Erasmus University Rotterdam
Erasmus School of Economics

Master Thesis - MSc Economics and Business Economics - Urban, Port and Transport Economics

Price Dynamics in the Dry Bulk Market: Investigating elements of individual spot-rate contracts through cross-sectional time-series analysis

Name student: Wouter Dirk Jan Peters
Student ID number: 604367

Supervisor: Dr. A.W. (Albert) Veenstra
Second supervisor: M. (Maurice) Jansen MSc

Date: 06-08-2023

Abstract

Over the past decade, the dry bulk shipping market has experienced significant growth, raising the question of whether data usage has kept pace with this expansion. Charter contracts record crucial information for transporting dry bulk, encompassing vessel and bulk characteristics, laycan period details, freight rates, and load and discharge destinations. However, it remains unclear what insights can be obtained from these contract elements. For that reason, a cross-sectional time-series analysis is conducted to investigate which elements of the individual spot-rate contracts in the dry bulk shipping market can be used to conduct a market analysis on the price dynamics.

Previous research focused on industry pricing patterns by creating timelines and using autoregressive models, but frequently missed out on investigating individual contracts and their components. For that reason, this research concentrates on specific spot-rate contract elements and investigates their impact on price trends on an individual contract basis. This research specifically investigates (laycan) time elements, geographical elements, freight elements and fleet availability and their effect on the price dynamics.

In this research, spot-rate contracts for Capesize vessels from 2020 to 2022 are used. The dataset consists of 3716 contracts based on $/Tonne freight rates and contains 609 distinctive Capesize ships. In order to determine the price dynamics, the Baltic Dry Index (BDI) is used as an economic indicator and linked to every contract date in the dataset. The analysis is done with Fixed Effects models, Random Effects models and a Random Forest model to determine the effect of different contract elements on price dynamics.

The analysis reveals that in order to conduct a market analysis on price dynamics, various components such as time elements, geographical elements and freight elements of individual spot-rate contracts for Capesize ships are insightful to investigate the dry bulk market. The individual spot-rate contracts are also used in this research to outline the overall fleet availability for Capesize ships and provided also relevant insights on the price dynamics. In conclusion, the relevance of many individual factors in influencing price dynamics in the dry bulk shipping industry is clarified by this research's findings. Understanding the interactions between these components provides valuable insights into pricing trends for Capesize ships in the dry bulk sector.

Keywords: Price Dynamics, Dry Bulk Market, Cross-sectional time-series, analysis, Capesize, Fixed Effects, Random Effects, Random Forests, Spot-rate contracts

Table of Contents

List of Tables

List of figures

Appendix tables

1. Introduction

Over the last centuries, it has become normal for humans to trade diverse items from all over the globe. The fact that shipping has been the primary mode of conveyance for goods for almost as long as people have been exchanging them shows the industry's ongoing significance to the global economy (Tan et al., 2019). The dry bulk industry is crucial to the worldwide flow of raw materials, however, it is also vulnerable to boom-and-bust cycles, speculative investments, and periods of irrational optimism (German, 2005). The demand for bulk commodities, consisting of petroleum and dry bulk, has more than tripled since the 1970s and reached 6.4 billion tons in 2018. Dry bulks consist of three major components: iron ore, coal, and grain, which account for 29.2 percent of global seaborne trade (UNCTAD, 2019). The maritime industry has emerged as the backbone of global trade in a manufacturing environment where trade distances are narrowing. The bulk cargo sector is particularly important for the global exchange of raw materials (Limao and Venables, 2001).

The shipping industry is characterized by a healthy industry of market analysts, data providers and brokers (Jacks & Stuermer, 2021). In dry bulk shipping, contracts are closed which involves the negotiation and agreements between the shipper and the carrier regarding the transportation of bulk commodities. It contains important elements such as the freight rate that was agreed upon, the quantity and type of cargo that would be moved, the designated loading and discharge ports, and the delivery schedule (Stopford, 2009). Over the past decade, the dry bulk shipping market has grown tremendously, but has the usage of data grown sufficiently with it? All the data for transporting dry bulk from one location to another is recorded in charter contracts. These contracts contain various elements such as vessel characteristics, characteristics of the bulk, information on the laycan period, the freight rate, the start and end destination However, it is not exactly evident what information could be obtained from these elements. For that reason, a cross-sectional time-series analysis of these elements of spot-rate contracts will be conducted to investigate the dry bulk shipping market and its effects on price dynamics. This research is based on spot-rate contracts data from Capesize ships, but before diving deeper into the topic, it is necessary to formulate the existing problem in dry bulk shipping.

1.1. Problem formulation

The dry bulk freight rates are generating more interest from a wide range of stakeholders in the shipping business, which necessitates the use of various forecasting approaches. Several academics have developed a few forecasting models throughout the years for the dry bulk shipping business, including simple and complex forecasting models, with varying degrees of success in terms of effectiveness (Alizadeh & Talley, 2011).

One of the most important challenges in the shipping business is the analysis of the spot and forward shipping costs in the freight market, including voyage rates, time-charter rates, and freight forward pricing. Spot and period pricing in the freight market has long drawn attention in the shipping business, because it may be useful to predict short-term price trends and minimize risks by revealing the internal shifting rules and pattern of fluctuation of these prices. Beenstock and Vergottis (1993) have turned their focus to the stationarity characteristics of data and modelling price dynamics in terms of its volatility (Chen, 2011). This resulted in the usage of modern econometric models in order to examine time-varying behaviour in price dynamics.

The most known modern econometric models in bulk shipping to investigate the price dynamics are the time-series forecast models. These models are a type of extrapolation that include fitting historical data to a model and then applying that model outside of that data fitting with the hope that subsequent data will be comparable to the historical data (Chatfield, 2001). Makridakis et al. (2018) discovered that while most statistical time-series models, like ARIMA, were built based on linear assumptions, they suffer from giving inaccurate forecast values for the series featuring nonlinear patterns. In this context, a number of time-series models, such as the threshold autoregressive (TAR) model and the autoregressive conditional heteroscedastic (ARCH) model, were developed to model nonlinear patterns in the data (Fischer et al., 2018). It was discovered that these models could only outperform a select few unique nonlinear situations. For generic nonlinear situations, they were unable to provide a good enough predicting performance. Engle (1982) developed the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, which successfully modelled the conditional volatility.

It is common to investigate the price dynamics for dry bulk by investigating historical freight rate data using time-series analysis. Although, for these time-series analyses mostly the modern econometric models such as the ARIMA, TAR, ARCH and GARCH models are used. While these methods of investigation consider the price dynamics within the industry, they do not look at specific contracts and specific elements of these contracts. These methods lump all contracts together and create a timeline and do not investigate individual contracts and the individual elements within the contracts. As a result, there is no substantive review of elements of the individual contracts, which potentially could lead to price dynamics modelling with incomplete information. For that reason, this research investigates price dynamics in a different way, by doing a cross-sectional time-series analysis by investigating specific elements in individual contracts of Capesize ships and examining which elements influence price dynamics. This research is answered by research questions, however before diving into the research questions, some background of the shipping industry is given.

1.2. Shipping industry background

The shipping market can be divided into two main sub-markets, namely the liner and bulk-shipping sectors. The shipping market can be divided into numerous divisions: bulk carriers, container vessels, oil tankers, LNG carriers, LPG carriers, chemical tankers, and cruise vessels (Fevre, 2018). Liner shipping is characterized by boats running along pre-specified, fixed routes on a regular, fixed schedule, while bulk shipping is distinct in that ships in this industry often operate when the charterer requests them. Since the freight rates in the liner sector are typically predetermined, there is typically very little room for bargaining between parties (Kang & Woo, 2017). On the other hand, bulk shipping is distinct in that the ships in this industry often operate when the charterer requests them. Additionally, the bulk industry's freight charges are very variable and subject to agreement between the ship-owner; as a result, most agreements are reached in private.

According to Stopford (2009), bulk shipping is most commonly characterized as the transportation of uniform bulk cargoes by bulk boats on an erratic timetable. Dry bulk carriers are substantial single-hull ships that specialize in the transportation of 'dry' commodities and transport massive amounts of bulk raw materials around the world as the primary sector of the shipping industry (Stopford, 2009). They are divided into big bulks and minor bulks, with the majority of big dry bulk shipments being divided by weight. The dry bulk shipping industry is a crucial component of the supply chains for the aluminium, metallurgical, and agri-food industries. Therefore, bulk trade will continue to play a significant role in global economic processes. There is a strong demand for bulk logistics to become more competitive as a result of the continual growth in the number of dry bulk shipments (Comtoi & Lacoste, 2021; Kanamoto et al., 2021).

The dry bulk shipping industry is characterised by individual contract data gathered and disseminated by charter brokers. The information in the shipping industry is in two ways distributed: in raw form and the form of unit value indices. The raw form can be distributed as monthly and even weekly average prices, depending on ship sizes and commodities on particular routes. This information is derived from fixtures, which are concise summaries of the transactions that were agreed upon between the charterers. This fixture data is transformed by brokers or other publishing companies into average freight rates, price indices, and other market indicators, such as the number of fixtures in a given trade or at a given period (Veenstra & van Dalen, 2008).

