that measured the world trade meaning the demand driving factors C1, the world economic indicators C2, prices and interest rates driving factors C3 and finally the supply driving factors C4. As mentioned before, variables were selected based on their correlation, so for example we expect that scrap prices for either bulk carriers or tankers would have a high correlation coefficient. A correlation matrix between the percentage returns of all variables can be found in the appendix. We therefore conclude with our six independent variables and BDI our dependent one.

Moving on to the actual process of co-integration, as we mentioned before, it is a relationship between time series variables. Assuming that all time series are integrated of order one, if there are coefficients α, β, γ such that the linear combination $ax + \beta y + \gamma z$ is integrated of order 0, then x, y and z are co-integrated. We will use the Johansen technique based on Vector Auto regression models in order to test and estimate co-integrating systems. Suppose that we have a set of g variables, all of them are integrated of order one or I (1), which we are examining for co-integration. A vector auto regressive model with k lags containing these variables could be the following: $y_t = b_1 y_{t-1} + b_2 y_{t-2} + \ldots . + b_k y_{t-k} + u_t$. In order to use the Johansen test, the previous model needs to be turned into a vector error correction model (VECM): $\Delta y_t = \Pi y_{t-k} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \ldots .. + \Gamma_{k-1} \Delta y_{t-(k-1)} + u_t$ where $\Pi = (\sum_{i=1}^{k} b_i) - I_g$ and $\Gamma_i = (\sum_{j=1}^{i} b_j) - I_g$. This model has g variables in first differenced form on the left hand side and k-1 lags of the dependent variables on the right hand side, each with a Γ coefficient matrix. The Johansen test focuses on the examination of the Π matrix which is interpreted as the long-run coefficient of the matrix. According to this method, the rank of this matrix determines the existence of co-integration, meaning that the rank of Π is equal to its characteristic roots or eigenvalues that are different from zero. If the rank of Π is equal to 0 then there is no co-integrating relationship between our variables and the VECM simply reduces to a VAR model in first difference form. On the other hand, if Π has a full rank (g) then this would correspond to the original VAR and all our time series would be stationary. For 1< rank (Π) < g there are r co-integrating vectors and Π is the product of two matrices $\Pi = \alpha\beta'$, α (g x r) and β' (r x g). There are two tests that help us determine the existence of co-integration and the number of co-integrating vectors in case co-integration exists applying an estimation technique such as Johansen (1991). Under this approach the two test statistics are:

$\lambda_{trance}(r) = -T \sum_{i=r+1}^{g} \ln(1-\hat{\lambda}_i)$ and $\lambda_{max}(r, r+1) = -T \ln(1-\widehat{\lambda_{r+1}})$.

Where T is the number of observations, r is the number of co-integrating vectors under the null hypothesis and λ_i is the estimated value of the characteristic roots or eigenvalues obtained from the Π matrix. λ_{trance} tests the null hypothesis that the number of co-integration vectors is less than or equal to r against a general alternative where λ_{max} tests that null hypothesis that the number of vectors is equal to r again the alternative of r+1 vectors. Johansen and Juselious (1990) provide critical values for the two statistics and these depend on the value of $g - r$. In STATA, this step is executed using the *vecrank* command, where a table compares λ with the tabulated critical values and presents the number of co-integrating vectors.

Having determined the co-integrating rank of our model, we move on to the actual representation where we apply the Vector Error Correction Model (VECM) and produce our final results which are the coefficients and other statistics regarding the relationship between our variables.

The final VECM equation provides the information regarding the I (1) variables, the input into the model and the choice for this type of modeling has a few advantages. Firstly, the Johansen approach and tests are considered more robust if there is presence of heteroskedastic disturbances (Lee and Tse, 1995) or non-normality (Cheung and Lai, 1993). Also, it is more efficient regarding the coefficient estimations compared to other models. Most importantly, ordering of the variables in the model does not affect the results which is not the case in the Engle-Granger approach (1987).

4.3 Hypothesis Testing – Lagrange Multiplier Test

Once we have estimated a model, we are interested in the significance and the accuracy of the values and coefficients obtained. A first and simple test is to examine the significance of each coefficient in the model using a t-test or an F-test to examine how a set of coefficients behaves jointly. “Passing” these initial test, our model we can say that the theory is consistent with the data (Popperian principle).

Estimations of the VECM are based on the assumption that the errors are not auto-correlated. The Lagrange Multiplier test examines this hypothesis as discussed in Johansen (1995). The formula for the LM test statistic for j lags is $LM_s = (T - d - 0.5)\ln\left(\frac{|\hat{\Sigma}|}{|\widetilde{\Sigma_s}|}\right)$ where T is the number of observations in the VAR (or the VECM in our case), d the number of coefficients in the augmented VAR, $|\hat{\Sigma}|$ is the maximum likelihood estimate of Σ, the variance-covariance matrix from the VAR and finally $\widetilde{\Sigma_s}$ is the maximum likelihood estimate of Σ from the augmented VAR. If we assume that we have K equations in the VAR we can define e_t to be K x 1 vector of residuals and after we create the K new variables $e_1 to\ e_k$ we can augment the original VAR with lags of these K new variables. For each lag we form a new regression with s lags where the missing values from these s lags are replaced with zeros. The asymptotic distribution of LM is x^2 with k^2 degrees of freedom.

4.4 Granger Causality, Impulse responses

Forecasting is the heart of time-series analysis. We want to be able to anticipate and envision the future. Non-stationary time series, like the ones we deal with at this paper, are often dominated by a trend which is the main driving factor behind future prices. Thus, an understanding of trends is the key to longer - horizon forecasts. Properties of the VAR model are usually summarized using structural analysis and Granger causality, impulse responses (IRF) and forecast error variance decompositions (FEVD). These are tools for analyzing the relations between the variables in a dynamic econometric model.

To further improve upon our findings from the VECM procedure we will apply and analyze the impulse response function of the VECM. The impulse response functions are regularly applied to capture the propagation mechanism of a shock across time (Sims, 1980). More specifically, in contemporary macroeconomic modeling, impulse response functions are used to describe how one variable will react over time to exogenous impulses (shocks). In other words, if the value of one of the endogenous variables, the EUR/USD exchange rate in our model, for instance, changes by one standard error, the impulse response function will predict how the value of the variable that we are analyzing, in our case the BDI, will react over time. Impulse-responses trace the effects of structural shocks on endogenous variables. Each response includes the effect of a specific shock on one of the variables of the system at time t, then on t+1 and so on. The goal here is to transform our structural autoregressive vector $AX_t = \beta_0 + \beta_1 X_{t-1} + u_t$ into a sum of shock or Word representation $X_t = \mu + \sum_{i=0}^{\infty} c_t u_{t-t}$. IRS not only answers to what will be the degree of the effect of one variable change to the other, but also how this will change evolve over time.

According to Sims (1980), impulse responses can be computed by orthogonalizing the underlying shocks to the model using a Cholesky Decomposition of the variance-covariance matrix. However, this approach will lead to impulses, which are not unique and depend on the ordering of the variables in the system (Lutkepohl, 1991). One generally accepted solution to this problem has been proposed by Pesaran and Shin (1998). They proposed the use of Generalized Impulse Responses (GIR).

IRS functions proceed after the VECM procedure has concluded. We use the *irf create* command after the fitting of our VEC model. "*Irf create*" estimates multiple sets of impulse-response functions (IRFs), dynamic-multiplier functions, and forecast-error variance decompositions (FEVDs) after estimation by *var, svar, or vec* (STATA Documentation). So we get a full set of estimates including the FEVDs. These estimates and their standard errors are known collectively as IRF results and are saved in an IRF file under the specified *irfname*, in our case "*myirf.irf*"

4.5 Variance Decomposition

This is an alternative method to the impulse response functions for examining the effects of shocks to the dependent variables. This technique determines how much of the forecast error variance for any variable in a system, is explained by innovations to each explanatory variable, over a series of time horizons. In other words, it gives us information about the proportion of movements of a variable due to shocks to itself and shocks to other variables. Usually own series shocks explain most of the error variance, although the shock will also affect other variables in the system. It is also important to consider the ordering of the variables when conducting these tests, as in practice the error terms of the equations in the VAR will be correlated, so the result will be dependent on the order in which the equations are estimated in the model. Also, the variance of the forecast error increases with the horizon which is consistent with the fact that there is more uncertainty the further we are attempting to forecast.

5. Results

5.1 Time series analysis

The first step towards an in-depth analysis of the relationship between our variables is to truly understand their nature. We have assumed that there exists some kind of relationship in previous chapters based on economic assumptions and the shipping literature; however we need to examine each individual variable in order to generate accurate results. Our examination begins with plots regarding our time series variables which will give us the ability to identify potential trends. From the graphs, available in the appendix, we observe that there are significant differences in the level values of the time series for many variables. Some of them have a clear trend towards one direction throughout the sample period while others display a behavior which may be closer to a random walk. The most important issue however, is the mean. It is clear that for most time series there is evidence for a non-zero mean as well as a trend.

Total Tanker Fleet Development for example, displays a constant upwards trend throughout the sample period while its mean is non-zero and does not remain constant and the same holds for the OECD World Trade Volume. On the other hand if we look at the BDI, our dependent variable, we observe that it resembles a random walk variable where the next observation is independent and unpredictable. There are still some trends for specific parts of the sample but do not last long. Another example is the MSCI index, where it appears that there is cyclicality. During 2000 MSCI reached its maximum value and then dropped again until reaching an even higher value during 2009. Then we observe the same pattern again, dropping in 2010 and maintaining an upwards trend as we go towards the future.

It is obvious that each variable has specific statistical properties and it is a great challenge to identify the relationship between them, when there doesn't seem to be one. Hopefully, there are some sophisticated tools that help us do that and we will start by examining the stationary of our variables.

5.2 Stationarity Investigation – Unit Root Testing

Based on the methodology we presented in previous chapters, we will investigate the stationarity of each time series, by examining the presence of unit roots, using the Augmented-Dickey Fuller and the Phillips Peron tests. It is important that variables which are stationary be treated different from those which are not; therefore a distinction is an important first step for further analysis.

Regarding the tests, we decided to perform them without trend due to uncertainty in our data. Also, since we have quarterly data and as we mentioned before, each variable has only 75 observations, our decision of the lag length was not so easy to make. Based on the frequency, a lag order of 4 would fit better. Eventually we went for 2 lags (covering 6 months) in order not to sacrifice many degrees of freedom in our tests. We must note here, that we are working within the 5% confidence level for the entire paper. Beginning with the Augmented-Dickey Fuller test, we get the following results.

Panel A: ADF Test (Z_t statistic) Without Trend, lags(2)		
Variable	Z_t	**Description**
BDI	-2.357	Non-stationary
VIX	-2.839	Non-stationary
EUECONOMICPOLICYUNCERTAINTY	-2.147	Non-stationary
USECONOMICPOLICYUNCERTAINTY	-2.726	Non-stationary
USFEDFUNDSRATE	-1.566	Non-stationary
EURUSDECHANGERATE	-1.296	Non-stationary
OECDWORLDTRADEVOLUME	-0.615	Non-stationary
BrentOilPricebbl	-0.699	Non-stationary
BulkcarrierNewbuildingPriceInd	-1.993	Non-stationary
OilTankerNewbuildingPrices	-1.827	Non-stationary
TankerSecondhandPriceIndex	-1.95	Non-stationary
BulkCarrierSecondhandPriceInd	-2.114	Non-stationary
TotalTanker10kDWTFleetDev	1.12	Non-stationary
TotalBulkcarrierFleetDevelopme	0.704	Non-stationary
BulkerScrapFarEastCapesizeP	-1.363	Non-stationary
TankerScrapFarEast	-1.429	Non-stationary
MSCI	-1.922	Non-stationary
CPI	1.729	Non-stationary

For each test, we get the t statistic which is then compared to the tabulated critical value of the Augmented-Dickey Fuller test. According to the test, the null hypothesis that the variable is non-stationary can be rejected if the t-statistic is more negative than the critical value (Fuller, 1976). All t statistics are lower than the critical value (in absolute terms) and therefore we fail to reject the null hypothesis and conclude that all of our variables are non-stationary.

We further investigate this issue with the help of the Phillips-Perron test. Following the same procedure, we get the table below.

Panel B: PP Test (Z_t statistic) Without Trend, lags(2)		
Variable	$\mathbf{Z_t}$	**Description**
BDI	-2.488	Non-stationary
VIX	-4.479	Stationary
EUECONOMICPOLICYUNCERTAINTY	-4.484	Stationary
USECONOMICPOLICYUNCERTAINTY	-4.986	Stationary
USFEDFUNDSRATE	-1.268	Non-stationary
EURUSDECHANGERATE	-1.476	Non-stationary

OECDWORLDTRADEVOLUME	-0.545	Non-stationary
BrentOilPricebbl	-0.974	Non-stationary
BulkcarrierNewbuildingPriceInd	-1.383	Non-stationary
OilTankerNewbuildingPrices	-1.344	Non-stationary
TankerSecondhandPriceIndex	-1.67	Non-stationary
BulkCarrierSecondhandPriceInd	-2.045	Non-stationary
TotalTanker10kDWTFleetDev	2.685	Non-stationary
TotalBulkcarrierFleetDevelopme	6.239	Non-stationary
BulkerScrapFarEastCapesizeP	-1.669	Non-stationary
TankerScrapFarEast	-1.764	Non-stationary
MSCI	-1.662	Non-stationary
CPI	1.35	Non-stationary

According to the Phillips-Perron test, the null hypothesis is non-stationarity. Again, we only reject this hypothesis when the t statistic is more negative than the Critical value at the 5% confidence level. Highlighted in yellow, are the variables that are stationary with the rest being non-stationary. In this test, we observe that three of our variables, VIX, EU and US Economic Policy Uncertainty are stationary which is in contradiction with the results from our previous ADF test. In order to solve this we must take a closer look at the critical values of the first ADF test for these three variables where we observe that all there have a t statistic that is very close to the critical value. Given our limited sample and the fact that we only slightly fail to reject non-stationary for the ADF test, we can say that for the purpose of this analysis, these three variables can be considered stationary.