The Baltic Freight Index (BFI) and the Baltic Dry Index (BDI) are two closely related indexes used by the shipping industry to measure and evaluate the costs of transporting dry bulk commodities across various routes. The BFI is determined by analyzing the average daily profits of Panamax boats on international trade routes, accounting for variables such as vessel size, distance travelled, and market demand. The BDI generates the average charter prices for four vessel classes, Capesize, Panamax, and

Supramax/Handysize vessels, which are based on the size in deadweight tonnage (DWT), along various international shipping routes. The BFI is a key economic indicator for assessing the health of the dry bulk shipping sector and global trade activity. The Baltic Dry Index is a useful instrument for evaluating overall performance and trends in the dry bulk shipping sector, as it covers a range of vessel sizes and commodities (Cullinane et al., 1999; German & Smith, 2012).

The bulk shipping business offers charterers flexibility in meeting their needs for sea transportation, and shipowners can make use of their fleet by offering sea transportation services. The market can be divided into three distinct types of transactions: time charter, voyage contract, and spot-rate contract. The first is a time charter, in which a ship is chartered on a day-to-day basis. The second is a voyage contract, whereby the shipper purchases transportation from the shipowner at a set cost per ton of goods. The third is a spot rate contract, which is a short-term contract at the prevailing market rate for single-voyage transportation of freight for a particular shipment between two ports (Stopford, 2009). The length of the contract, method of calculating the freight rate, cost distributions, and commercial and operational obligations are the primary distinctions between these contracts, all for different vessel sizes (Koekebakker et al, 2007; Alizadeh & Nomikos, 2009; Stopford, 2009; Alizadeh & Talley, 2011). In summary, there is a lot of information within these contracts and for that reason, the following sub-chapter dives into the examination of the spot-rate contracts of Capesize ships. These are investigated with various sub-question in order to answer the research question.

1.3. Research questions

This research adopts a unique method for time-series analysis by using a cross-sectional time-series analysis that concentrates on examining various parts of individual Capesize contracts. This research seeks to find hidden patterns and dynamics that conventional time-series approaches can miss by examining individual components in each contract. For that reason, the following research question needs to be answered: ***Which elements of individual spot-rate contracts in the dry bulk market can be used to conduct a market analysis on price dynamics?***

The research is carried out step by step, therefore the research question is answered by the various sub-questions. It is necessary to investigate different elements of individual contracts and therefore every sub-question investigates a certain element.

The first sub-question investigates the element of time. Dobre (2016) states that the laycan period, which is the time allocated for loading and unloading the vessel, is negotiated depending on the previous route and the coming route and has a major impact on the freight rate. Alizadeh and Talley (2011) conclude that the laycan period is a crucial element in freight rate analysis as they are related in a simultaneous matter. Kavussanos & Alizadeh (2001) investigate the seasonal behavior of dry bulk freight prices for a range of ship sizes and freight contracts with a range of durations and conclude that there are predictable seasonality patterns. In other words, many elements of time are investigated previously, which indicates the importance of time elements in contracts.

However, many researchers investigate a specific time element such as the laycan period and not all time-related elements of contracts. For that reason, the first sub-question is: *Can insights into time elements of contracts provide information about price dynamics in dry bulk freight rates?* This investigates which time elements of contracts have an effect and if so, what these effects are on the price dynamics in dry bulk freight rates.

Throughout the past few decades, international marine transportation has increased, playing a significant role in the expansion of global trade (Alizadeh & Talley, 2011). The dry bulk cargo shipping market is a significant part of the global shipping market and is acknowledged as a submarket with high risk and volatility. Elements such as the volume and pattern of international trade, the state of the world economy, and governmental policy influence the dry bulk cargo shipping routes (Jing et al., 2008). As Shibasaki et al. (2017) state, the competition among shipping routes intensifies as maritime shipping becomes more globalised, suggesting the geographical distribution of shipping activities also becomes more important. Alizadeh & Talley (2011) conclude with their research that the voyage route is an important determinant in the dry bulk shipping freight rate. The more recent research of Guan et al. (2019) investigated the effects of geographical distribution relative to an economic performance parameter, the Baltic Supramax Index and concluded that geography is an important element. Therefore, the geographical distribution of shipping activities on price dynamics is investigated as the next element. This results in the second sub-question: *What are the effects of the geographical elements of shipping activities on the price dynamics?* This sub-question examines the effects of major shipping routes and the overall effects of geographical distribution on price dynamics.

As stated earlier, dry bulk is divided into major bulks and minor bulks. Wada et al. (2018) evaluated the effects of vessel size and newly built bulk carriers focusing on iron ore, coals and grains and used a cargo prediction model to model the shipbuilding market and found that vessel size affects the type of cargo. Comtoi & Lacoste (2021) state that vessels with different commodities are handled in a restricted number of terminals, which results that 75% of iron ore being carried by Capesize ships. Adland et al. (2016) conducted a study intending to determine the variables that affect how a ship's cargo-carrying capacity is utilized during distinct iron ore-related voyages. According to empirical studies, general market conditions are mostly influenced by vessel-specific characteristics (DWT), with smaller boats frequently having lower capacity utilization (Adland et al., 2016).

It appears from the literature that there is a connection between different dry bulk commodities and different vessel sizes. Consequently, it might be interesting to examine how freight elements influence the price dynamics in the market. This is investigated with the following sub-question: *How do freight elements of dry bulk influence the price dynamics?*. This sub-question investigates if elements of the freight such as the quantity, the type of bulk and the rate influence the price dynamics.

Forecasting is considered a very important feature of the shipping industry since it helps shipowners and other parties involved, such as lenders and shipyards, in the industry make more profitable choices about charter agreements, the purchase of vessels, and other business-related decisions (Stopford, 2009). In many industries throughout the decennia, time series forecasting techniques have been used to attempt and predict the future. Chatfield (2001) states that there are three types of forecasting methods: judgmental forecasts, univariate time series forecasts and multivariate time series forecasts. With the judgmental forecast, predictions are made under the influence of judgment which is founded on any kind of information. Univariate time series forecasts enhance a time function and only take into account past and current values. For multivariate time series forecasting, predictions are based on values of one or more explanatory factors from additional time series (Chatfield, 2001).

Univariate and multivariate time series forecasts are considered relatively commonly used methods to make predictions about dry bulk shipping. However, this research focuses on judgemental forecasting in order to look at current and expiring contracts and place them next to the ship list, in order to get an overall picture of fleet availability. This makes it also possible to see whether fleet availability has positive or negative effects on price movements in the markets. For that reason, the last sub-question is: *What are the effects of fleet availability on price dynamics in the market?*

Ultimately, after answering the four sub-questions, it is possible to answer the research question appropriately. The aim of the research is to provide insights into what elements and how these elements of individual Capesize contracts provide information on the price dynamics in dry bulk freight rates for Capesize ships.

1.4. Content of the research

In order to provide insights into these elements in individual contracts and provide information on price dynamics, the context of the shipping market must be explained first. The following chapter dives into the context of the literature, which is divided into two sub-chapters: the shipping market and the dry bulk market. The third chapter consists of the literature review, which is divided into chartering, freight rates and the freight market. Then, the methodology chapter discusses how the data is handled and how the insights are obtained. This is handled in three parts, first the literature of the methodology, then the chosen method per sub-question and finally the mathematical background of the models.

The fifth chapter contains the data, which consists of three sub-chapters. The first one describes how the data is obtained and what is in the data and what is not. The second one dives into the data preparation and the manipulation of the data. The third and last one presents the final dataset. Eventually, the result chapter presents all the results obtained through the discussed methods with the available data, which is divided into sub-chapters per sub-question. Finally, there is a concluding section which summarizes the research by answering all sub-questions and eventually it answers the research question. At last, the limitation section describes some limitations of the research. As the content of the research is clear, it is necessary to start with the literature context.

2. Literature context

In the following chapter, the theoretical context for this research is established. The context consists of two sub-chapters, which are divided into smaller sub-sections. The first sub-chapter is about the shipping market. First, the shipping market is introduced and then the sub-chapter is divided into three sub-sections. The first sub-section discusses the different types of dry bulk ships. The second sub-section explains the ship sale and purchase market. The third at last section gives an overview of the total fleet size. The second sub-chapter is about dry bulk, which is also divided into three sub-sections. First, an overview of the distribution of the dry bulk is given, then the different types are discussed and finally, the different dry bulk routes are mapped out. Ultimately, this should provide a clear picture of the context of shipping and can then be followed by a review of the substantive literature that the study of various elements of the contracts and the impact on price dynamics requires.

2.1. Shipping Market

As stated in the introduction, the shipping market can be divided into numerous divisions: bulk carriers, container vessels, oil tankers, LNG carriers, LPG carriers, chemical tankers, and cruise vessels (Fevre, 2018). By the end of 2010, dry bulk carriers accounted for 35.5% of the global fleet's tonnage and 9% of all ships. The 7,300 ships in the world's dry bulk fleet have a combined tonnage of 459.3 million tons (UNCTAD, 2010). Dry bulk cargo ships have carrying capacities ranging from 10,000 deadweight tonnes to 400,000 deadweight tonnes. The number of ships on the market and their cargo-carrying capacity, expressed in deadweight tonnes (DWT), determines the dry bulk fleet's overall capacity (USDA, 2020). According to Stopford (2009), the DWT is the most weight of cargo that a ship can carry before being deemed overloaded. This measurement includes all fuel, supplies, water ballast, freshwater, passengers, and baggage. Depending on their size, these vessels typically have a different number of holds and hatches, ranging from one to nine and other diverse characteristics (Plomaritou & Papadopoulos, 2017).

The economies of scale led to the growth and specialization of dry bulk carriers. The demand for bulk commodities: oil, gas, iron ore, coal and grain reached 6.4 billion tons in 2018, which accounted for 29% of the overall world seaborne trade (UNCTAD, 2019). According to Drewry Shipping Consultants Ltd., there were 12,312 dry bulk vessels active in the fleet as of December 2020 with a 912.2 million deadweight tonnes cargo capacity (USDA, 2020).

From this, it can be concluded that the total fleet capacity has almost doubled in the last decade, indicating the relevance of the shipping market. First, the different types of vessels will be reviewed, then the sale and purchase market will be examined. After that, the built and demolition market will be covered and finally, the total fleet size is discussed to summarize this chapter.

2.1.1. Types of dry bulk ships

As stated, deadweight tonnes (DWT) describes the weight of cargo, fuel, water and other supplies that the vessel can carry. The DWT is frequently used to describe the size of dry bulk carriers and according to their DWT ranges, they are divided into major groups. The following categorization is frequently used: Handysize, Handymax/Supramax, Panamax, Post-Panamax, Capesize and Vloc and the DWT range from 10,000 to over 400,000. As this research investigates spot-rate contracts of Capesize ships, only this type of dry bulk ship is discussed in the main text and the other vessel sizes are briefly summarized. More information about the other vessel sizes can be found in Appendix A.1.

The smallest vessels are the Handysize vessels, which carry between 10,000 and 40,000 DWT. Handymax/Supramax vessels can carry between 40,000 and 65,000 DWT. The Handysize and Handymax vessels are considered somewhat smaller and are often categorized as Handyclass vessels. Panamax vessels have a DWT capacity between 65,000 and 85,000, while Post-Panamax vessels range from 85,000 and 100,000 DWT. As the names suggest, the Panamax and Post-Panamax vessels are defined by certain conditions to transit through the Panama Canal (Plomaritou & Papadopoulos, 2017). Over, 2017; USDA, 2020).

As the industry is evolving, the size of the vessels is evolving too. The conventional definition of Capesize bulk carriers in terms of deadweight tonnage has changed from 80,000 to above 100,000 (Darie et al., 2013). Since they are too large and too wide to travel via the Panama Canal, they must travel through the Cape of Good Hope. Due to the fact that these vessels have to sail past Cape of Good Hope, the name Capesize was born. These ships are divided into Small Capes that can carry from 100,000 to 130,000 DWT, Normal Capes that can carry from 130,000 to 200,000 DWT, and Large Capes, also known as Very Large Bulk/Ore Carriers which can carry over 200,000 DWT and can carry up to 400.000 DWT (Plomaritou & Papadopoulos, 2017). The Capesize industry concentrates on the long-distance trading routes for coal and iron ore and as a result, 75% of iron ore is carried by Capesize ships (Comtoi & Lacoste, 2021). Only a relatively small number of ports worldwide have the facilities to handle Capesize vessels because of their size. For that reason, the large Capesize vessels often serve the route Brazil-China and therefore they are also known as Chinamax or Valemax vessels (Over, 2017).

To summarize, DWT is used to describe the size of the carriers and according to that it is divided into six major groups ranging from 10,000 to 400,000+ DWT. Now that it is clear which major categories of vessels there are, what they transport and where they operate globally, it is necessary to look at the sale and purchase market of the vessels.

2.1.2. The ship sale and purchase market of dry bulk vessels

The maritime industry's sale and purchase market of dry bulk vessels entails deals in which shipowners, operators, or investors buy or sell dry bulk carriers for a variety of reasons, including fleet development, fleet replacement, or investment goals. Three parties make up the sale and buy market: the new building market, the used market, and the scrapping market, which are further separated into several types of vessels (Stopford, 2009; Over, 2017).

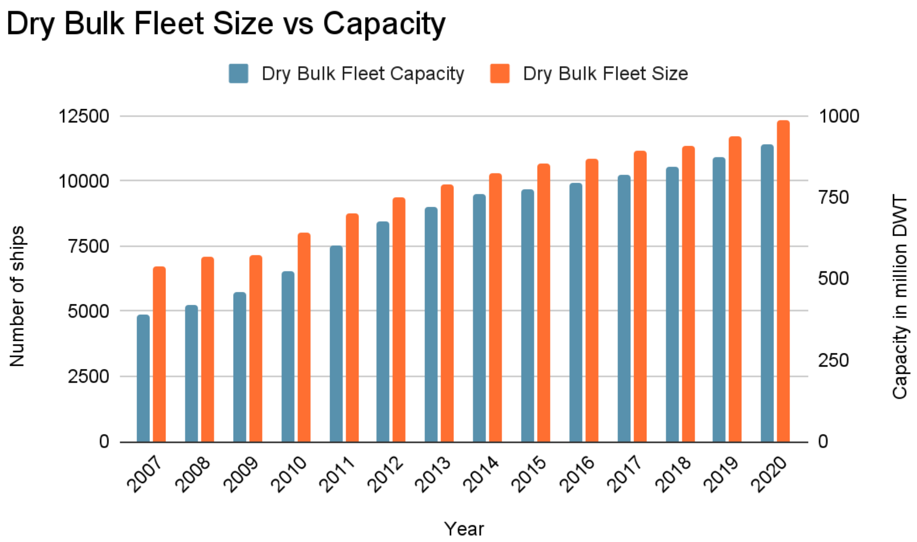
Various variables might impact the sale and purchase market of dry bulk vessels, such as supply and demand patterns, freight costs, market attitude and world economic situations. Alizadeh & Nomikos (2003) found evidence that price fluctuations can be used to anticipate trading volume, which implies that bigger capital gains stimulate more market activity. Additionally, they contend that increased trading activity lowers market volatility since volume has a negative impact on the volatility of price changes. The unusual foundations of the shipping market, such as thin trading, which posits that an increase in trading activity results in price transparency and stability, may be used to explain this (Alizadeh & Nomikos, 2003; Lun & Quaddus, 2009). In sum, the sale and purchase market consists of three different parties, the new building market, the second-hand market and the scrapping market. The sales and purchase conditions are impacted by various factors such as the freight rate, the prices of new ships, the investors' landscape, the second-hand market, the scrap market, and other world economic factors. Additional information on the new building, second-hand and scrapping market can be found in Appendix A.2.

2.1.3. The total size of the dry bulk fleet

The dry bulk fleet size consisted of over 12,000 ships with a total capacity of over 912 million tons in December 2020. The efficiency of the world fleet, as well as shipbuilding and shipwrecking activities, all affect its capability. The supply is increased by adding new ships to the fleet while it is decreased by the retirement or scrapping of older ships. Fleet performance is affected by vessel breakdown frequency, vessel operating speed, and vessel traffic congestion at busy ports (Stopford, 2009; USDA, 2020).

To assess the overall size of the fleet, it is useful to investigate the trend of the number of ships in recent years. If a ship is added to the fleet and no ship will be demolished, the total fleet capacity will increase. For that reason, the fleet capacity will also be taken into account. The following figure will provide a timeline of the dry bulk fleet size and capacity:

Figure 1. Dry bulk fleet size vs fleet capacity



Source: USDA, 2020

As can be seen from the figure, fleet size and fleet capacity have the same positive trend. It can be seen that the number of dry bulk ships has consistently increased annually from 2007 till 2020 and that the capacity consistently increased too. It is notable, that the capacity has doubled since 2007, while the number of ships has increased by one-third, suggesting that the new ships are on average likely to be bigger. As a next step, the distribution of the dry bulk fleet of 2020 is presented:

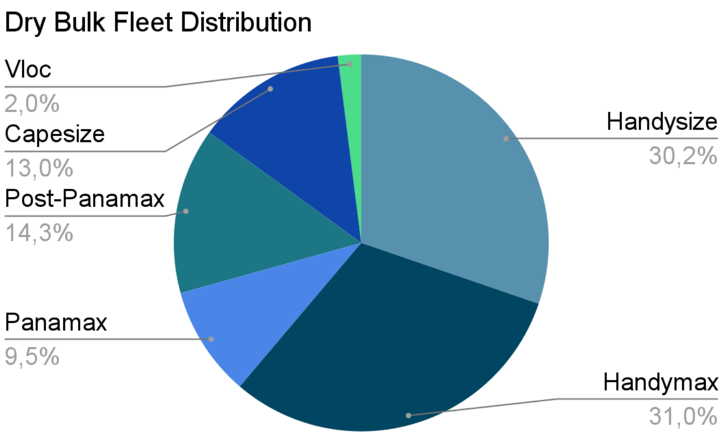
Table 1. Dry bulk fleet distribution

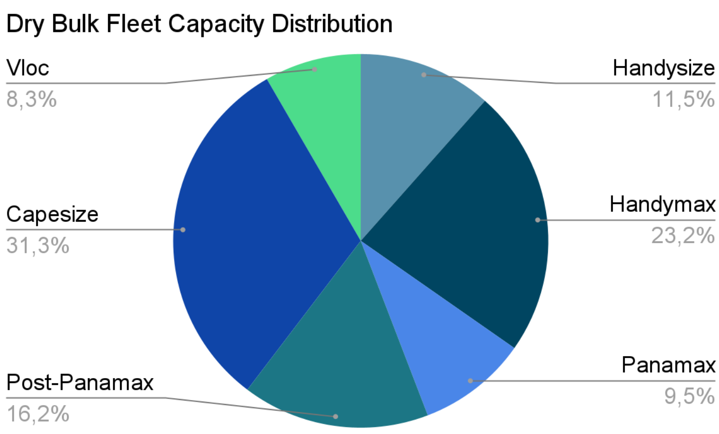
Name	DWT	Number of ships	Total capacity in million DWT
Handysize	10,000 - 40,000	3696	104.9
Handymax/Supramax	40,000 - 65,000	3788	211.4
Panamax	65,000 - 85,000	1160	86.1
Post-Panamax	85,000 - 100,000	1743	147.1
Capesize	100,000 - 200,000	1593	285.3
Vloc	200,000 +	246	76.0

Source: USDA, 2020

Figure 2. Dry bulk fleet distribution

Figure 3. Dry bulk fleet capacity distribution





Source: USDA, 2020

As can be seen from the table and figures, the distribution of the fleet and the fleet capacity are different. Handymax/Supramax consist of 31% of the fleet and has the largest number of ships in the dry bulk fleet, while it distributes 23.7% of the DWT capacity. It is notable that only 246 Very large bulk/ore carriers in the dry bulk fleet, consisting of barely 2% of the fleet, has 8.3% of the total DWT capacity. This is logical since these ships are also costly and can only operate at limited ports and therefore sail limited routes. A similar trend can be seen in the Capesize fleet, in which 13% of the fleet has 31.3% of the DWT capacity.

To summarize, the shipping market is divided into three segments: oil tankers, container ships and dry bulk carriers. The dry bulk carriers are divided by size characteristics and most importantly on the DWT. Certain types of boats often carry the same commodities and travel the same routes. It can be concluded that the total size of the fleet and the total fleet capacity has increased steadily over the last decade. Also, it can be stated that economies of scale have become more important in the shipping industry, as the vessels have become larger over the years. As the dry bulk shipping market is now generally covered, it is necessary to investigate what is transported with these dry bulk ships. For that reason, the next sub-chapter discusses the dry bulk market.

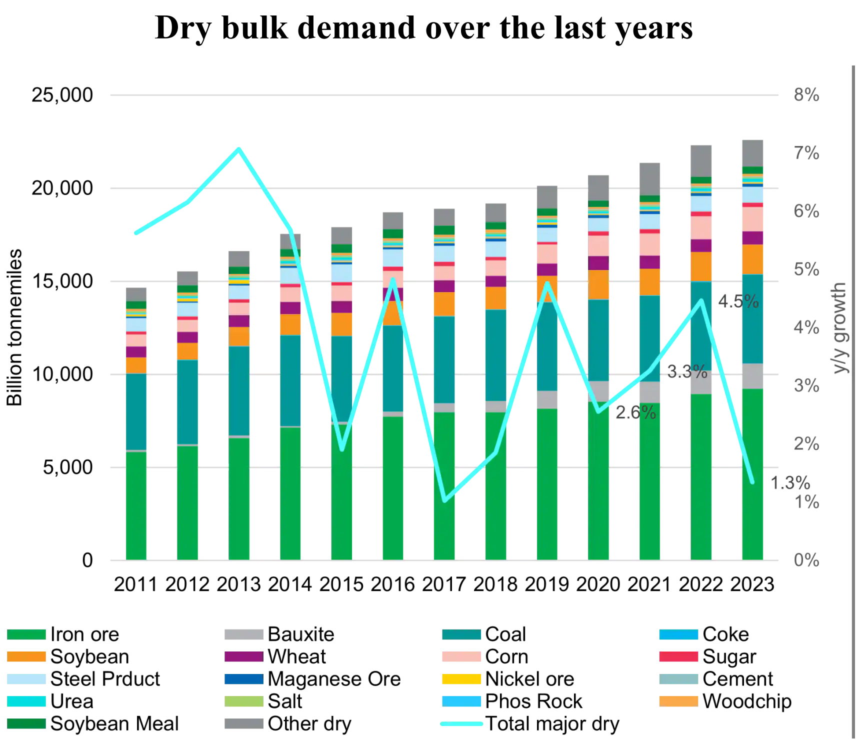
2.2. Dry bulk market

It is generally known that the whole shipping sector, and in particular the bulk shipping business, is extremely unstable because it is a global market impacted by political unrest and cyclical changes in the global economy. Stopford (2009) states that the most typical definition of bulk shipping is the irregular timetabled conveyance of homogenous bulk cargoes by bulk boats. The market for dry cargo consists of large quantities of unpackaged, homogenous commodities that are transported without any intermediate packaging. This market is strongly related to the industrial sector, as raw materials are transported to producing nations (Over, 2017). Prior to examining the dry bulk market, is it necessary to investigate the different types of dry bulk.

2.2.1. Distribution of dry bulk

As stated earlier, the transportation of dry bulk cargoes is often divided into major bulks and minor bulks. The significant bulk cargoes make up the great majority of dry bulk shipments by their total DWT and comprise mainly: grain, coal, and iron ore. Minor bulk cargoes consist mainly of agricultural goods, and mineral cargoes including metal concentrates, cement, forest products, and steel products (Kanamato et al., 2021). In the transportation sector, tonne-miles are frequently used to measure the volume of cargo transported over a distance. The following figure shows the demand for different dry bulk types over the years:

Figure 4. Dry bulk demand over the years



Source: S&P Global Market Intelligence (2021)

As can be seen, the left y-axis represents the tonne-miles in billions and the right y-axis the year-to-year growth of it. Total capacity in 2012 the total demand for dry bulk reached almost 15,000 billion tonne-miles and grew steadily over the years and in 2020 reached over 20,000 billion tonne-miles. It is notable that the demand for iron ore grows steady, while the demand for coal remains roughly the same.

2.2.2. Types of dry bulk

One of the most common dry bulk commodities handled nowadays is iron ore. Prior to arriving at their final destination, iron ore is also processed into precious metals in developing nations for a variety of reasons, the primary ones being cost reductions for the end user and the collection of foreign exchange for these nations. Most iron ore comes from Australia and Brazil and ships to China mostly by Capesize and VLOC ships and to a lesser extent by Panamax ships. The demand for iron ore transportation is affected by many factors, such as the Baltic Dry Index, the price of iron ore, geopolitical situations, and the global market in general. These factors all have an impact on the demand for iron ore transportation and thus affect the freight rate (S&P Global Market Intelligence, 2021; Over, 2017).

The second-largest dry bulk commodity handled in modern times is coal. Coal is a very important source of energy for many businesses and nations. Maintaining a steady supply of primary energy is one of the fundamental challenges in the endeavour to keep the nation's electrical supply available, for many countries, such as Indonesia, coal has been one of the primary energy sources utilized to sustain electrical supply (Yunianto et al., 2018). However, coal is a substance with a wide range of characteristics, which can be divided into steam coal and coking cool. Steam coal is mostly used to generate power in electric power plants, while coking coal is primarily used as a heat source in industrial applications, such as making steel. Numerous nations contain varying amounts of coal. The consumption of the world's main industrial hubs is what essentially determines coal transport patterns and thus there are no ‘standard’ routes but there are major importers and exporters. Australia and Brazil are considered major exporters and the Far East is considered a major importer of coal and is mainly transported by Capesize ships and to a lesser extent by Panamax ships. The demand for coal transportation depends on the dynamics of the coal market, the availability of coal, geopolitical factors and the global market in general (S&P Global Market Intelligence, 2021; Over, 2017).

The third major bulk is known as grain, but it includes a wide variety of grains, including soybeans, wheat, oats, corn, and barley. It can be seen in the figure that these types of grain consist of the third major bulk category in the total tonnes-miles distribution. Grain is mostly distributed by smaller ships such as the Handymax, Handysize and the Panamax. This is partially due to the difficulties involved in shipping grain, such as the need to clean the cargo compartments before loading and fumigate them to get rid of rats and insects. Numerous factors, such as grain supply and demand, harvest-related weather, geopolitical concerns, and other market dynamics such as the Baltic Dry Index, have an impact on grain transportation and the freight rate (S&P Global Market Intelligence, 2021; Over, 2017). The distribution of the major dry bulk commodities by their respective cargo can be summarized in the following table:

Table 2. Dry bulk commodities per different vessel size (percentage of total shipments)

	Iron ore	**Coal**	**Grain**
Capesize	70%	45%	7%
Panamax	22%	40%	43%
Handyclass	8%	15%	50%

Source: Grammenos 2013 page 323

Apart from grain, coal, and iron ore, there are a number of additional dry bulk commodities that are sent in significant amounts globally. These include cement, salt, urea, wood chips, coke, bauxite, Phos rock, and sugar. Every one of these commodities has unique traits, demand patterns, and transportation needs. For sustaining industries including energy, construction, agriculture, and manufacturing as well as for promoting global commerce and economic growth, the shipment of these various dry bulk commodities is essential (S&P Global Market Intelligence, 2021).