We found out that most of our data set is non-stationary which means that many of the statistical properties change over time. Looking at the graphs we can observe that most of our time series follow a random path or a trend and the mean changes over time depending on the time period we examine. So, both the variance and the mean are time-dependent while the variance goes to infinity as time approached infinity. For the three variables which are stationary, the previous assumptions do not hold. These revert around a constant mean while there is not a specific trend.

Concluding about the stationarity of our variables, we observe that most of our time series have a unit root. Based on financial theory, this is the case for most variables and it is something to be expected. Examining the variables which are non-stationary we can say that they are I (1) but this does not guarantee that the series does not have a second unit root. Therefore, we will investigate this possibility by producing the first difference of each factor. Generating the difference $\Delta Y_t = Y_t - Y_{t-1}$ for all I (1) variables we will again investigate the presence of a unit root. Following the same procedure, we perform both unit root tests, ADF and PP, at lag order 2 without the presence of trend and get the following table.

First Difference Form	Panel A: ADF Test		Panel B: PP Test	
Variable	Z_t	2nd Difference Z_t	Z_t	2nd Difference Z_t
BDI	-4.856		-8.016	
USFEDFUNDSRATE	-3.514		-7.525	
EURUSDECHANGERATE	-4.9		-8.236	
OECDWORLDTRADEVOLUME	-4.58		-4.47	
BrentOilPricebbl	-5.889		-6.748	
BulkcarrierNewbuildingPriceInd	-3.495		-4.431	
OilTankerNewbuildingPrices	-3.589		-5	
TankerSecondhandPriceIndex	-3.761		-6.264	
BulkCarrierSecondhandPriceInd	-4.375		-6.804	
TotalTanker10kDWTFleetDev	-2.814	-7.156	-3.857	
TotalBulkcarrierFleetDevelopme	-1.226	-5.293	-1.476	-11.133
BulkerScrapFarEastCapesizeP	-5.624		-9.853	
TankerScrapFarEast	-5.561		-10.112	
MSCI	-4.401		-7.253	
CPI	-5.589		-7.656	

Highlighted in yellow are the variables that have a t statistic that is more negative than the critical value and therefore we reject the hypothesis of non-stationarity. However, there are two variables according to the ADF test, the Tanker Fleet Development (TotalTanker10kDWTFleetDev) and the Bulk Carrier Fleet Development (TotalBulkcarrierFleetDevelopme), that appear to have a second unit root. Differencing for the second time and performing the ADF test, confirms that these two variables are indeed integrated of order 2 or I (2). Except for these two controversial variables, both tests produce the same results for the rest and we can confirm that they are integrated of order 1 and contain one unit root. Regarding the Tanker Fleet Development and the Bulk Carrier Fleet Development, Phillips-Perron tests concludes that the first is stationary and I(1) while the second one is not and is later confirmed that it is integrated of order 2. Again, we can solve this inconsistency of the Tanker Fleet Development variable by looking at the t statistic and the critical value of the ADF test. We observe that we fail to reject the hypothesis by a small margin and in combination with a relatively small sample size; it is difficult to make a firm decision. On the other hand, the Phillips-Perron test gives us a definitive answer about the order of integration and due to the uncertainty of the ADF test we will conclude that the Tanker Fleet Development variable is integrated of order 1.

To conclude this part of the analysis, we found that three of our variables are stationary and I (0) and these are, VIX, EU and US Economic Policy Uncertainty. The majority of the data set however is non-stationary and I (1) which are the BDI, the US Federal Funds Rate, the EUR/USD Exchange rate, OECD World Trade Volume, Brent Oil Price, Bulk carrier new-building price index, Tanker new-building price index, Bulk carrier second-hand price index, Tanker second-hand price index, Tanker Fleet Development, Bulk Carrier Scrap Price, Tanker Scrap Price, MSCI and finally CPI. We found that only the Bulk Carrier Fleet Development is non-stationary and integrated of order 2.

The next step is to proceed to the reduction of the dimension of our model. This means that we want to keep the most important and significant variables in our final model. To do this, we categorized our candidate variables into four distinct macroeconomic categories depending on their nature and role in the macroeconomic theory. We discussed this notion in the previous chapter. Here is again the table that contains all the variables divided in their respective categories.

C1 (Demand driving factors)	C2 (Word Economic Indicators)	C3 (prices and interest rates driving factors)	C4 (supply driving factors)
OECD World Trade Vol.	VIX	EURUSD Exchange rate	Bulk Newbuilding price
	US Economic Policy Uncertainty	USFedFundsRate	Tank Newbuilding price
	EU Economic Policy Uncertainty	Brent Oil Price	Bulker secondhand price
	MSCI	CPI	Tanker secondhand price
			Bulk Fleet Development
			Tank Fleet Development
			Bulk Scrap price
			Tank Scrap price

Table 1: Categories of variables

We generated a correlation matrix between every single pair of the percentage returns of the variables (see appendix for full info), that can help us trim off our model any redundant variables.[2] For example, the correlation between the movements of the “Bulk Scrap price” and “Tank Scrap price” variables is about 0.94. Indeed, as their combined graphical representation suggests, these two variables are almost alike, thus we can safely exclude one of them from our model.[3]

[2] The choice of percentage returns has been made because we are interested in the movements of the variables and how these changes are similar or not.

[3] A high correlation between two variables may suggest a strong relationship between them but it is not always an indicator of causality between the two.

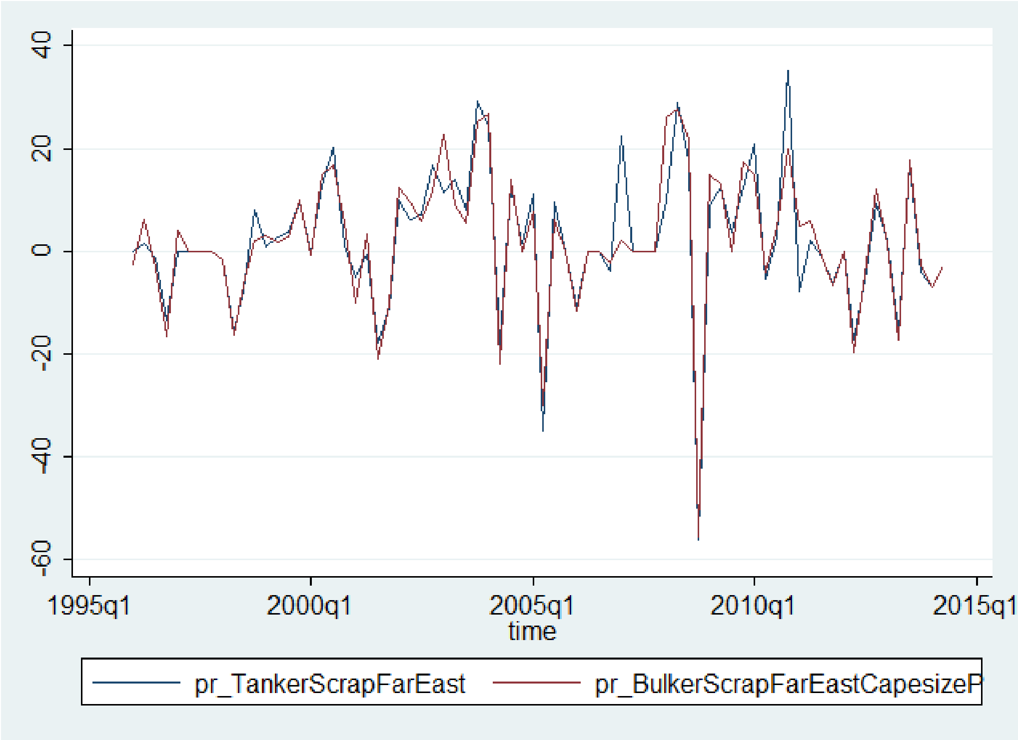


Figure 1: Scrap prices

In the same manner, we decided to exclude from our model variables that show a correlation of over 50% with another variable. Our final minimized model will include these variables from each category:

C1 (Demand driving factors)	C2 (Word Economic Indicators)	C3 (prices and interest rates driving factors)	C4 (supply driving factors)
OECD World Trade Vol.	MSCI	EURUSD Exchange rate	Bulk Newbuilding price
		CPI	Tanker Fleet Development

Table 3: Final Categories for modeling

A few notes: The Tanker Fleet Development might seem like an odd choice to include in a model that analyzes the Baltic Dry Exchange Index, which is an index used in the bulk carrier market. The Bulker Fleet Development though, came out as an I (2) variable, so it could not be used in the Johanssen approach of the Vector Error Correction Model (which requires all variables to be integrated of order 1). Moreover, their respective correlation of their returns was computed to be almost 50% which is an extremely high value especially for real datasets like the one we are analyzing. So, we could safely disregard it from our analysis and use the Tanker Fleet Development instead, without any meaningful loss of credibility for our model.[4]

Lastly, for practical reasons, we needed to scale the Euro-USD exchange rate and the CPI index up accordingly. The Euro/USD rate was scaled up with a factor of 1000 (*EURUSDECHANGERATE_new*) and the CPI with a factor of 10 (CPI_new).
This will not affect the results of our model fitting; it will help us with a better presentation of the coefficients (for them to be in about the same scale). The co-integrating relationships that may exist will not be affected qualitatively.

5.3 VECM – Johansen's Approach

We are now ready to fit the VECM. We can rewrite a general, p_{th}-order VAR for a K-element vector y_t as a Vector Error Correction Model.

$$\Delta y_t = \Pi y_{t-k} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \ \ldots.. + \Gamma_{k-1} \Delta y_{t-(k-1)} + \ u_t$$

where $\Pi = (\sum_{i=1}^{k} \ b_i \ ) - I_g$ and $\Gamma_i = (\sum_{j=1}^{i} \ b_j \ ) - I_g$.

We can also rearrange some terms in order to make the possible co-integrating relationships better presented.

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \ \ldots.. + \Gamma_{k-1} \Delta y_{t-(k-1)} - \ \Pi y_{t-k} + \ u_t$$

Where $\Gamma_i = -(I - A_1 - \cdots - A_i)$ i=1,2,..,k-1 and Π $= -(I - A_1 - \cdots - A_k)$

We discussed the theory behind the VECM approach in the previous chapter and what each matrix means. Let us now employ the method to our specific model.

$$BDI = f(WorldTradeVolume, \frac{Euro}{USD} rate, CPI, Bulk\ carrier, Newbuilding\ Prices, TankerFleetDevelepoment)$$

5.4 Pre-Estimation of VECM

Step 1: Identify the number of co-integrating relationships.

The first part of the Johansen approach is to estimate the rank of the Π matrix. The rank of αβ'specifies the number of co-integrating relationships. As mentioned previously, the number of these relationships

[4] . For the booming period.

lies between 0 and $k-1$, where k is the number of our variables. Therefore, we expect that $0 < r < 6$. The *vecrank* command calculates two statistics, the λ_{trace} and λ_{max} and some other information criterion tests, utilizing the SBIC, HQIC and AIC criteria, in the final table.

```
. vecrank BDI CPI_new MSCI EURUSDECHANGERATE_new OECDWORLDTRADEVOLUME BulkcarrierNewbuildingPriceInd TotalTanker10kDWTFleetDev, trend
> (constant) max ic levela

                       Johansen tests for cointegration
Trend: constant                                   Number of obs =       73
Sample:  1996q2 - 2014q2                                   Lags =        2
-------------------------------------------------------------------------------
maximum                                      trace     5% critical  1% critical
  rank    parms      LL        eigenvalue  statistic      value        value
    0      56    -2371.9034                172.6549      124.24       133.57
    1      69    -2341.3607    0.56690     111.5695       94.15       103.18
    2      80    -2319.8352    0.44553      68.5183*1*5   68.52        76.07
    3      89    -2303.9894    0.35217      36.8269       47.21        54.46
    4      96    -2296.2054    0.19205      21.2588       29.68        35.65
    5     101    -2289.1106    0.17665       7.0693       15.41        20.04
    6     104    -2285.8897    0.08446       0.6274        3.76         6.65
    7     105     -2285.576    0.00856
-------------------------------------------------------------------------------
maximum                                       max      5% critical  1% critical
  rank    parms      LL        eigenvalue  statistic      value        value
    0      56    -2371.9034                 61.0854       45.28        51.57
    1      69    -2341.3607    0.56690      43.0512       39.37        45.10
    2      80    -2319.8352    0.44553      31.6915       33.46        38.77
    3      89    -2303.9894    0.35217      15.5680       27.07        32.24
    4      96    -2296.2054    0.19205      14.1896       20.97        25.52
    5     101    -2289.1106    0.17665       6.4419       14.07        18.63
    6     104    -2285.8897    0.08446       0.6274        3.76         6.65
    7     105     -2285.576    0.00856
-------------------------------------------------------------------------------
maximum
  rank    parms      LL        eigenvalue     SBIC         HQIC        AIC
    0      56    -2371.9034                 68.27497    67.21812    66.5179
    1      69    -2341.3607    0.56690      68.20224*   66.90005    66.03728
    2      80    -2319.8352    0.44553        68.259    66.74922    65.74891
    3      89    -2303.9894    0.35217      68.35383     66.6742*   65.56135
    4      96    -2296.2054    0.19205      68.55199    66.74025    65.53987
    5     101    -2289.1106    0.17665      68.65147    66.74538    65.48248
    6     104    -2285.8897    0.08446      68.73955    66.77684    65.47643
    7     105     -2285.576    0.00856      68.78973    66.80815    65.49523
-------------------------------------------------------------------------------
```

For each method STATA places an asterisk by the test statistic associated with the recommended rank. Both the trace statistic and the max statistic suggest a rank of 2. The trace statistic is lower than the critical value at rank 2 and the same is true for the max statistic. This confirms that there will be 2 co-integrating relationships in our model which justifies the use of a VECM.

Step 2: Identify the number of lags

Next, we want to identify the number of lags in the VECM from of our model. We will perform the *varsoc* test and for each method STATA places an asterisk by the test statistic associated with the recommended lag length.

```
. varsoc BDI CPI_new MSCI EURUSDECHANGERATE_new OECDWORLDTRADEVOLUME BulkcarrierNewbuildingPriceInd TotalTanker10kDWTFleetDev, lutsta
> ts

   Selection-order criteria (lutstats)
   Sample:  1996q4 - 2014q2                     Number of obs      =        71
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-3031.13				3.5e+28	65.5189	65.5189	65.5189
1	-2310.48	1441.3	49	0.000	2.1e+20	46.5992	47.2201	48.1607*
2	-2225.69	169.59	49	0.000	8.0e+19	45.5908	46.8328*	48.7139
3	-2174.4	102.57	49	0.000	8.3e+19	45.5264	47.3894	50.2111
4	-2095.69	157.43*	49	0.000	4.4e+19*	44.6894*	47.1733	50.9357

```
   Endogenous:  BDI CPI_new MSCI EURUSDECHANGERATE_new OECDWORLDTRADEVOLUME
                BulkcarrierNewbuildingPriceInd TotalTanker10kDWTFleetDev
    Exogenous:  _cons
```

The likelihood ratio test suggests four lags, as well as the Akaike's information criterion. HQIC suggests a lag order of two and the Schwarz Bayesian Information Criterion suggest only one lag. These tests can be very sensitive to the maximum lag length considered, so, to be sure, we rerun these tests with a maximum of eight lags.

```
. varsoc BDI CPI_new MSCI EURUSDECHANGERATE_new OECDWORLDTRADEVOLUME Bulkcarrier
> NewbuildingPriceInd TotalTanker10kDWTFleetDev, maxlag(8) lutstats

   Selection-order criteria (lutstats)
   Sample:  1997q4 - 2014q2                     Number of obs      =        67
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-2849.74				2.6e+28	65.2017	65.2017	65.2017
1	-2183.81	1331.9	49	0.000	2.6e+20	46.7859	47.4239	48.3983*
2	-2104.64	158.33	49	0.000	1.1e+20	45.8854	47.1615	49.1102
3	-2052.2	104.9	49	0.000	1.1e+20	45.7825	47.6966	50.6196
4	-1968.86	166.66	49	0.000	5.2e+19	44.7577	47.3098	51.2072
5	-1891.05	155.62	49	0.000	3.5e+19	43.8977	47.0878	51.9596
6	-1837.03	108.05	49	0.000	6.5e+19	43.7477	47.5759	53.422
7	-1738.79	196.48	49	0.000	6.0e+19	42.2779	46.7441	53.5646
8	-1468.3	540.97*	49	0.000	1.2e+18*	35.6664*	40.7706*	48.5654

```
   Endogenous:  BDI CPI_new MSCI EURUSDECHANGERATE_new OECDWORLDTRADEVOLUME
                BulkcarrierNewbuildingPriceInd TotalTanker10kDWTFleetDev
    Exogenous:  _cons
```

Lutkepohl (2005) has shown that the Hannan and Quinn's information criterion (HQIC) and the Schwarz Bayesian Criterion (SBIC) statistics provide more consistent estimates of p, the true lag length while the FPE and the AIC statistics overestimate p even for infinite samples. The rerun of the test shows exactly that. All criteria except of the SBIC seem very sensitive to the maximum lag length employed in the test, while only the SBIC remains consistent to its suggestion. For all the above reasons, in combination with our limited sample size, we decided to settle to one lag length for our modeling.

5.5 VECM Estimation

We estimated that the optimal lag order for the VECM would be a lag order of 1 and the suggested rank of the Π matrix would be 2. We are ready to fit the VECM. We execute the *vec* command in STATA, indicating a lag order of 1 and 2 co-integrating relationships. Also, we specified that we have a constant trend, which defines a model with an unrestricted constant (Johansen notation $H_1(r)$, 1995). This allows for the undifferenced data and the co-integrating equations to be stationary around a non-zero mean and it is the default option. The most important output regarding our analysis is presented here. The rest of the model results can be found in the appendix of the paper.

```
. vec BDI CPI_new MSCI EURUSDECHANGERATE_new OECDWORLDTRADEVOLUME BulkcarrierNewbuildingPriceInd TotalTanker10kDWTFleetDev, trend(con
> stant) rank(2)

Vector error-correction model

Sample:  1996q2 - 2014q2                          No. of obs      =        73
                                                  AIC             =  65.74891
Log likelihood = -2319.835                        HQIC            =  66.74922
Det(Sigma_ml)  =  9.44e+18                        SBIC            =    68.259

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
D_BDI                10    1079.15   0.1892   14.46658   0.1528
D_CPI_new            10    3.11441   0.9026   574.6093   0.0000
D_MSCI               10    92.8867   0.2594   21.71886   0.0166
D_EURUSDECHANG~w     10    61.4295   0.1151   8.067874   0.6222
D_OECDWORLDTRA~E     10    42.6451   0.7906   234.1398   0.0000
D_BulkcarrierN~d     10     5.5242   0.6470   113.6351   0.0000
D_TotalTanker1~v     10    1.96364   0.8292   301.0828   0.0000
----------------------------------------------------------------
```

Figure 1: Sample information

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D_BDI						
_ce1 L1.	-.3677713	.1962794	-1.87	0.061	-.7524718	.0169292
_ce2 L1.	8.833176	6.996291	1.26	0.207	-4.879302	22.54565
BDI LD.	.1971739	.1749535	1.13	0.260	-.1457286	.5400763
CPI_new LD.	-64.99102	37.63839	-1.73	0.084	-138.7609	8.778864
MSCI LD.	-.5130351	1.542911	-0.33	0.740	-3.537086	2.511016
EURUSDECHANGERATE_new LD.	2.871193	2.514817	1.14	0.254	-2.057758	7.800144
OECDWORLDTRADEVOLUME LD.	2.894477	2.408613	1.20	0.229	-1.826319	7.615272
BulkcarrierNewbuildingPriceInd LD.	-7.692333	22.92641	-0.34	0.737	-52.62727	37.2426
TotalTanker10kDWTFleetDev LD.	20.07977	49.08843	0.41	0.683	-76.13177	116.2913
_cons	1.842721	460.7099	0.00	0.997	-901.132	904.8174

Figure 2: Short Run Parameters

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	5	163.4089	0.0000
_ce2	5	2322.169	0.0000

Identification: beta is exactly identified

Johansen normalization restrictions imposed

beta	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ce1						
BDI	1	.	.	.	.	.
CPI_new	0	(omitted)				
MSCI	-1.85389	.9683462	-1.91	0.056	-3.751813	.044034
EURUSDECHANGERATE_new	-11.01584	2.267128	-4.86	0.000	-15.45933	-6.57235
OECDWORLDTRADEVOLUME	-2.997094	.6416054	-4.67	0.000	-4.254617	-1.73957
BulkcarrierNewbuildingPriceInd	-3.687074	10.18151	-0.36	0.717	-23.64247	16.26833
TotalTanker10kDWTFleetDev	47.79742	8.485434	5.63	0.000	31.16628	64.42856
_cons	5919.798	.	.	.	.	.
_ce2						
BDI	0	(omitted)				
CPI_new	1	.	.	.	.	.
MSCI	-.0182552	.0231923	-0.79	0.431	-.0637113	.0272009
EURUSDECHANGERATE_new	-.2822453	.0542987	-5.20	0.000	-.3886688	-.1758218
OECDWORLDTRADEVOLUME	-.2137206	.0153667	-13.91	0.000	-.2438389	-.1836024
BulkcarrierNewbuildingPriceInd	1.31741	.2438517	5.40	0.000	.8394693	1.79535
TotalTanker10kDWTFleetDev	.4163099	.2032298	2.05	0.041	.0179868	.8146331
_cons	-59.90071	.	.	.	.	.

Figure 3: Co-integrating Equations

The output provides with two very important pieces of information regarding the short-run and the long-run relationships between our variables. Looking at the second table, Figure 2, we have information regarding the short run part of our analysis. We observe the relationship of ΔBDI as our dependent variable with itself but also other macro-economic factors. The first step towards identifying the importance of those factors is looking at the p-value of each variable in order to determine whether these are statistically significant or not. At the 5% confidence level, none of the variables seem to have a statistically significant relationship with ΔBDI and for the majority of them we strongly reject this hypothesis so there is no point in examining the qualitative and quantitative effect on the ΔBDI. We conclude that the ΔBDI is not Granger Caused by a variable in the short-run. A simplified version of our model is the following equation (1):

$$\begin{aligned}\mathrm{\Delta B\Delta I_t} = {} & 0.19\mathrm{\Delta BDI_{t-1}} - 64.99\mathrm{\Delta CPInew_{t-1}} - 0.51\mathrm{\Delta MSCI_{t-1}} \\ & + 2.87\mathrm{\Delta EURUSDEXCHANGERATEnew_{t-1}} + 2.89\mathrm{\Delta OECDWORLDTRADEVLULE_{t-1}} \\ & - 7.69\mathrm{\Delta BulkcarrierNewbuildingPriceInd_{t-1}} \\ & + 20.07\mathrm{\Delta TotalTanker10kDWTFleetDev_{t-1}} + 1.84 - 0.36(\mathrm{CV1}) + 8.83(\mathrm{CV2})\end{aligned}$$

Moving on to the second part of our analysis, we will examine the variables inside the vectors which are tied by a long-run relationship. We have proven and shown that there exist two co-integrating vectors, but in our case, we will focus only on the first one where BDI is not omitted from the analysis. The second vector, not only omits the BDI but also its coefficient is statistically insignificant so there so incentive to examine the vector. Therefore, we examine the first vector presented in Figure 3 and we can write this relationship (Equation _ce1) as follows (2):

$$\begin{aligned}\mathrm{BDI_t} = {} & 1.86\ \mathrm{MSCI_t} + 11.01\ \mathrm{EURUSDEXCHANGERATE_t} + 2.99\ \mathrm{OECDWORLDTRADEVOLUME_t} \\ & + 3.68\ \mathrm{BulkcarrierNewbuildingPriceInd_t} - 47.79\ \mathrm{TotalTanker10kDWTFleetDev_t} \\ & - 5919.79\end{aligned}$$

CPI_{new} is omitted from the analysis (loading factor 0) but its effect is captured by the constant. We observe that all variables are statistically significant (p-value is 0) except two variables. $BulkcarrierNewbuildingPriceInd_t$ has a p-value of 0.717 and therefore we can safely say that it is statistically insignificant for our analysis. MSCI is the second variable that is only slightly insignificant at the 5% confidence level with a p-value of 0.056 but not at the 10% level. We shouldn’t sorry about this rejection and for the purpose of our analysis we can assume that MSCI is in fact statistically significant.

Moving to the actual interpretation of the coefficients we will examine each variable one by one to see the effect on the dependent variable. Beginning with MSCI, its coefficient is equal to 1.86 which means that a unit increase in the MSCI index will cause BDI to increase as well by 1.86 points towards the same direction indicating a positive relationship. Next, EUR/USD Exchange Rate which is multiplied by 1000, will cause the BDI to move in the same direction, by 11.01 index points, when it increases by one unit (1/1000 regarding the actual data) indicating a positive relationship as well. OECD World Trade Volume has a positive coefficient of 2.99 so an increase in the world trade volume by one billion USD will cause the BDI to move up by 3 points. The Bulk Carrier New-building Price index, has also a positive relationship with the BDI and when it increases by one index point, we can expect the BDI to follow in the same direction by 3.68 points. However, we found that this relationship is statistically insignificant and it won't in fact have an important effect. Finally, the Tanker Fleet Development, working as our measure of supply, has a negative impact on the BDI and an increase by one million deadweight tonnage will cause the BDI to drop by 47.79 index points. The constant is equal to -5919.79.

This relationship, inside the first vector is stationary and provides the information about the adjustment process towards the equilibrium. However, as with all variables, this vector has a coefficient of its own which is denoted by _ce1 and its significance must be examined. From Figure 2, we observe that the coefficient _ce1 which STATA calls L1._ce1 corresponds to -0.36 from equation (1). Looking at the p-value of this coefficient, it is calculated at 0.061 and is only slightly insignificant at the 5% level but not at the 10%. Again, we can assume for the purpose of our analysis that this coefficient is significant and proceed accordingly. This measures, the speed of adjustment towards the equilibrium. Essentially, it is the force that restores the equilibrium relationship and predicts that some variables might decrease in the future while other may increase for the purpose of correcting the disequilibrium error. Another way to interpret this relationship is calculate the time needed for this adjustment. Since _ce1 measures speed, then $\frac{1}{|ce1|}$ will measure time or frequency at which the model will revert back to the equilibrium. This is equal to 2.77 which means that, because of the quarterly frequency of our variables, every 2.77 quarters the model will be at equilibrium, or in other words BDI about every 8 months.

In conclusion, we saw that the BDI has no short-term relationship with either variable but it is affected in the long run. The coefficient of the first co-integrating vector is statistically significant and the BDI is positively affected by three variables and the Bulk Carrier New-building Price index which, however, is statistically insignificant. A negative relationship was found for the Tanker Fleet

Development variable while the CPI was omitted from vector and its effect can be found in the constant. This long-run equilibrium is returned to at a speed of -0.36 or in other words every 8 months.

5.6 Post-Estimation

After the estimation and fitting the VECM and as with any model, we must run some diagnostic tests to examine the stability of our model. The one we are going to employ here is the Lagrange Multiplier (LM) test for autocorrelation in the residuals of the Vector Error-Correction Model. Executing the *veclmar* command in STATA, at a maximum of one lags, we get the following table.

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	51.8688	49	0.36264

H0: no autocorrelation at lag order

At the 5% level we fail to reject the null hypothesis of no autocorrelation in the residuals at lag order 1 and thus there is no evidence of model misspecification.

5.7 Forecasting

We are now ready to make some predictions! We will first generate the impulse responses we discussed in chapter 4. They will help us analyze how the system is affected by dynamic shocks in its endogenous variables. That way, we can track what might happen to the BDI in each case, for many time periods ahead and maybe paint a clearer picture of the causal relationship of it with the other macroeconomic factors.

As we discussed in earlier chapters, the IRFs show the response in one standard error shocks in the equations of the VECM. We already discussed how Sims(1980) described a procedure of orthogonalizing the shocks, using a Cholesky Factor decomposition of the covariance matrix of the basic VECM formula and how this approach leads to dependable and not unique impulses as shown by Lutkepohl. Pesaran and Shin (1998) proposed a solution to this problem, the Generalized Impulse Responses (GIR), which are the ones we are going to use in our paper as well. Let us examine each one individually.