2.2.3. Dry bulk trading routes

The previous section clarified the different types of boats, the distribution of each type and the capacity of these ships. This section focused on the different types of dry bulk commodities and discussed the developments in the total tonnes-miles demand over the years. As the distribution of boats relative to bulk is clear, it is necessary to investigate the geographical aspects. For that reason, the standard trading routes will be clarified for the major dry bulks.

Most iron ore comes from Australia and Brazil and ships to China and Western Europe with Capesize ships. In general, Australia is rich in natural resources such as iron ore and coal and one of the leading exporters of these two bulk goods. For that reason, the routes between Australia and the major importing continent of Asia, mainly the Far East, are well-established and optimized. Through the years, these routes developed and are now even taking advantage of the currents and weather conditions to reduce fuel consumption and voyage duration. Over the years, strong business ties have been forged between Australia and significant coal and iron ore importers in the Far East. A steady flow of goods is frequently made possible by long-term contracts and commercial agreements, attracting Capesize ships from Australia to the Far East. For that reason, Australia has built a reliable and effective port infrastructure that can manage the loading and unloading of massive amounts of bulk goods.

Capesize boats may be accommodated in ports in Western Australia such as Port Hedland and Dampier, which streamlines the loading procedure. The combination of well-developed routes, many long-term trade agreements and efficient port infrastructures have made routes for transporting iron ore and coal from Australia to the Far East for Capesize ships very popular (Grammenos 2013; Over, 2017, S&P Global Market Intelligence, 2021).

Asia and Western Europe are the largest consumers of coking coal. Australia exports most coking coal to Asia, while South Africa and the United States are the main exporters of Western Europe. Steam coal is mainly exported by Australia, Colombia, Indonesia, South Africa and Russia to Asia. Coal is transported about as much by Capesize as Panamax on these routes. Most grain is produced in the United States, followed by Argentina, Canada and Australia. The main importers are Asia, the Middle East, Africa and Latin America. Only on the route from Argentina and River Plate to Near East and East Europe make use of Capesize vessels. All other routes make use of smaller vessels such as Panamax and Handyclass vessels (Grammenos 2013; Over, 2017, S&P Global Market Intelligence, 2021). The major routes are summarized and can be found in Appendix B Table B.1.

In conclusion, the dry bulk shipping industry still plays a crucial role in world trade by enabling the movement of vital goods around the globe. The total tonnes-miles demand increased steadily over the years, suggesting the demand for dry bulk is growing year to year. Iron ore is the most transported major dry bulk commodity, followed by coal and grain. These commodities are transported on different routes with different types of vessels. This is accompanied by the demand, as iron ore is mostly transported by Capesize vessels, coal mainly with Panamax vessels and grain with Handyclass ships. Given that this has become clear, it is necessary to investigate the chartering of these ships' and their commodities in the next chapter.

3. Literature review

In the following chapter, the theoretical framework for this research is established. The framework consists of three sub-chapters. The first chapter elaborates on chartering, which is divided into four sub-sections. Starting with a systematic review of basic elements of contracts, followed by a section about time chartering and then voyage chartering. At last, the differences and links between time chartering and voyage chartering are discussed. This chapter provides general elements of contracting and previous research in different contracting.

The second sub-chapter examines freight rates, this is broken down into three sub-sections. The first sub-section starts with an overview of the Baltic Freight Index. The second sub-section continues the timeline and deals with the Baltic Dry Index. The third and last sub-section investigates the existing prediction methods in previous research. This chapter elaborates on the used integrated index, the Baltic Dry Index, which will be used to investigate the price dynamics.

The final sub-chapter elaborates on the freight market and is divided into four sub-sections. Firstly, the demand and supply of dry bulk shipping and its effects on freight rates are discussed. Secondly, there is an elaboration on all the different elements influencing the freight rate and pricing in dry bulk shipping. Thirdly, the risk premium is investigated. Thereafter, it should be clear which elements from individual contracts will and will not be investigated for this research in order to investigate price dynamics.

3.1. Chartering

There are three distinct sorts of transactions in the freight industry as stated in the introduction. The first is a time charter, in which a ship is chartered on a day-to-day basis. The second is a voyage charter, whereby the shipper purchases transportation from the shipowner at a set cost per ton of goods. The third is a spot rate contract, which is a short-term contract at the prevailing market rate for single-voyage transportation of freight for a particular shipment between two ports (Stopford, 2009). Shipowners (or their representatives) and charterers often negotiate and sign contracts for the chartering of freight in the dry bulk shipping sector. Charter party agreements are the conventional name for these contracts. Dry bulk commodities must be transported, thus charterers, who may be dealers, cargo owners, or operators, enter into these contracts to lease vessels for a specified time or journey. These negotiations of transportation are made up of certain contracts, which consist of a couple of basic elements in order to make sure both parties fulfil the agreements (Pirrong, 1993).

3.1.1. Basic elements of contracts

Chartering contracts consist of two parts, a not-negotiable part with specifics about the vessel and a negotiable part with the terms of the contract. The terms of these contracts are divided into three groups: conditions, guarantees and intermediate terms (Dobre, 2016). The conditions of the contract are of great significance to the execution of chartering. If conditions are not met, the innocent party has the right to regard the contract as having not been fulfilled. In the case of a charter agreement, a guarantee is a small clause that has no bearing on how the journey will proceed. Failure to adhere to such a condition does not grant the innocent party the right to consider the contract to be fulfilled, but instead allows him to pursue compensation. Contractual provisions of a complex character exist in every chartering agreement and cannot be viewed as terms or assurances, but are taken into account as intermediate terms. Failure to do so can be regarded as fundamentally important in one circumstance and unimportant in another. The loss incurred by the innocent party is typically what establishes the line between condition and warranty (Dobre, 2016).

It is known that some elements in contracts for chartering freight in the dry bulk shipping market are highly negotiable and others are not. Specific information and elements of the vessel are not negotiable of the vessel but must be in the contract. Information such as the name of the ship, the year of construction, the nationality of the ship, the dead weight, gross and net tonnage, the draft of the ship, length of the ship and height of the ship above the waterline. These are not negotiable and are a given fact but are always implemented into the contract if there is an agreement on the conditions and the ship is nominated (Alizadeh & Talley, 2011; Dobre, 2016).

Other elements, such as freight rates, laytime, demurrage, cargo handling responsibilities, and insurance provisions, are negotiated. The freight rate is the price or cost of transporting cargo, negotiations of the freight rate are a bargaining process between shipowners (or representatives) and the charterer. Often the laytime, which is the time which is allocated for loading and unloading the vessel, is negotiated depending on the previous route and the coming route. The demurrage serves the purpose in contracts to keep shippers to their agreement, which is an agreed penalty clause for the demise of the laytime.

Laytime and demurrage are closely linked and for that reason, both parties negotiate about both of them. Other elements that may be important for the contract are the lifting appliances of the ship, the number of storerooms, the sizes of the storerooms and the cargo hatch. Both shippers and shipowners must request and supply all the elements necessary to carry out the economic calculations and the planning of loading, transport and unloading operations (Alizadeh & Talley, 2011; Köhn & Thanopoulou, 2011; Dobre, 2016).

3.1.2. Time chartering

A time charter is hiring a ship for a set period of time, usually months or years. It is important to keep in mind that it is not a lease, charters do not acquire possession of the ship (Coghlin et al., 2014). In this situation, the shipowner provides the vessel and its crew as well as paying for running expenses, and the charterer pays a set fee per day or per month to use the vessel (Stopford, 2009).

Time charter contracts are conceived as a constant income and for that reason they are preferred over voyage chartering in times the freight prices are unfavourable for shipowners. In economic analysis, this can be seen that time charter contracts have a smoother line with less variability compared to voyage charter contracts. The shipowner is only accountable for the vessel's capital expenses and operating expenses, which cover all maintenance expenditures, whereas the charterer is accountable for the costs associated with the voyage. The shipowner is also entitled to payment of rent, which is paid semi-monthly in advance, and the charter party contract includes many additions and deductions (Stopford, 2009; Coghlin et al., 2014).

3.1.3. Voyage chartering

In a voyage charter, the shipowner consents to move a particular amount of goods between ports in a predetermined amount of time. The shipowner (or their agent) commits to providing the vessel and its crew, and the charterer (typically a cargo owner, merchant, or operator) commits to chartering the vessel for the transportation of their products. The amount of cargo carried or the distance travelled is typically depending on the freight rate. In voyage chartering, the shipowner undertakes to conduct one or more specified voyages under a charter agreement in exchange for the payment of freight and demurrage. The freight is paid to the shipowner in exchange for the transportation and is typically calculated pro rate based on the DWT cargo loaded or on the net tonnage capacity of the ship (Cooke et al., 2014).