Figure 3: BDI responding to shocks from all the variables in the model

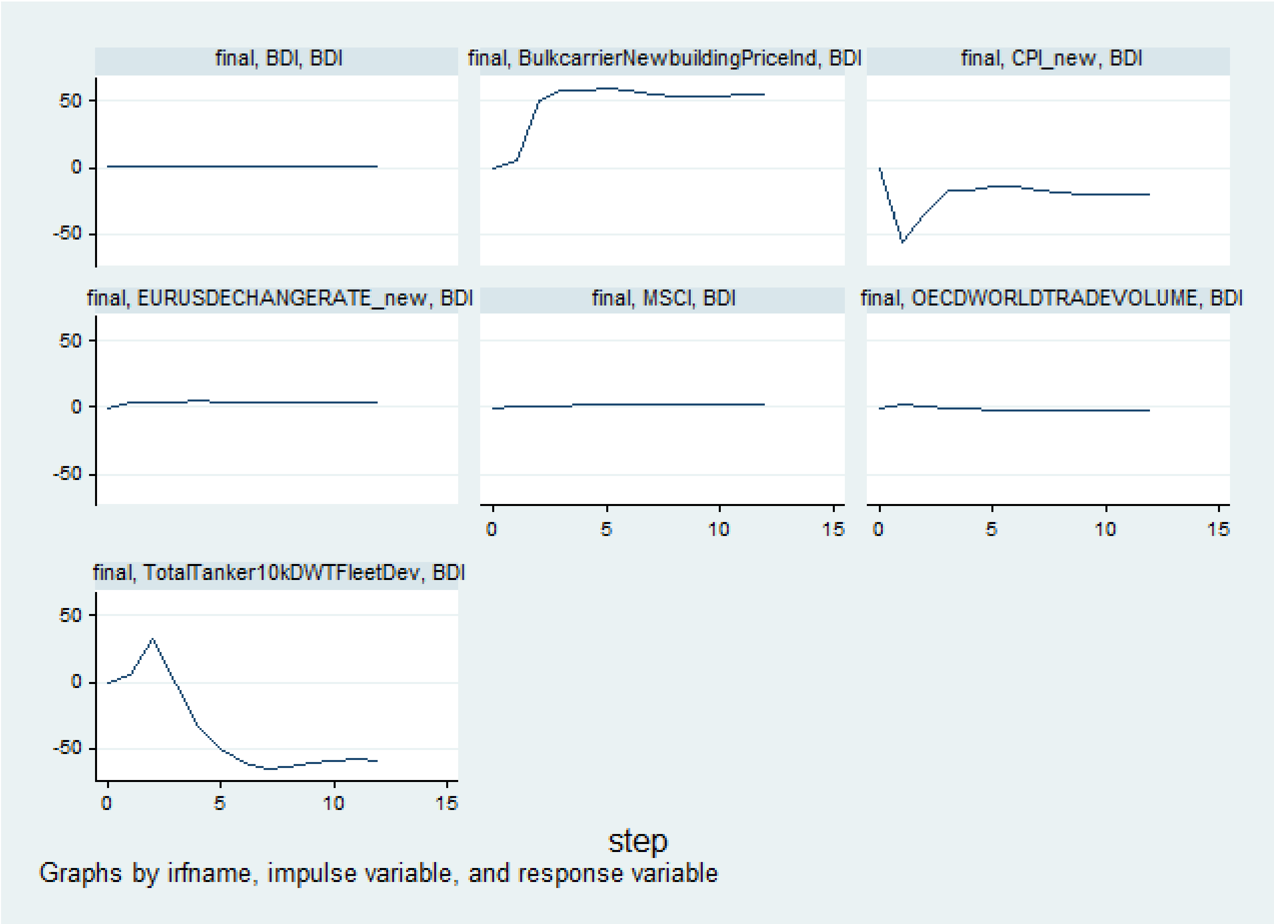
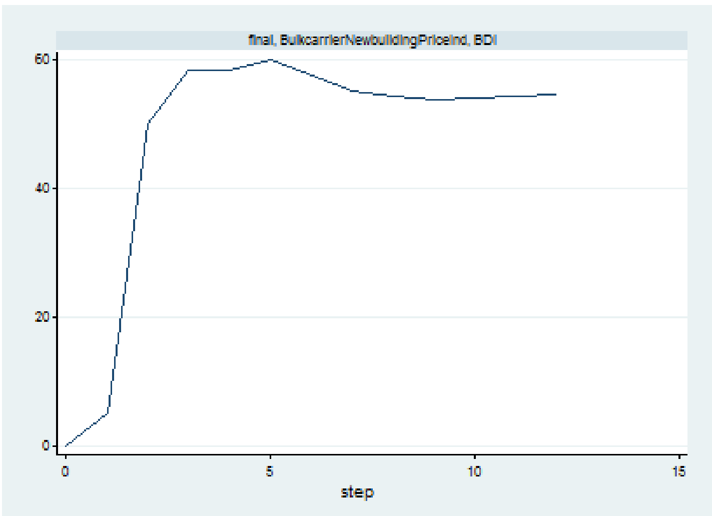


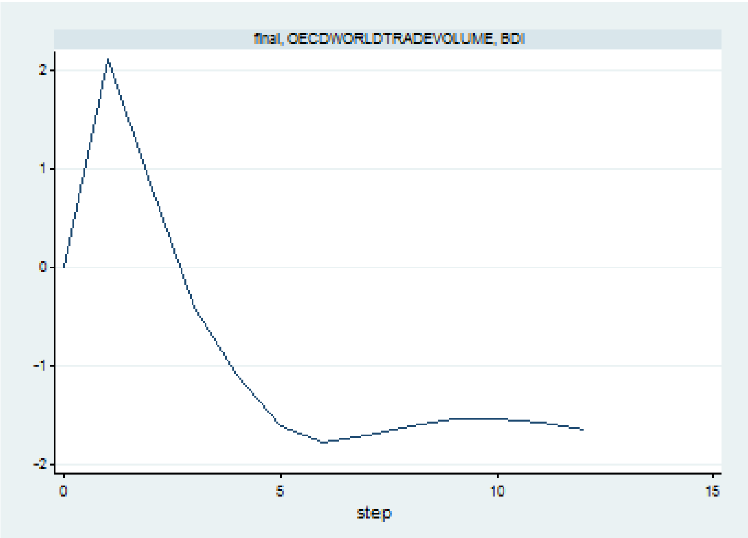
Figure 4: New Building prices of bulk carriers shocks to the BDI



If shipyards require a higher price for new building bulk carriers, the BDI will probably follow an upward trend that will be very steep until it smooths out after almost a year (4 steps = 12 quarters). The BDI may remain in high level even after 2 or 3 years due to this change in prices asked by construction shipyards. This follows the general economic and empirical understanding of the industry shipping cycle:

As new building prices rise and the economy is booming, freight rates tend to go up and remain to a high level and only starting to return to more moderate levels after 2-3 years when the new ships are delivered and the supply catches up to the initial demand that spurred the initial growth. This is shown by the above graph where the impulse generates a response which is sustained for at least 2-3 years, which is exactly the period needed for a ship to be delivered. Another point of interest is that the response is almost immediate in terms of long term analysis. The BDI starts to rise in no less than a month and its curve's slope is huge. This indicates that someone has to be really aware of the changes in prices that shipyards ask and to be well informed in order to be benefitted by the growth of the BDI in a relatively short time frame.

Figure 5: World trade volume shock in the BDI



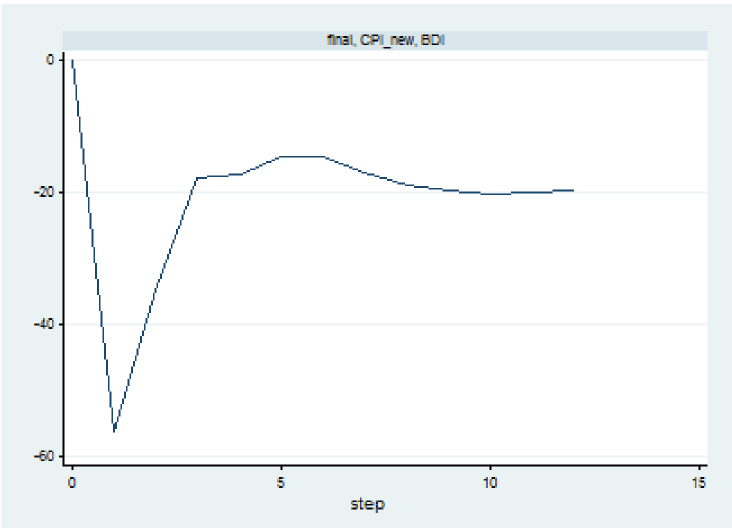
Here, we observe another interesting interaction. We have a one variance shock to the volume of the international trade and how this affects the BDI. There are some similarities with the previous case

(New Building Prices shock). For example an increase in the world trade volume is possible to get followed by an extremely high increase in the BDI as well. After the initial shock, BDI will reach a high value only after 3-4 months and then, it will start declining in almost the same speed. If no other shocks to the economy are present in that period to sustain the high BDI, the index will not only reach its initial value in about the same time (2-3 months), but it will continue to drop immensely and reach very low levels after a year.

All the above are also supported by what we already know about the dynamics of the bulk market: A potential growth will then be followed by a period of oversupply and the booming economy will start declining after new ships enter the market and increase the fleet size. This is why initial growth is often followed by an even bigger decline in the market. The BDI does not simply return to previous values but industry needs more positive shocks to sustain this growth.

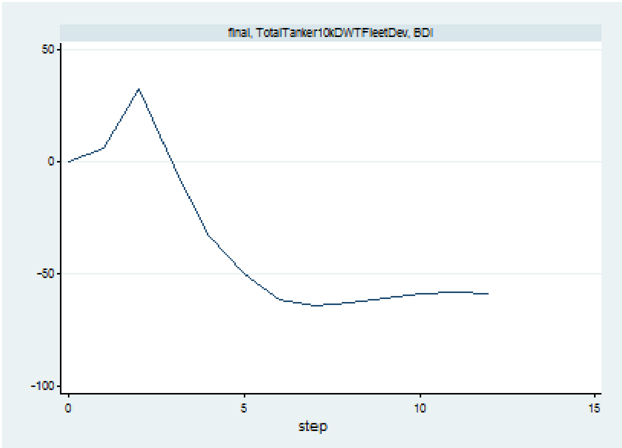
A last thing to note is that after one year the BDI tends to rise a bit due to the initial impact which can be attributed to the fact that after the huge decline in the index's value, a lot of companies decided to scrap their ships, lay off them or even exited the market as a whole. But the BDI cannot restore its previous year values just of that factor alone (a valid hypothesis would be that the most common choice of ship-owners would be the lay-off and/or the selling of the ships and not "immediately" scrap them because of 3-4 unprofitable quarters. So the actual supply does not tend to drop by a lot and is extremely inelastic).

Figure 6: Shocks of CPI to the BDI



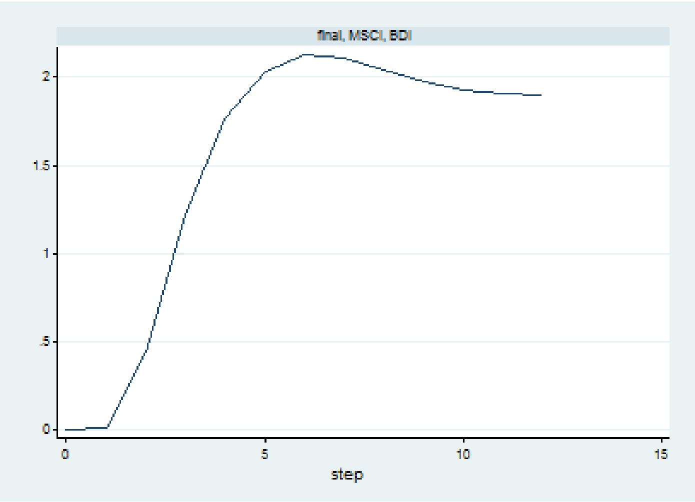
We again, observe a graph that an initial shock to one variable (now the CPI) has a big effect in the BDI for the first few months. In this case, though the relationship is the opposite. Take for example the scenario that the Consumer Price index drops. This will cause the BDI to rise for the first couple of months, which is logical because a low CPI means a higher purchasing power for the consumers, thus, an increase in the demand of goods and as a result an increase in demand for shipping services. The BDI, if no other shocks affect the system will start dropping but in this case, it will return to about 75% of its initial value and stabilize there (with little fluctuations over the next quarters). This does not follow the “trend” we have seen before regarding the ship prices and trading volume impulses which are signs of potential future 180 degrees’ turnaround of the market after a year period. CPI will not affect the BDI in the long run in a big way, but we have to be aware of the first couple of months after a change in the CPI.

Figure 7: Tanker Fleet Development shock to the BDI



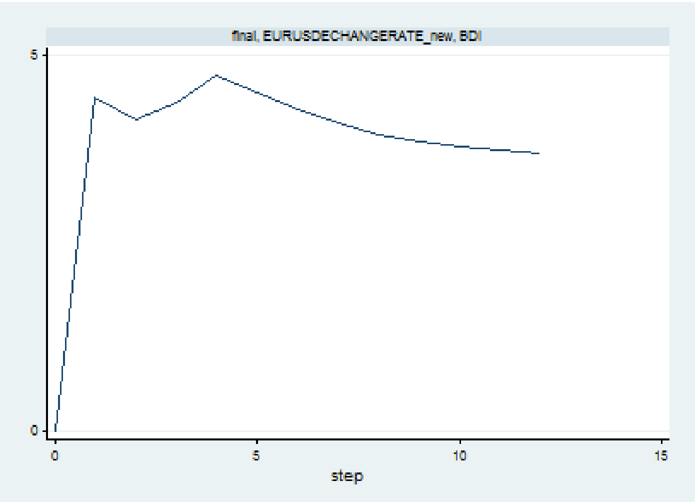
A possible one variance positive shock in the tanker fleet development, will result in minor increase of the BDI. Notice that this increase resembles the changes we saw in most previous impulses, although in a smaller change. In almost 3 quarters BDI will return to where it was and will likely drop quite a lot again resembling the same moderate growth – immense decline (or moderate decline- immense growth) scheme we have discussed before.