A voyage charter's main distinguishing feature is that it concentrates on a single journey rather than a certain amount of time. The shipowner and charterer will normally negotiate and come to an agreement on the terms and conditions of a trip charter after taking into account details like the kind and quantity of cargo, loading and discharge ports, laytime, freight rate, demurrage, and any other particular needs. Both shipowners and charterers have freedom with voyage charters. By securing individual journeys, shipowners may make the most of the use of their fleet, while charterers are free to choose the best ships and routes for their cargo requirements and market circumstances (Stopford, 2009; Cooke et al, 2014).

3.1.4. Spot-rate contracts

The spot rate contract is a significant contract form in the dry bulk shipping sector in addition to voyage charters. Spot rate contracts are different from voyage charters in terms of their scope and length. In a spot-rate contract, the transportation of products is agreed upon for immediate or almost immediate delivery at the going market prices. Spot-rate contracts concentrate on a particular cargo rather than a defined period, in contrast to voyage charters, which entail pre-planned travels and predetermined intervals. They give both shippers and charterers flexibility, enabling them to react rapidly to shifting market circumstances and urgent shipping demands (Stopford, 2009).

Spot pricing agreements often entail the charterer leasing a ship for a single trip or delivery of cargo. While the charterer, who is often a cargo owner, merchant, or operator, charters the vessel for the transportation of its products, the shipowner or their agent commits to supplying the vessel and its crew. The freight rate, which establishes the payment to the shipowner for the transportation, is often determined using variables like the amount of cargo, its weight, or the ship's net tonnage capacity.

Spot-rate contracts offer flexibility and freedom of choice to both shipowners and charterers, which is one of its main benefits. By negotiating unique spot-rate contracts for each of their vessels, shipowners can maximise the usage of their fleet, while charterers have the freedom to choose the ships and routes that best fit their cargo needs and market circumstances. In the dry bulk shipping sector, spot-rate contracts are essential because they provide quick shipping options and the flexibility to adjust to shifting market conditions. Both shipowners and charterers who are looking for effective and adaptable transportation solutions for their goods have a great choice in them (Kavussanos & Alizadeh, 2001; Stopford, 2009).

3.1.5. Time vs Voyage chartering and Spot-rate contracting

To summarize, a time charter involves the hiring of a vessel for a specific duration, while a voyage charter involves the hiring of a vessel for a specific journey or voyage and spot rate contracts are different from voyage charters in terms of their scope and length. Kavussanos (1996) investigates market volatility as a risk indicator for dry-bulk ships. The research specifically contrasts time-charter and spot rate volatility estimations, while the volatility of dry-bulk boats of various sizes is also contrasted. It is discovered that when spot rates are employed, small boats are less dangerous than bigger ones and time-charters are more unpredictable. The policy repercussions for risk-averse ship owners include using the spot market rather than time charters and investing in smaller boats rather than larger ones (Kavussanos, 1996).

The research of Axarloglou et al. (2013) empirically examines the properties of the spread between voyage and time-charter rates that change over time and provides evidence that these characteristics are closely related to the maritime industry's business cycle (market demand), expectations for the future market demand, and market volatility. The results suggest that managers decide to commit firm resources for a brief period of time during a market upswing in order to keep flexibility in better utilising the impending business possibilities. On the contrary, in a market downswing, the commitment is for a long period, protecting the company resources from a lack of business opportunities. Overall, changes in the time-varying gap between voyage and time-charter rates provide managerial perspectives on resource allocation that can improve decisions on chartering, budgeting, and financial management about the time commitment of resources in the maritime sector (Axarloglou et al., 2013).

Zhang & Zeng (2015) examine the link between time charter and spot freight rates as well as the price discovery function of time charter contracts using three different types of dry bulk ships: Capesize, Panamax, and Supramax. The impact of time charter prices on spot freight rates is examined using an impulse response function and a vector error correcting model. Empirical investigations show that time charter contracts have a price discovery function and that there are two-way lead-lag correlations between time charter and spot freight prices. Zhang & Zeng (2015) conclude that a greater price discovery function is produced by smaller ship sizes and longer voyages for time charter rates.

To conclude, the decision-making of contracts consists of many elements for the shipowners as for the charterers. The choice for shipowners and charterers between a time charter, a voyage charter or a spot-rate contract depends on the type of vessel, the type of dry bulk, the geographics, the specifics of the vessel, the economic landscape and many other macroeconomic factors. It can be concluded that there is not one particular element to focus on, but it is necessary to focus on the bigger picture. However, to create the bigger picture it could be useful to investigate which basic elements could be used and to compare to price dynamics. As the price dynamics can be investigated by an integrated index, it is necessary to dive into an appropriate price index for dry bulk shipping. For that reason, the following sub-chapter dives into the Baltic Dry Index, as it is considered an important economic parameter in the shipping industry and closely linked to freight rates and price dynamics.

3.2. Freight rates

The Baltic Freight Index (BFI) is a well-known (integrated) economic indicator and is launched on June 11th, 1984, with a base value of 1000 index points. It is created to serve as a representation of the freight rates that shipowners charge for the chartering of their vessels. The freight rate is the unit of measurement used to indicate the cost of marine transportation (Duco, 2010). Since freight rates affect how much money the shipping industry makes, changes in the BFI time series can (at least in part) be used to gauge changes in the sector's financial health and for that reason freight rate and BFI are correlated (Cullinane et al., 1999). It monitors the demand for transporting dry bulk commodities and illustrates the overall health and activity of the dry bulk shipping business on a worldwide scale. It offers insightful information about changes and patterns in freight prices for moving important raw commodities including iron ore, coal, grain, and other bulk cargoes (Stopford, 2009).

3.2.1. The Baltic Freight Index

The Baltic Exchange, an independent company with headquarters in London, compiles and publishes the BFI every day. Based on the typical daily freight prices for various dry bulk vessel types and itineraries, the index is calculated. The freight rates charged on eleven important trade routes define the value of the BFI at any given moment. The panellists from the Baltic Exchange in London provide their analysis of the freight rates for each of these trade routes every morning. Where applicable, these evaluations are based on real freight rates; alternatively, where actual fixtures are not accessible, they are based on educated guesses as to what the freight rates would be (Cullinane et al., 1999).

The top and lowest ratings for each trade route are eliminated, and the average of the remaining ratings is determined as a preventative step to avoid any one broker having an undue influence on the market. The weighting of these average freight rate figures then reflects the relative significance of each trade route to the dry bulk shipping industry. They are then combined to create the BFI, which is released each working day at 1 o'clock (London time). To guarantee that the index stays representative, the trade routes and their corresponding weightings are continuously examined. In actuality, there have been several changes made to the BFI's makeup since its debut (Cullinane et al., 1999; Stopford, 2009).

The majority of these modifications have been very few tweaks, but a handful have been more significant. On November 3, 1993, all trade routes served by vessels between 25,000 and 50,000 deadweight tonnes (DWT) (also known as Handysize routes) were removed from the index, bringing the total number of component routes inside the BFI down from 13 to 11. The new index at that time was made up of four Capesize routes, size range exceeding 75,000 DWT and a weighted average of seven Panamax routes, with vessels size range of 50,000 – 75,000 DWT (Cullinane et al., 1999).

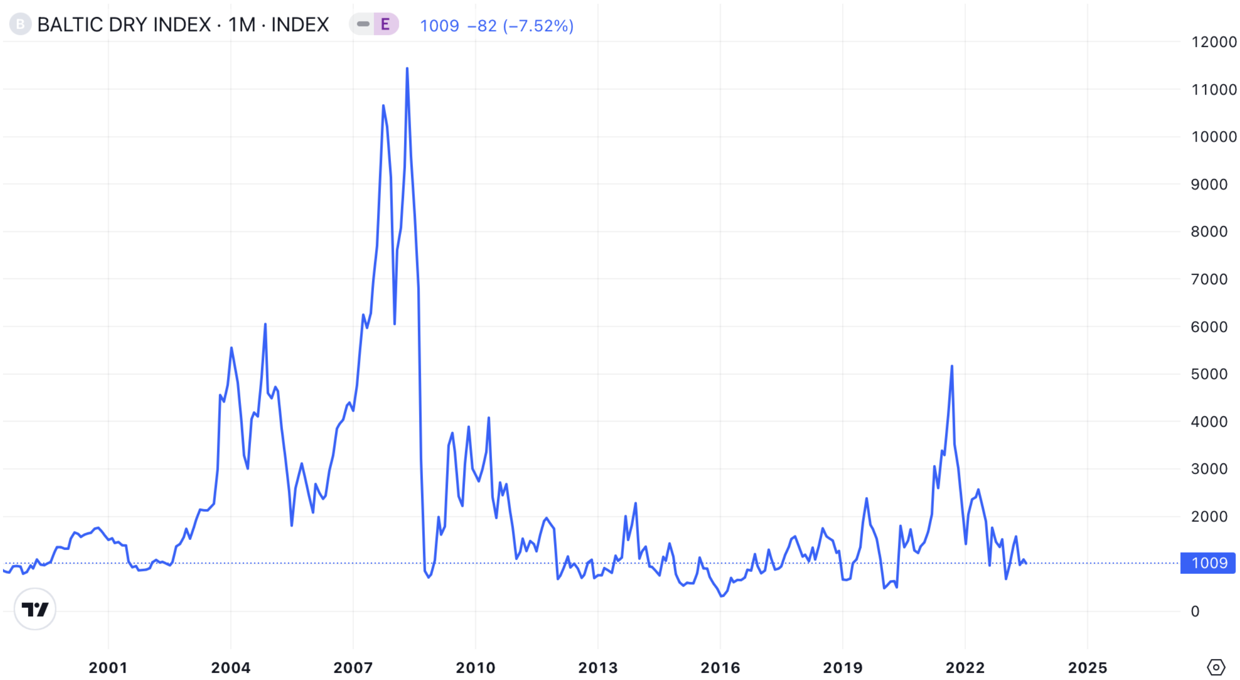
3.2.2. The Baltic Dry Index

The Baltic Exchange in London revolutionised the BFI in 1999, to the Baltic Dry Index (BDI) which represents the dry bulk international shipping freight index (Duro, 2010). The Capesize, Panamax, Supramax, and Handysize sub-indices reflect each distinct vessel type and are weighted averaged to create the BDI. Several variables that affect the supply and demand dynamics in the dry bulk shipping industry have an impact on the BDI. The demand for shipping services is largely influenced by economic circumstances, industrial activity, infrastructural development, and worldwide trade patterns. These factors also have an impact on the need for raw materials. In addition, the BDI may be impacted by environmental factors, geopolitical developments, and regulatory changes (Ghiogre et al., 2013; Stopford, 2009; Over, 2017).