Figure 8: Shock of MSCI to the BDI



MSCI is a good benchmark index regarding the conditions of the world economy as we discussed in previous chapters. A positive shock to the MSCI, meaning a booming of the world economy, will result in a positive response of the BDI (and vice versa). This response will be smoother regarding the results we have seen for other impulses and will last longer (the BDI will keep increasing for over a year). Only note the "delay" before the BDI picks up on the MSCI change and starts responding accordingly, which, from the graph seems about a quarter of a year long.

Figure 9: Euro/USD rate shocks to the BDI



Finally, regarding the exchange rate shocks to the BDI we see a familiar picture. Positive shocks will be followed by steep positive responses of the BDI (and vice versa). A positive shock in the rate means a strengthening of the Euro vis-a-vi the dollar and an incline of the BDI as well. One reason might be that, since the world's shipping economy currency is the US dollar, many companies that want to denominate their cash flows to Euro, may need to ask for higher freight rates to offset their losses and keep a steady income flow. This interaction is especially interesting for the European based shipping companies which constantly translate revenue and expenses into their domestic currency and a change in the exchange rate might have a significant effect on their operations.

An interesting thing to note is that after 2 or 3 quarters, disregarding any other shocks, BDI's rise will stall for a while (about a quarter) and then start climbing again for a couple of quarters before starting to smoothly drop in the long term.

5.8 Conclusion

What have we learned from all this? Most changes in the variables result in an "overshooting "in the responses of the BDI. Overshooting suggests that after a response of our variable to an impulse, the return to the equilibrium is not rapid in most cases but the BDI "overshoots" that point. This means, as said before, that a moderate impulse will be followed by an even bigger response from the BDI in most cases, either positive or negative.

This overshooting can be explained along the lines of Zannetos's (1966) "Elastic expectation hypothesis" (Kavussanos, Nomikos 2003). What Zannetos suggested is that there are some behavioral factors to be considered when freight rates change. He stated that an increase in the rates may make ship-owners wait for higher rates, while charterers rush to fix contracts to lock in profit for themselves. This different behavior of these two big players of the market results in greater surpluses than expected which slingshots the BDI way above the equilibrium. A decrease in rates, again results in extreme surpluses because ship-owners and charters tend to behave differently to this change (owners now rush to offer their vessels and services while charterers do not want to lock unprofitable agreements for them), again making the BDI "miss" the equilibrium and reach way worse levels than expected.

5.9 Factor Error Variance Decomposition

In order to supplement to our previous IRF results, we will also perform the Forecast Error Variance Decomposition procedure (FEVD). This method gives us information about the proportion of movements of a variable due to shocks to itself and shocks to other variables for a given horizon (In literature the most common horizon for quarterly data is 12). FEVD offers a slightly different method to examine VECM dynamics. They give the proportion of the movements in the dependent variable that are due to their “own” shocks, versus shocks from other variables. The thinking is that the shock from one variable will affect itself but the shock will be transmitted to the other systemic variables as well.

The table can be analyzed by looking its row individual to assess the participation of each variable to the error variance (rows must sum up to 1 or close to 1 due to rounding up). It can also be analyzed vertically, to examine the effect of each variable over time and how it evolves. To some extent, impulse responses and variance decompositions offer very similar information.

Forecast Error Variance Decomposition Table

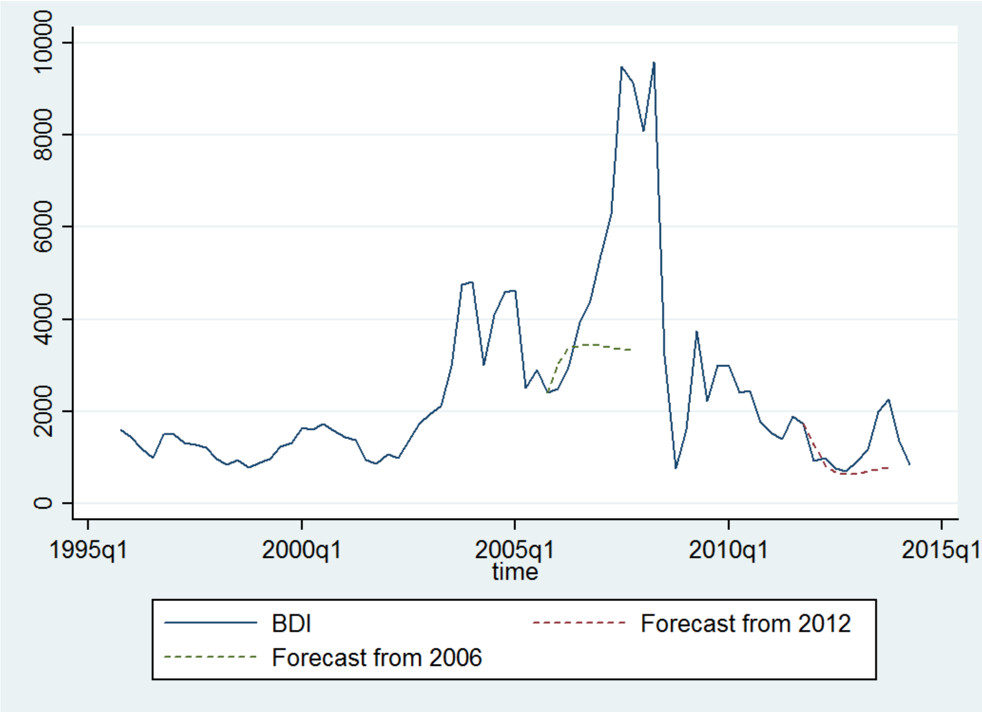
step	(1) fevd	(2) fevd	(3) fevd	(4) fevd	(5) fevd	(6) fevd	(7) fevd
0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0
2	.956987	.011585	.001609	.0258	.003777	.00019	.000053
3	.922015	.01699	.005328	.03494	.006411	.013187	.001129
4	.893549	.017668	.013864	.041074	.006178	.026712	.000954
5	.86281	.018269	.026159	.047203	.00557	.038267	.001722
6	.834275	.018622	.038829	.050712	.005107	.04909	.003366
7	.811042	.018769	.050054	.052425	.004801	.057337	.005571
8	.793294	.018982	.059039	.053154	.004537	.063307	.007687
9	.780265	.019253	.065799	.053204	.00427	.067856	.009354
10	.770594	.019541	.070879	.052903	.00401	.071464	.010607
11	.762924	.019834	.074871	.052476	.003776	.07454	.01158
12	.756383	.02011	.078197	.052009	.003577	.077327	.012397

(1) irfname = final, impulse = BDI, and response = BDI
(2) irfname = final, impulse = CPI_new, and response = BDI
(3) irfname = final, impulse = MSCI, and response = BDI
(4) irfname = final, impulse = EURUSDECHANGERATE_new, and response = BDI
(5) irfname = final, impulse = OECDWORLDTRADEVOLUME, and response = BDI
(6) irfname = final, impulse = BulkcarrierNewbuildingPriceInd, and response = BDI
(7) irfname = final, impulse = TotalTanker10kDWTFleetDev, and response = BDI

We observe that most of the error forecast is explained by the BDI itself (BDI’s effect to itself is dominant in earlier periods, meaning in a short-term horizon). As time goes by, BDI explains less of the forecast error which is subsequently attributed to the other endogenous variables in the model. Each variable’s contribution to the system’s dynamics is ascending overtime with the exception of the world trade volume which after 8 periods (8 quarters = 2 years) starts declining, something very familiar with the results we got from the IRF’s.

Finally, we attempted to examine the ability of our model to predict future values and trends on the BDI by comparing the forecast model to actual historical values. For this reason, we generated two forecasts spanning 8 periods (2 years), one starting from the first quarter of 2006 and the second from the first quarter of 2012, in order to see how well the predictions perform in different market dynamics. Below we see a graph of the 2 forecasts and the "true" historical BDI values. We observe that in both cases, and for a period of up to 3 to 4 quarters, the predictions follow quite closely the actual values of the BDI. In both cases, forecasts can capture the upward (2006) and downward (2011) trend for a short-term horizon (9 months to 1 year). However, in the long term we can observe that there are significant differences between the actual and the predicted values, which indicates that our model has the ability to accurately predict short-term values and performs not so well in the distant future (more than one year).

Figure 10: Predictions against the historical BDI



6. Conclusion

With the introduction of the BDI, the Baltic Exchange gave us the opportunity to measure the cost of transportation through an index which would be internationally recognized as the standard for the cost of transportation regarding dry bulk cargo. As with any other index, the relationship of the BDI with other macro-economic factors was important for the players of the shipping industry as well as individuals and corporations outside that market.

Firstly, we examined the macro-economic factors and the BDI individually and their statistical properties and then with the help of the VECM model we tried to determine if there is a long-term relationship between them. We observed that, there are actually no factors that significantly affect the BDI in the short-run but this was not the case for their long-term relationship. We saw that there is in fact such a relationship and most of the variables have a positive effect to the BDI except for the supply factor, the Total Tanker Fleet Development, where an increase would cause the BDI to drop by a significant margin which translates to about 50 points in the index. All other results agree with the economic theory regarding supply and demand but the importance of this thesis was to try and quantitatively calculate these effects. Finally, we looked into the future and tried to capture the effect of each variable on the BDI. The best predictor of the BDI is obviously its past values, however variables like the MSCI and the EUR/USD Exchange rate are increasingly important especially as we go into the distant future (more than 3 years) where the effect of past values of the BDI on itself is only about 75% and the rest is dividend among the other factors.

7. References

Davidson, R., and J. G. MacKinnon. 1993. Estimation and Inference in Econometrics. New York: Oxford University Press.

Johansen, S. 1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford: Oxford University Press.

Johansen, S. (1991). “Estimation and Hypothesis Testing of Cointegration. Vectors in Gaussian Vector Autoregressive Models”, Econometrica, 59, 1551-1580.

Faust (1976). The influence of exogenous factors on freight rate development.

Beenstock, M. and Vergottis, A. (1989a) An Econometric Model of the World Shipping Market for Dry Cargo, Freight and Shipping. Applied Economics, 21, 339-356.

Brooks C. (2001). Introductory Econometrics for Finance. Cambridge University Press: Cambridge, UK.

Sean Becketti (2013). Introduction to Time Series Using STATA. Stata press publication, College Station, Texas.

Stavros Tsolakis (2005) Econometric Analysis of Bulk Shipping Markets Implications for Investment Strategies and Financial Decision-Making. (Erasmus University Rotterdam, Ph.D. Thesis).

Harris, R. (1995). Cointegration Analysis in Econometric Modelling, Prentice-Hall, Harlow, UK.

Kavussanos M. G. and Visvikis, I.D. (2002). Market Interactions in Returns and Volatilities between Spot and Forward Shipping Markets. International Association of Maritime Economists (IAME). Annual Conference, Panama November 2002.

Drakos K. Analysis of Time Series DATA: Non-stationarity, Volatility, VARs and Cointegration. (University of Essex, Summer School in Social Science Data Analysis and Collection).

Thien, Le Duc, "Forecasting the dry bulk freight market" (2005). World Maritime University Dissertations. Paper 242.

Petter Hvalgard Bakke & Daniel Reinsborg (2012). Empirical impacts of macroeconomic risk factors on tanker shipping equities. (Copenhagen Business School, Master Thesis).

Kavussanos, M.G. & Nomikos, N.K. Review of Derivatives Research (2003) pp 203-230.

J.G.M Oomen (2012). The Baltic Dry Index: A predictor of stock market returns. (Tilburg University, Department of Finance, Master Thesis).

Stopford, M. 2003. Maritime economics. Routledge, New York

Shun Chen (2011). Modeling and Forecasting in the Dry Bulk Shipping Market. (Shanghai Maritime University, P.H.D thesis)

Alizadeh and Talley, 2010. Microeconomic determinants of dry bulk shipping freight rates and contract times. Transportation, 38 (2010), pp. 561–579

E. Bildirici , Fazıl Kayıkçı, Işıl Şahin Onat (2015) BDI as a major economic policy indicator.

Geman and Smith, 2012. Shipping Markets and Freight Rates: An Analysis of the Baltic Dry Index. The Journal of Alternative Investments, 1 (2012), pp. 98–109

Faqin Lin and Nicholas Sim (2013). Trade, income and the Baltic Dry Index. European Economic Review

Apergis and Payne (2013). New Evidence on the Information and Predictive Content of the Baltic Dry Index .
International Journal of Financial Studies

Han, Liyan, Wan, Li, Xu, Yang (2020). Can the Baltic Dry Index predict foreign exchange rates?. Finance Research Letters.

Yuliang Zhang, Gang Zhao (2020). Analysis of Baltic Dry Index based on GARCH Model. International Journal of Social Science and Education Research.

Jason Angelopoulos (2017). Time–frequency analysis of the Baltic Dry Index. Maritime Economics and Logistics.

Kamal, Bae, Sunghyun, Yun (2020),DERN: Deep Ensemble Learning Model for Short- and Long-Term Prediction of Baltic Dry Index

Lan Kuo, Chien-Liang Chiu, Chan-Sheng Chen, Mei-Chih Wang, (2020). Dynamic Relationships Between The Baltic Dry Index And The BRICS Stock Markets: A Wavelet Analysis

EXTRA REFERENCES

Websites

Lloyd’s List http://www.lloydslist.com/

World Bank http://www.worldbank.org/

International Monetary Fund http://www.imf.org

The Baltic Exchange http://www.balticexchange.com/

Investopedia www.investopedia.com

8. Appendix

	p~Tank~v	p~Bulk~v	pr_BDI	pr_VIX	pr_EUE~Y	pr_USE~Y	pr_USF~E	pr_EUR~E	pr_OEC~E	pr_Bre~l	pr_Bul~e	pr_Oil~s	pr_Tan~x
pr_TotalTa~v	1												
pr_TotalBu~v	0.3727	1											
pr_BDI	0.194	-0.2619	1										
pr_VIX	-0.0158	0.0025	-0.017	1									
pr_EUECONO~Y	-0.004	-0.0491	0.018	0.5504	1								
pr_USECONO~Y	0.0127	0.0261	-0.1001	0.5117	0.6466	1							
pr_USFEDFU~E	0.0054	0.1941	0.1229	0.3397	0.0907	0.0362	1						
pr_EURUSDE~E	0.1034	-0.0055	0.1586	-0.0809	0.12	0.063	-0.0754	1					
pr_OECDWOR~E	-0.1305	-0.0018	-0.0529	0.0377	0.0369	-0.0265	0.0464	0.0682	1				
pr_BrentOi~l	0.0869	-0.0077	0.2419	-0.0335	0.0467	-0.0332	-0.0347	0.1574	0.5231	1			
pr_Bulkcar~e	-0.0984	-0.0661	0.1185	-0.0414	0.0177	-0.062	-0.0653	0.1069	0.4819	0.3451	1		
pr_OilTank~s	-0.0881	-0.0803	0.087	0.0458	0.072	0.0015	0.1004	0.0369	0.323	0.1847	0.756	1	
pr_TankerS~x	-0.0694	-0.105	0.0065	0.0394	0.1026	0.0312	0.0906	-0.0513	0.5158	0.3729	0.6191	0.5862	1
pr_BulkCar~d	0.0919	-0.2037	0.4928	-0.1143	0.0521	-0.092	-0.0104	0.0537	0.3693	0.4387	0.7233	0.4754	0.5535
pr_BulkerS~P	0.0558	-0.1058	0.3763	-0.1219	-0.0086	-0.037	-0.0085	0.0935	0.2114	0.4358	0.3532	0.2188	0.3797
pr_TankerS~t	0.0465	-0.1131	0.3822	-0.1307	-0.0486	-0.0659	-0.0188	0.058	0.2523	0.4088	0.3891	0.2225	0.324
pr_MSCI	0.1587	0.0114	0.1532	-0.6545	-0.4287	-0.4803	-0.1103	0.2048	0.3272	0.1786	0.1386	-0.0785	0.0718
pr_CPI	-0.1777	-0.2016	-0.1256	0.0259	0.1707	0.0182	-0.0412	-0.1853	0.2296	0.2339	0.0672	0.0475	0.4123

	pr_Bul~d	pr_Bul~P	pr_Tan~t	pr_MSCI	pr_CPI
pr_BulkCar~d	1				
pr_BulkerS~P	0.5505	1			
pr_TankerS~t	0.5546	0.9377	1		
pr_MSCI	0.2653	0.1821	0.26	1	
pr_CPI	0.1651	0.3375	0.2833	0.0852	1