For shipowners, charterers, dealers, and analysts is the BDI a crucial instrument. In order to arrange their cargo, negotiate freight rates, and evaluate market circumstances and trends, they can use this information to their advantage. A high BDI is a sign of healthy global economic growth and increasing trade volume, as well as strong demand and higher freight costs. A low BDI, on the other hand, denotes less demand, lower freight rates, and perhaps difficult market circumstances. Similar to the BFI, the BDI is representative and correlated with the freight rate (Ghiogre et al., 2013; Lyridis et al., 2014).

The index is a reference point for the cost of shipping the main raw commodities by sea. The index, which measures different sizes of dry bulk carriers, is made up of three sub-indices: Capesize, Panamax, and Supramax. The Baltic Dry Index considers 23 distinct maritime lanes used to transport grains, coal, iron ore, and many other commodities (Trading Economics, 2023). The following figure shows the trend of the BDI since 1999:

Figure 5. The Baltic Dry Index month-to-month



Source: Trading Economics: The Baltic Dry Index month-to-month data (2023)

As can be seen from the figure, the BDI fluctuates relatively heavily over the years. Especially in May 2008, during the financial crisis, the BDI spiked to 11,140 and dropped to 652 in December 2008. It can be stated that the trend of BDI fluctuates a lot over the years. Therefore, the next section investigates which models are used to predict and forecast the BDI in the literature.

3.2.3. Forecasting and the Baltic Dry Index

Cullinare et al. (1999) state that due to the shipping industry's thin profit margins, even a little change in freight rates can have a significant impact on profitability. The apparent potential benefit of being able to precisely forecast the market is that shipowners will be able to better plan how to optimise earnings and/or avoid losses sustained in the shipping market.

Michail & Melas (2020) investigates the connection between trade in seaborne commodities and freight rates. To overcome the problem of data availability, a Bayesian Vector Autoregressive technique is used on data from the Baltic Dry Index, the Baltic Dirty Tanker Index and the Baltic Clean Tanker Index. It is found that the volume of seaborne commodity trade has no effect on all three indices. However, the authors suggest that there appears to be a link between the freight indices since changes in one might have an impact on the other (Michail & Melas, 2020). A recent study by Michail (2020) has demonstrated evidence of a positive association between global economic development and seaborne commerce carried by all sorts of vessels (containerships, dry bulk vessels, and tankers), as well as evidence of a negative relationship between oil prices and seaborne trade. It can be stated, there is a connection between the BDI and other indices.

Bakshi et al. (2010) investigate the relationship between BDI and stocks, in order to see if the growth of the BDI has the capacity to forecast a variety of stock markets. In-sample experiments and out-of-sample data on the three-month growth rate of the BDI as a predictor of global stock market returns can be used as predictors in commodity index returns and growth in actual economic activity. The findings presented in this research thus contribute to highlighting the usefulness of the BDI as a predictor that captures heterogeneity across the real and financial sectors of several global economies. Our line of inquiry may also be expanded to investigate the potential contribution of the BDI growth rate as a price component in elucidating differences in the cross-section of predicted stock returns (Bakshi et al., 2010).

The artificial neural networks (ANNs) model of Lyridis et al. (2014) attempts to anticipate the BDI using real data for a twenty-year period for a variety of macroeconomic parameters (nineteen) and nautical indices (four), for which data are accessible on a daily, weekly, monthly, or three-monthly basis. The data set, which covers the 20 years from 1991 to 2011, includes market movements through the more severe upheavals and downturns. They conclude that the ANNs model is more precise than current techniques, by accounting for macroeconomic factors in addition to other criteria to forecast the development of the Baltic Dry Index (Lyridis et al., 2014).

In summary, many authors suggest different approaches to forecast the Baltic Dry Index or other indices using the BDI. This suggests that the BDI, as an economic performance parameter, could be influenced by many other elements of the dry bulk freight market. From this, it can also be validated that the BDI can be used to capture price dynamics. For that reason, it could be interesting to investigate what other elements could be used to predict BDI. Therefore, the following section addresses these elements by clarifying the dry bulk freight market.

3.3. The freight market

This sub-chapter investigates the freight rate market and which corresponding elements could be interesting to use in determining price dynamics. On the freight market, shipping services for the transportation of products are traded. Midway through the nineteenth century, trading as a commodities and shipping exchange began at the Baltic Shipping Exchange, the initial freight market. The freight market still functions as a market where maritime cargo is bought and sold today, although most transactions are now done over the phone, over email, or through messaging apps (Stopford, 2009).

Truett & Truett (1998) state that the demand for shipping services and freight rates are related to one another. According to McConville (1999), the freight market generates a situation where freight rates rise to a point where shippers' demand and transportation companies' supply are equal. The level of freight rates in the market is determined by the demand and supply functions of the dry bulk shipping sector (Stopford, 2009).

3.3.1. The effects of demand and supply on freight rate

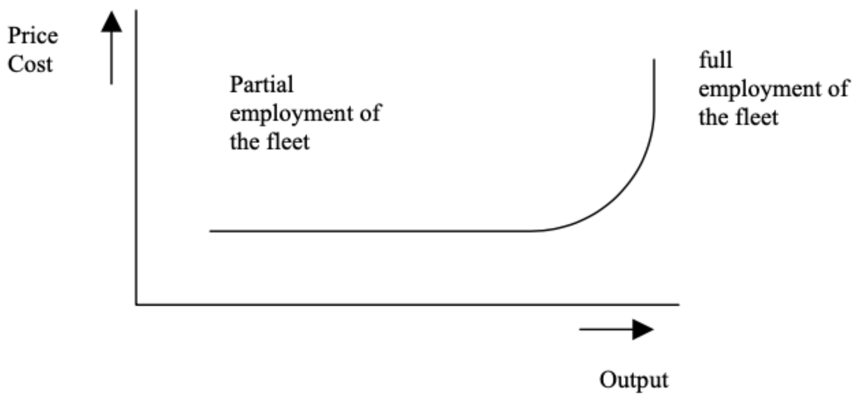
The cost of the transportation of goods is called freight and depends on the type of goods being transported, the distance that the vessel must travel to convey the cargo, and the stowage factor of the cargo. The balance of demand and supply determines eventually the cost of transporting the freight (Over, 2017). This section is divided into three smaller parts, which will discuss the effects of demand and supply on the freight rate.

3.3.1.1. The effects of supply on freight rate

Concentrating on the supply side of the market, Stopford (2009) argues that the supply of shipping services depends on five main factors, namely: 1. The fleet stock; 2. Ship-building production; 3. Scrapping; 4. Fleet productivity; 5. The current level of freight rates in the industry. The supply depends on the fleet stock, it is clear that a higher fleet stock will result in a stronger supply of bulk shipping, due to the fact that more ships will be available to transport cargo. Shipbuilding increases the fleet, as more ships hit the market, supply grows. An increase in scrapping and losses would result in a comparable drop in the availability of bulk shipping. Whenever the fleet is slow-steaming, or not working at its highest rate, it further reduces the efficiency of shipping and therefore reduces the productivity of the fleet.

The freight rate impacts the supply side of bulk shipping activities both in positive and negative ways, a higher freight rate increases profitability for shippers and thus increases the supply for ships. The effects of employment of the fleet on the price costs are modelled by Veenstra (1999):

Figure 6. Bulk market supply curve



Source: Veenstra (1999)

As can be seen from the figure, the variation in the freight rate is demonstrated by the shape of the supply curve. The flat segment of the supply curve, suggests elasticity and should cross the demand curve when the fleet is partially operational. The flat part of the curve suggests that fluctuations in demand have little impact on freight rates, as long as the fleet is not entirely employed because the fleet can quickly adapt to them. However, while the entire fleet is working, the supply cannot simply keep up with the rise in demand. As a result, the freight rates (Price Costs) significantly increase as can be seen at the right (Veenstra, 1999). However, the demand curve is not indicated in the figure because there are various factors influencing the demand. For that reason, the following section investigates the demand and its effects on the freight rate.