Figure 1: Correlation Matrix

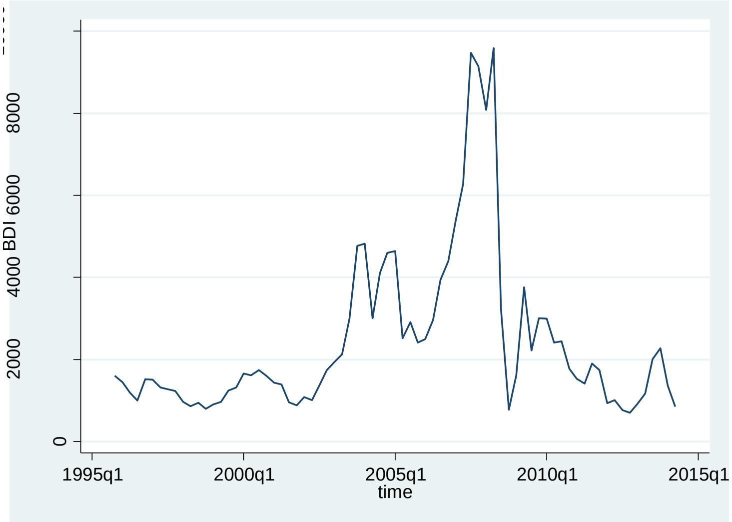


Figure 2: BDI time series chart

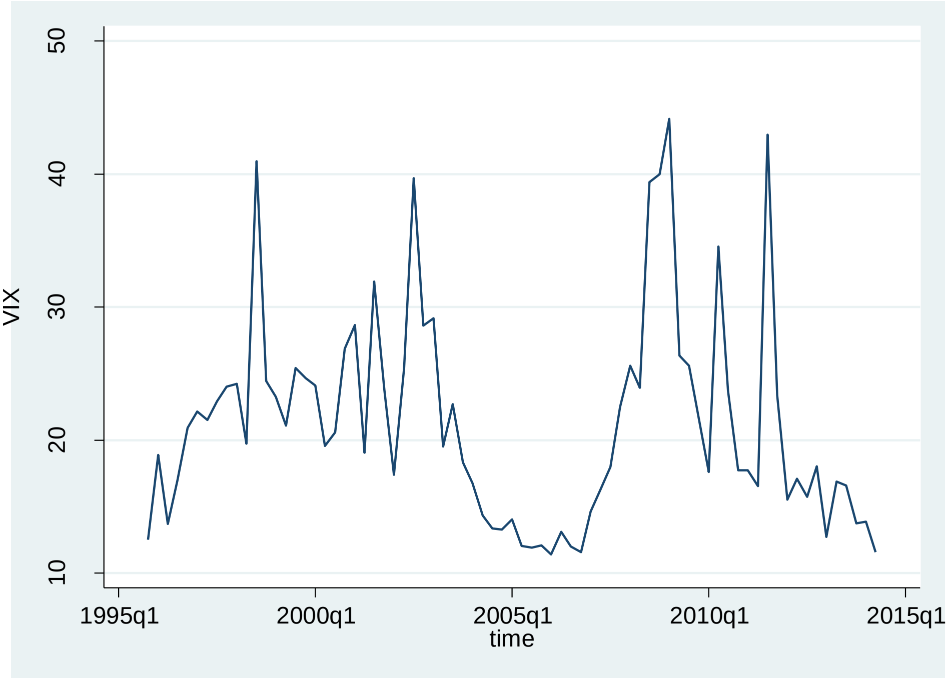


Figure 3: VIX time series chart

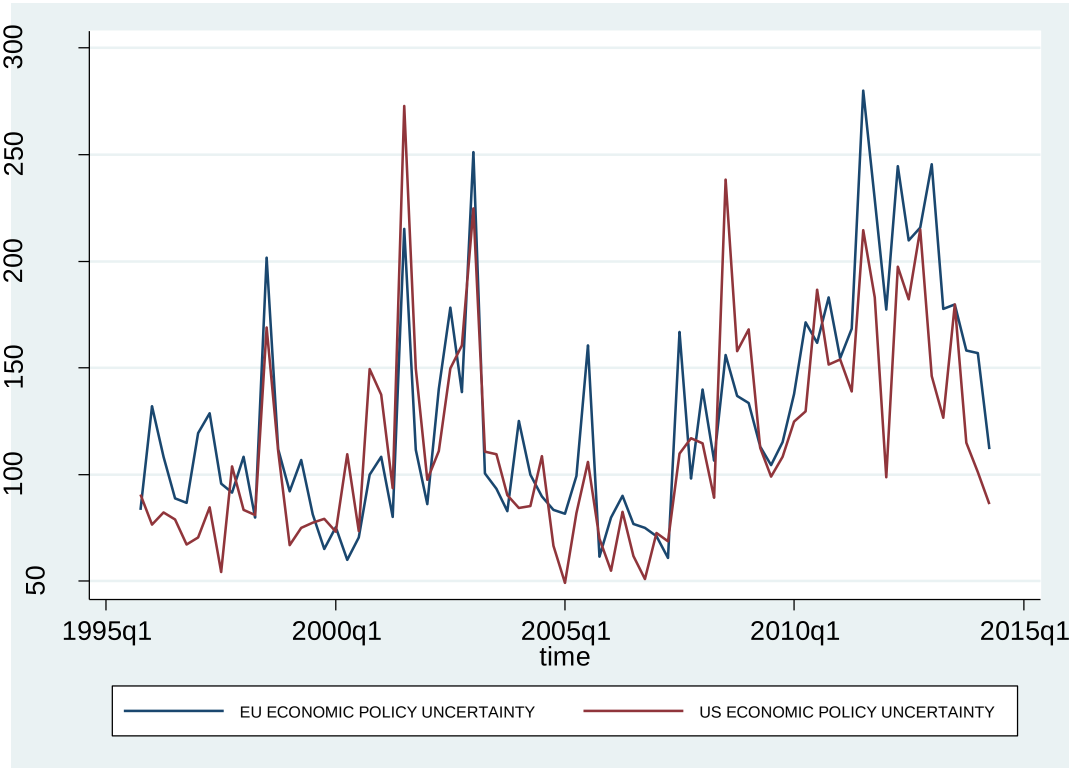


Figure 4: Economic Policy Uncertainty time series chart

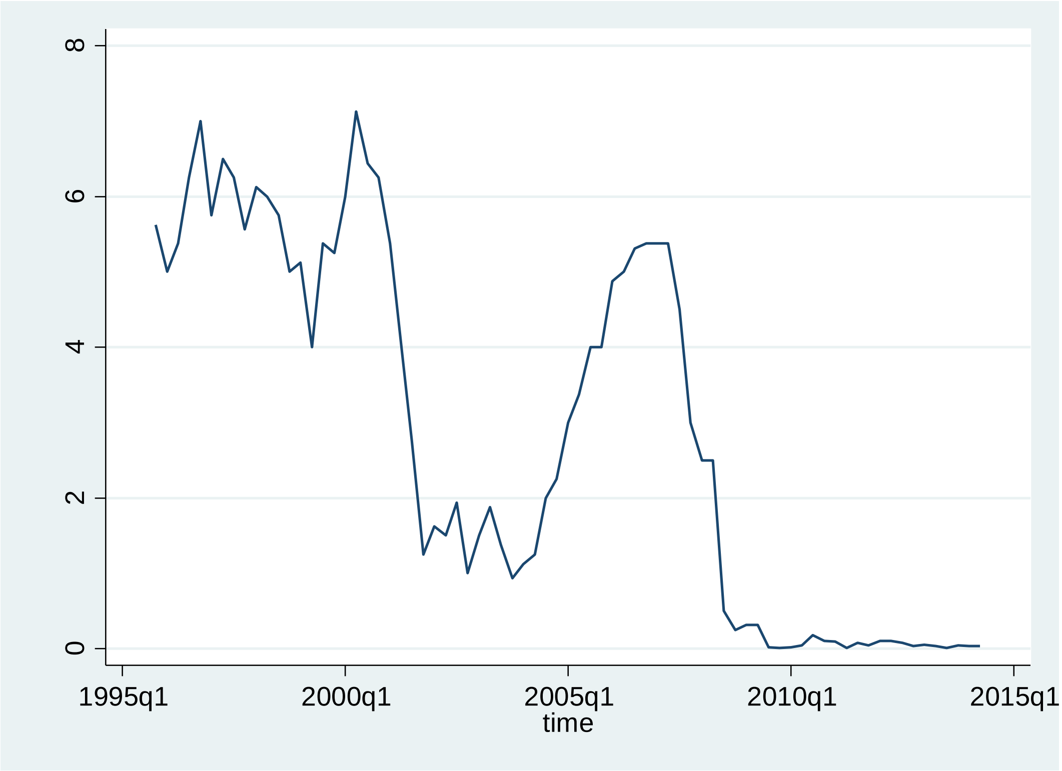


Figure 5: US federal funds rate time series chart

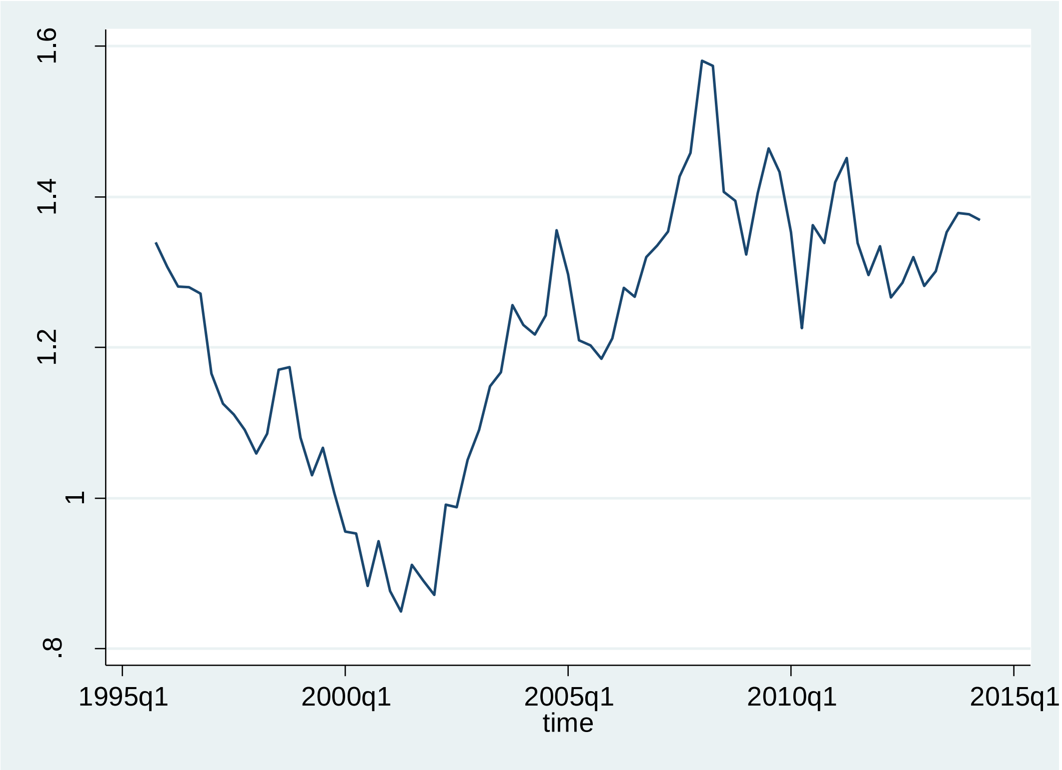


Figure 6: Euro/USD exchange rate time series chart

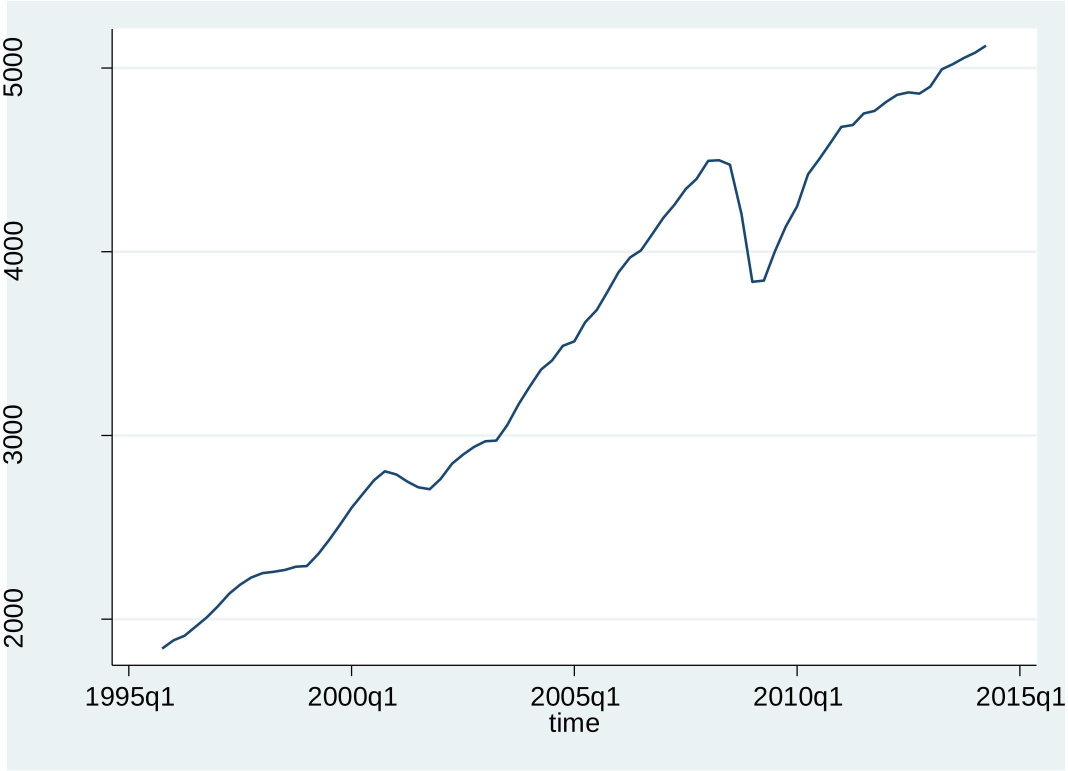


Figure 7: OECD World trade volume time series chart

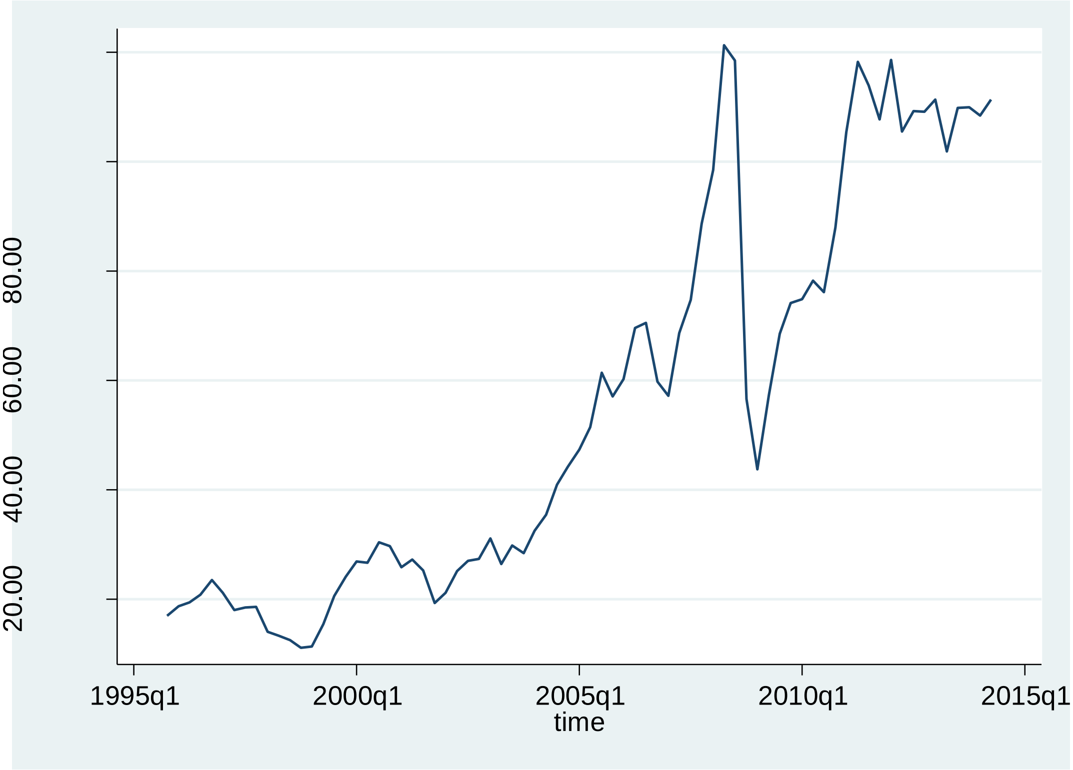


Figure 8: Brent oil price (bbl) time series chart

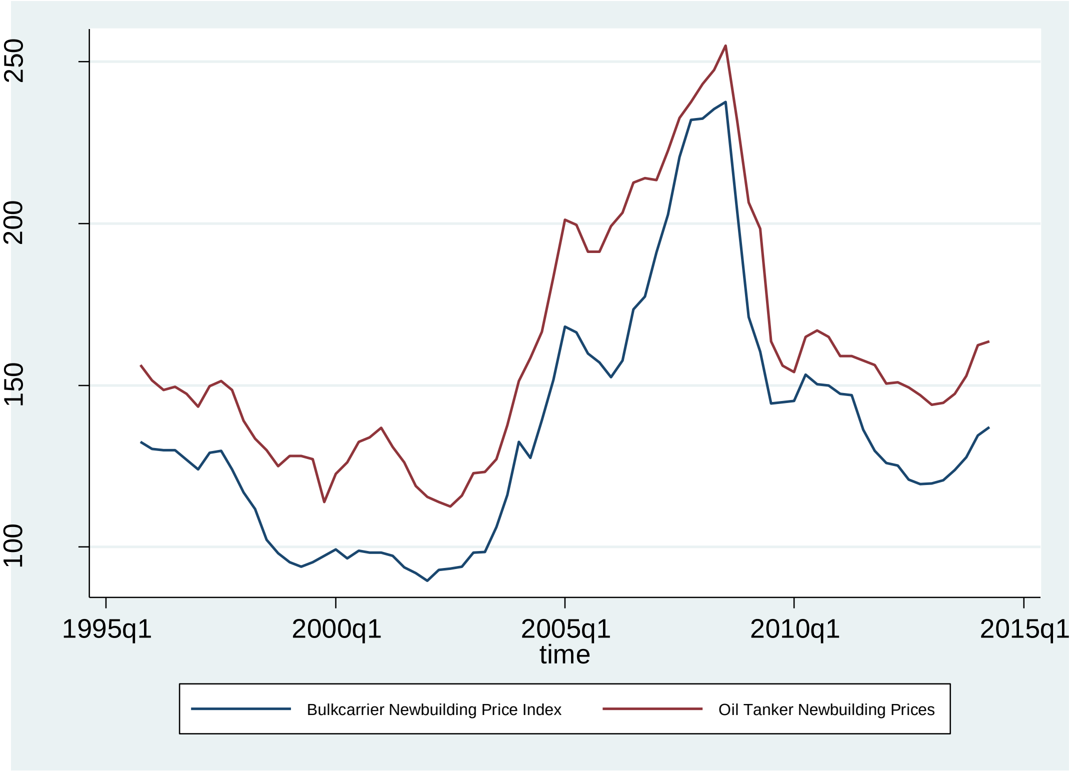


Figure 9: New building Price index time series chart

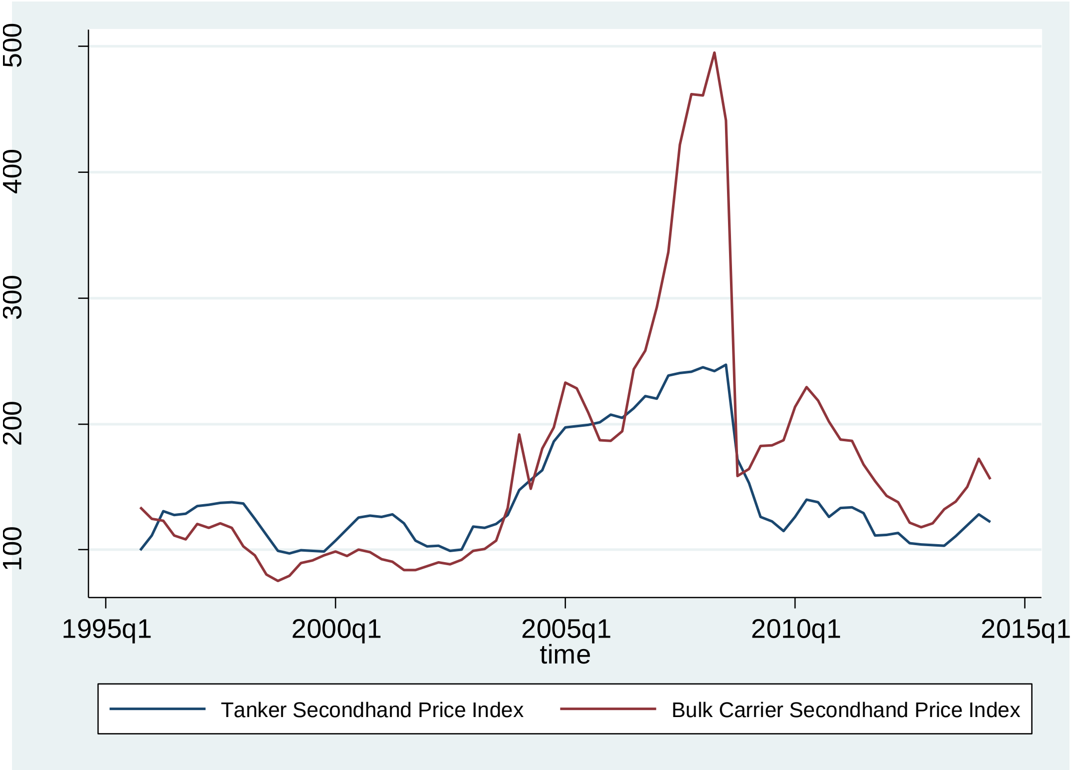


Figure 10: Second-hand Price index time series chart

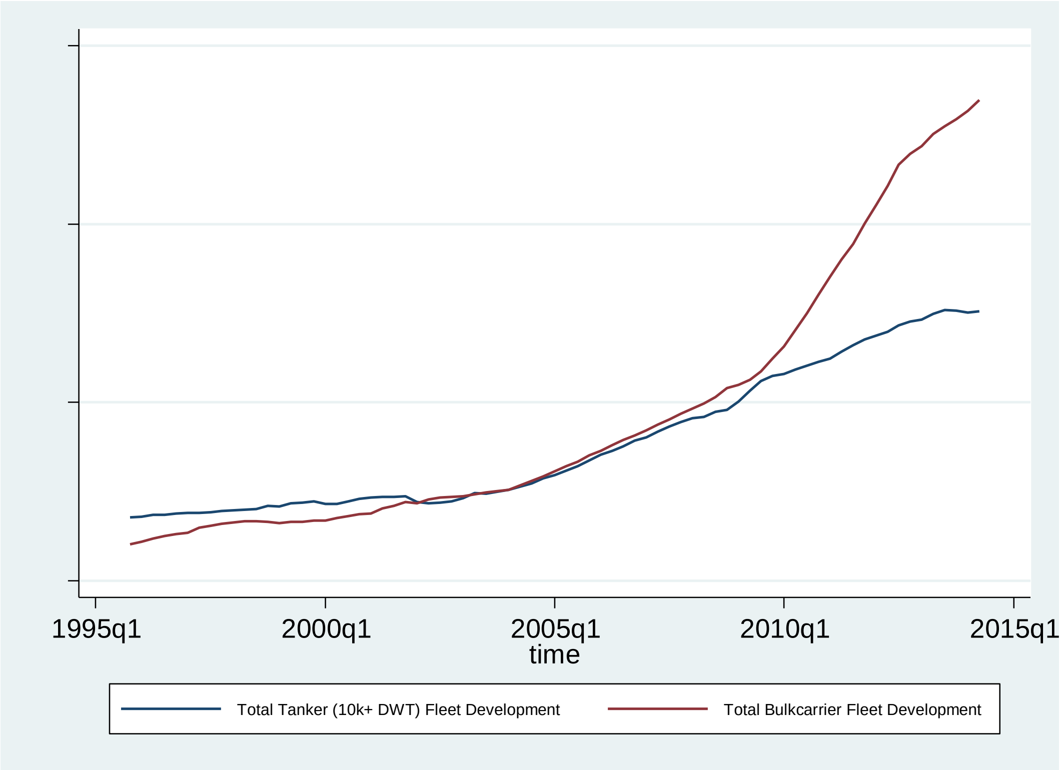


Figure 11: Total Fleet Development (Dry Bulk & Tanker) time series chart

Figure 12: Scrap Price (Dry bulk & Tanker) time series chart

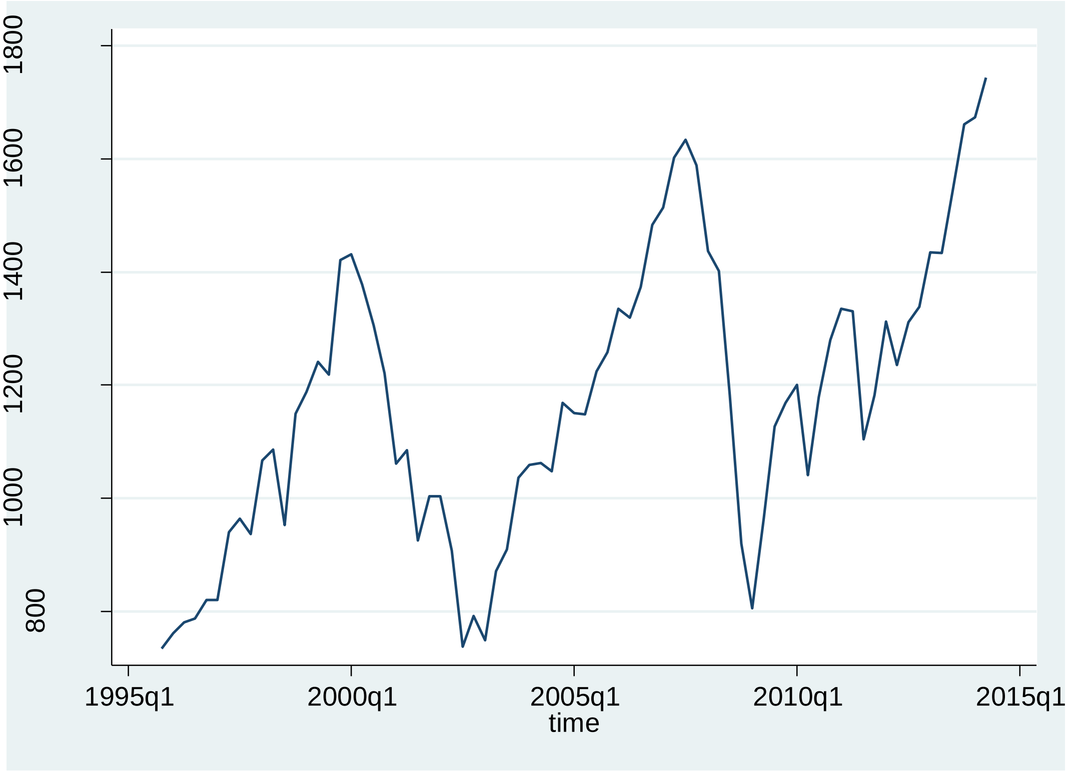


Figure 13: MSCI time series chart

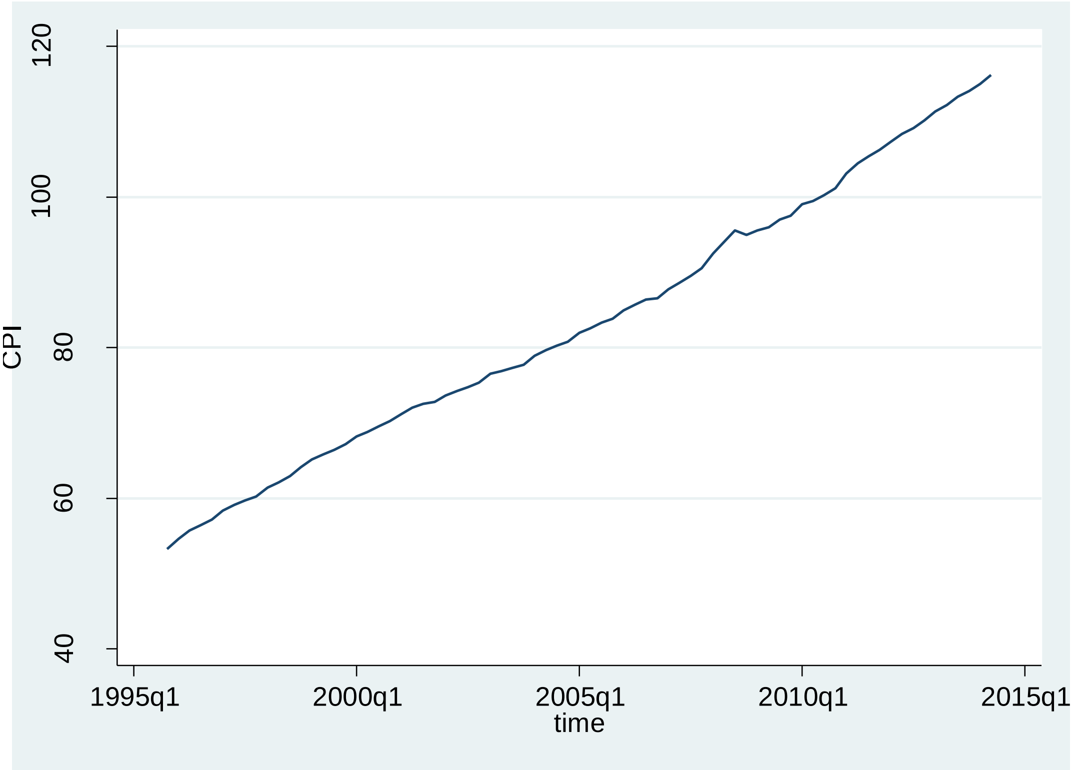


Figure 14: CPI time series chart

Vector error-correction model

Sample: 1996q2 - 2014q2		No. of obs =	73
		AIC =	65.74891
Log likelihood = -2319.835		HQIC =	66.74922
Det(Sigma_ml) = 9.44e+18		SBIC =	68.259

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_BDI	10	1079.15	0.1892	14.46658	0.1528
D_CPI_new	10	3.11441	0.9026	574.6093	0.0000
D_MSCI	10	92.8867	0.2594	21.71886	0.0166
D_EURUSDECHANG~w	10	61.4295	0.1151	8.067874	0.6222
D_OECDWORLDTRA~E	10	42.6451	0.7906	234.1398	0.0000
D_BulkcarrierN~d	10	5.5242	0.6470	113.6351	0.0000
D_TotalTanker1~v	10	1.96364	0.8292	301.0828	0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D_BDI						
_ce1						
L1.	-.3677713	.1962794	-1.87	0.061	-.7524718	.0169292
_ce2						
L1.	8.833176	6.996291	1.26	0.207	-4.879302	22.54565
BDI						
LD.	.1971739	.1749535	1.13	0.260	-.1457286	.5400763
CPI_new						
LD.	-64.99102	37.63839	-1.73	0.084	-138.7609	8.778864
MSCI						
LD.	-.5130351	1.542911	-0.33	0.740	-3.537086	2.511016
EURUSDECHANGERATE_new						
LD.	2.871193	2.514817	1.14	0.254	-2.057758	7.800144
OECDWORLDTRADEVOLUME						
LD.	2.894477	2.408613	1.20	0.229	-1.826319	7.615272
BulkcarrierNewbuildingPriceInd						
LD.	-7.692333	22.92641	-0.34	0.737	-52.62727	37.2426
TotalTanker10kDWTFleetDev						
LD.	20.07977	49.08843	0.41	0.683	-76.13177	116.2913
_cons	1.842721	460.7099	0.00	0.997	-901.132	904.8174

D_CPI_new						
_ce1						
L1.	.001803	.0005665	3.18	0.001	.0006927	.0029132
_ce2						
L1.	-.0970316	.0201913	-4.81	0.000	-.1366058	-.0574574
BDI						
LD.	.0001044	.0005049	0.21	0.836	-.0008852	.0010941
CPI_new						
LD.	-.0736198	.1086244	-0.68	0.498	-.2865197	.1392801
MSCI						
LD.	.0059129	.0044528	1.33	0.184	-.0028145	.0146403
EURUSDECHANGERATE_new						
LD.	.0013016	.0072578	0.18	0.858	-.0129234	.0155265
OECDWORLDTRADEVOLUME						
LD.	.0022242	.0069513	0.32	0.749	-.0114	.0158484
BulkcarrierNewbuildingPriceInd						
LD.	-.2509111	.0661656	-3.79	0.000	-.3805934	-.1212289
TotalTanker10kDWTFleetDev						
LD.	-.2058315	.1416692	-1.45	0.146	-.483498	.071835
_cons	10.38547	1.329609	7.81	0.000	7.779483	12.99145
D_MSCI						
_ce1						
L1.	-.0302229	.0168946	-1.79	0.074	-.0633357	.00289
_ce2						
L1.	.9145724	.6022009	1.52	0.129	-.2657196	2.094864
BDI						
LD.	.0469722	.015059	3.12	0.002	.0174571	.0764873
CPI_new						
LD.	-3.773722	3.239698	-1.16	0.244	-10.12341	2.575969
MSCI						
LD.	-.0303891	.132805	-0.23	0.819	-.2906822	.2299039
EURUSDECHANGERATE_new						
LD.	-.1412845	.2164611	-0.65	0.514	-.5655405	.2829716
OECDWORLDTRADEVOLUME						
LD.	.3338005	.2073197	1.61	0.107	-.0725387	.7401397
BulkcarrierNewbuildingPriceInd						
LD.	-.5550553	1.973375	-0.28	0.779	-4.422799	3.312688
TotalTanker10kDWTFleetDev						
LD.	6.063012	4.225252	1.43	0.151	-2.21833	14.34435
_cons	-13.88485	39.65528	-0.35	0.726	-91.60777	63.83807

D_EURUSDECHANGERATE_new						
_ce1						
L1.	.0071357	.011173	0.64	0.523	-.0147631	.0290345
_ce2						
L1.	.2207082	.3982584	0.55	0.579	-.5598639	1.00128
BDI						
LD.	-.0031514	.0099591	-0.32	0.752	-.0226708	.0163681
CPI_new						
LD.	-2.6649	2.142536	-1.24	0.214	-6.864192	1.534393
MSCI						
LD.	-.0346617	.087829	-0.39	0.693	-.2068034	.13748
EURUSDECHANGERATE_new						
LD.	.0639245	.143154	0.45	0.655	-.2166522	.3445012
OECDWORLDTRADEVOLUME						
LD.	-.0026262	.1371084	-0.02	0.985	-.2713537	.2661014
BulkcarrierNewbuildingPriceInd						
LD.	1.693037	1.305068	1.30	0.195	-.8648489	4.250923
TotalTanker10kDWTFleetDev						
LD.	4.81284	2.79432	1.72	0.085	-.6639271	10.28961
_cons	22.00082	26.22555	0.84	0.402	-29.40031	73.40195
D_OECDWORLDTRADEVOLUME						
_ce1						
L1.	-.007084	.0077565	-0.91	0.361	-.0222864	.0081183
_ce2						
L1.	.6461183	.2764755	2.34	0.019	.1042362	1.188
BDI						
LD.	.0358174	.0069137	5.18	0.000	.0222668	.0493681
CPI_new						
LD.	1.76487	1.487373	1.19	0.235	-1.150327	4.680067
MSCI						
LD.	.1711708	.0609719	2.81	0.005	.0516681	.2906736
EURUSDECHANGERATE_new						
LD.	-.0967919	.0993792	-0.97	0.330	-.2915714	.0979877
OECDWORLDTRADEVOLUME						
LD.	.5904843	.0951822	6.20	0.000	.4039305	.7770381
BulkcarrierNewbuildingPriceInd						
LD.	.197091	.9059931	0.22	0.828	-1.578623	1.972805
TotalTanker10kDWTFleetDev						
LD.	4.213732	1.939849	2.17	0.030	.4116974	8.015766
_cons	-11.15677	18.20608	-0.61	0.540	-46.84002	24.52649

D_BulkcarrierNewbuildingPriceInd						
_ce1						
L1.	.0011822	.0010048	1.18	0.239	-.0007871	.0031515
_ce2						
L1.	-.0890291	.0358144	-2.49	0.013	-.159224	-.0188342
BDI						
LD.	.0023903	.0008956	2.67	0.008	.000635	.0041456
CPI_new						
LD.	-.1825412	.1926729	-0.95	0.343	-.5601732	.1950907
MSCI						
LD.	.0149749	.0078982	1.90	0.058	-.0005054	.0304551
EURUSDECHANGERATE_new						
LD.	.0096361	.0128735	0.75	0.454	-.0155955	.0348677
OECDWORLDTRADEVOLUME						
LD.	.0192875	.0123298	1.56	0.118	-.0048785	.0434535
BulkcarrierNewbuildingPriceInd						
LD.	.1094248	.1173615	0.93	0.351	-.1205995	.3394492
TotalTanker10kDWTFleetDev						
LD.	-.5376099	.2512863	-2.14	0.032	-1.030122	-.0450979
_cons	2.420047	2.358398	1.03	0.305	-2.202328	7.042423
D_TotalTanker10kDWTFleetDev						
_ce1						
L1.	-.0012372	.0003572	-3.46	0.001	-.0019372	-.0005371
_ce2						
L1.	.0065504	.0127306	0.51	0.607	-.0184012	.031502
BDI						
LD.	.0013221	.0003184	4.15	0.000	.0006981	.0019461
CPI_new						
LD.	.0712945	.0684879	1.04	0.298	-.0629393	.2055282
MSCI						
LD.	-.0105587	.0028075	-3.76	0.000	-.0160613	-.005056
EURUSDECHANGERATE_new						
LD.	-.0039533	.004576	-0.86	0.388	-.0129221	.0050156
OECDWORLDTRADEVOLUME						
LD.	-.0022103	.0043828	-0.50	0.614	-.0108004	.0063798
BulkcarrierNewbuildingPriceInd						
LD.	-.0217907	.0417175	-0.52	0.601	-.1035556	.0599741
TotalTanker10kDWTFleetDev						
LD.	.5036181	.0893227	5.64	0.000	.3285489	.6786873
_cons	-.3621239	.8383205	-0.43	0.666	-2.005202	1.280954

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	5	163.4089	0.0000
_ce2	5	2322.169	0.0000

Identification: beta is exactly identified

Johansen normalization restrictions imposed

beta	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ce1						
BDI	1	.	.	.	.	.
CPI_new	0	(omitted)				
MSCI	-1.85389	.9683462	-1.91	0.056	-3.751813	.044034
EURUSDECHANGERATE_new	-11.01584	2.267128	-4.86	0.000	-15.45933	-6.57235
OECDWORLDTRADEVOLUME	-2.997094	.6416054	-4.67	0.000	-4.254617	-1.73957
BulkcarrierNewbuildingPriceInd	-3.687074	10.18151	-0.36	0.717	-23.64247	16.26833
TotalTanker10kDWTFleetDev	47.79742	8.485434	5.63	0.000	31.16628	64.42856
_cons	5919.798	.	.	.	.	.
_ce2						
BDI	0	(omitted)				
CPI_new	1	.	.	.	.	.
MSCI	-.0182552	.0231923	-0.79	0.431	-.0637113	.0272009
EURUSDECHANGERATE_new	-.2822453	.0542987	-5.20	0.000	-.3886688	-.1758218
OECDWORLDTRADEVOLUME	-.2137206	.0153667	-13.91	0.000	-.2438389	-.1836024
BulkcarrierNewbuildingPriceInd	1.31741	.2438517	5.40	0.000	.8394693	1.79535
TotalTanker10kDWTFleetDev	.4163099	.2032298	2.05	0.041	.0179868	.8146331
_cons	-59.90071	.	.	.	.	.

Figure 15: VECM Results