3.3.1.2. The effects of demand on freight rate

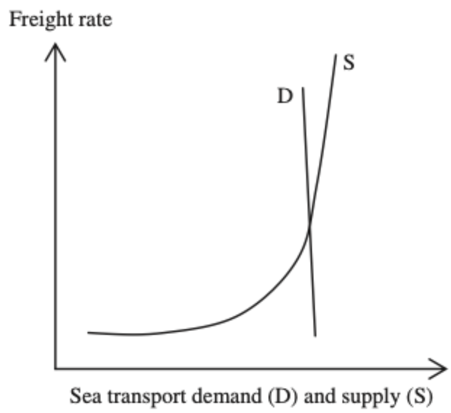
The demand for bulk shipping services derives from the desire for the products being transported, hence it is subject to the cyclical nature of global trade since it is a demand for the movement of goods. Despite the fact that this demand is derived, Stopford (2009) shows how it can be broken down into five main factors: 1. The level of global economic activity; 2. The level of seaborne commodity trade; 3. The average haul, or the distance that the commodities must be transported; 4. The state of current politics; 5. The level of transportation costs. These five factors all impact demand, but the shippers can barely impact one of these factors, because it is beyond their control. It can be stated that these factors are more macroeconomic and for that reason mainly impacted by world economic activity. If there is a high demand for goods, will result in a high demand for the means by which these would be transported, leading to an increase in the demand for bulk-shipping services (Jacks & Stuermer., 2021).

3.3.1.3. Freight rate and fleet availability

Since the early 1930s, marine economics has conducted extensive research on the link between freight rates and the fundamentals of supply and demand. In classic literature, the need for bulk shipment is treated as independent of freight prices (Koopmans, 1939). As stated in the previous section, Zannetos (1966) concluded that in price setting there is an association between voyage charter rates and fleet utilization. The results of Eriksen's (1983) analysis of the potential impact of freight prices on the demand for freight services point to an inverse relationship between demand and freight rates.

Beenstock and Vergottis (1993) provided a historical overview and in-depth investigation of the modelling of the bulk shipping sector and according to them the relationship between supply and demand determines the freight rate. According to Strandenes (2004), supply and demand also affect fleet utilization too. When every vessel is in use, increasing fleet utilization through faster speeds, shorter ballast legs, shorter port stays, and postponing routine maintenance will enhance supply in the short run. which will also affect the price. In the dry bulk shipping market supply and demand balance is essential in order to investigate the freight rates. By anticipating and adapting to changes in freight rates, stakeholders may optimise their operations and pricing strategies by having a solid understanding of the dynamics of fleet availability (Strandenes, 2004). Stopford (2009) outline the demand and supply and the impact on freight rate in the following figure:

Figure 7. Sea transport demand and supply relative to freight rate



Source: Stopford (2009)

The fleet supply function (S), which has the form of a hockey stick, functions by putting and taking ships out of service in response to freight rates. It can be seen that whenever the freight rates are low, the ship supply function is elastic, but whenever the freight rate is high, it becomes inelastic. The vertical fleet demand function (D) demonstrates how charterers respond to changes in freight rates. Shippers send the goods regardless of cost since there is no other available route of transportation (Stopford, 2009).

Regli & Nomikos (2019) demonstrate that the location and employment status of the fleet of very large crude oil carriers affects the evolution of crude oil tanker freight rates. By modelling the voyage charter market's short-term capacity in order to serve as a stand-in for the proportion of ships that are open to orders, they conclude that fleet availability explains the variation in the freight rate (Regli & Nomikos, 2019). To summarize, in previous literature there is a relationship detected between price and fleet utilization, however there is not much literature about this topic.

3.3.1.4. Price setting based on demand and supply

According to Thorburn (1960), freight rates will eventually be equal to costs if transportation markets are completely competitive. This resulted in individual investigations of ships, their expenses and the shipowners in the freight market. The research of Zannetos (1966) is considered very revolutionary in the freight market, as his major argument focused on the part expectations play in setting prices. He stated that the voyage charter rate has a considerable impact on the expectation because it is based on the short-term supply and demand in the shipping industry. In price setting the association between the voyage charter rate and the number of idle ships was also established (Zannetos, 1966).

In sum, the main drivers of demand are the number of commodities to be carried, the distance travelled, and political decisions impacting commerce and shipping. The main drivers of the supply are determined by the total operational tonnage, the average ship life, and operational efficiency. In other words, these studies suggest there are a lot of macroeconomic factors influencing the supply and demand in dry bulk shipping. However, these macroeconomic factors are not the only factors influencing the freight rate, it can be stated from other studies that prices and fleet ability are also both important factors for the supply and demand in dry bulk shipping. In addition to the supply and demand, there are other factors linked to the freight rate and thus pricing. For that reason, the next section dives into elements of individual contracts and investigates which other factors are relevant to the freight rate.

3.3.2. Elements influencing freight rate and pricing

Alizadeh & Talley (2011) state that most researchers focus on macroeconomic determinants of shipping freight rates, while the factors influencing shipping freight prices on a microeconomic level have not been systematically studied. Their research aims to explore the microeconomic factors that affect freight prices in the dry bulk shipping industry by investigating individual dry bulk charter contracts. Investigations are also conducted into the variations in freight prices along the principal dry bulk shipping routes, the geographic distribution of shipping activity worldwide, and the length of the laycan time for shipping contracts. They concluded that these microeconomic factors, such as the vessel’s deadweight, age, route, freight rate level and laycan period are important in determining the dry bulk shipping freight price (Alizadeh & Talley, 2011). For that reason, it is necessary to investigate different elements of individual contracts in the following sections.

3.3.2.1 Freight rate and the element time

As there are multiple time elements in contracts such as the date when the contract is published, the start date of the contract, and the start of the laycan period. There is also quite some 'end' dates of the contracts, such as the end of the laycan period, the desired arrival date, the estimated date and eventually the date when the vessel arrived at the location (Stopford, 2009). The time series characteristics of maritime freight rates, such as their reliance on historical data and whether univariate or multivariate time series models adequately reflect the dynamics of freight prices, have also been investigated in other research. Researchers used these models in order to predict the shipping freight rates and their volatility on different parameters (Kavussanos and Alizadeh, 2002; Adland and Cullinane, 2005).

The element of time has been an interesting topic for dry bulk shipping, as articles related to the stochastic process characterise the dynamics. Poblacion (2015) examines how freight prices change with the seasons and discover that models with stochastic seasonality perform better than models with deterministic seasonality. In his research, he came up with a factor model with one element being a seasonal factor for the stochastic behaviour of TCE (Time Charter Equivalent) and WS (World Scale) pricing for five routes defined by the Baltic ranging from February 2009 to February 2014. Poblacion (2015) concluded that freight rates are higher in the winter and spring than in the summer and fall. It is stated that these variations are stochastic rather than deterministic and therefore ship owners and charterers have to adapt their business strategy accordingly. In other words, the pricing is also depending the element of time, but time is a broad element in dry bulk shipping.

For various ship sizes and for freight contracts with various periods, Kavussanos & Alizadeh (2001) explore the seasonal behaviour of dry bulk freight rates. They also look at seasonality patterns under various market situations. They used spot and time-charter (1-year and 3-year) rates based on the average of daily fixtures from January 1980 to December 1996 for Capesize, Panamax and Handysize vessels. ARIMA, VAR and OLS models are used to model the time series data and identify the effects of seasonality. The results show that deterministic seasonality, or predictable seasonal patterns, is far more prevalent and that stochastic seasonality has been disproved. The type and pattern of seaborne commerce, according to the authors, is what causes the seasonal trends. The conclusion suggests that seasonal oscillations are more pronounced and sharper during market expansion times than during market contraction periods. Kavussanos & Alizadeh (2001) state that the differences diminish as contract duration grows, suggesting longer contract durations result in more stability in freight rates.

Alizadeh & Talley (2011) examines if there is a simultaneous link between the choice of laycan time and the ship freight rate. They used data on Panamax and Capesize trip-charter fixtures from January 2003 to July 2009 for their regression estimates. According to the estimation results, the length of the laycan time affects the shipping freight rate significantly, and vice versa. The findings also show a negative and substantial correlation between the length of trip-charter contracts' laycan periods and the volatility of Capesize freight prices (Alizadeh & Talley, 2011).

Holguin-Veras et al. (2011) identified time-dependent effects on freight distribution and freight generation. These effects are investigated by using a panel formulation with time-dependent parameters and fixed time effects, and then the related cross-sectional models were compared to find time-dependent effects. Results show that all freight generation models, freight distribution models, and empty trip models all exhibit statistically significant time-dependent impacts on freight rates (Holguin-Veras et al., 2011).

In summary, it can be concluded that the time element in contracts certainly has an impact on the freight rate. However, the element of time is also linked to geography, since vessels have to travel certain distances. The distance and routes the vessels have to travel influence elements of time in the contracting and for that reason, the next section investigates the effects of geography on freight rates.

3.3.2.2 Freight rate and the element geography

As previously stated, the BFI and the BDI are based on various ‘standard’ routes operated by Capesize, Panamax and Handysize dry bulk vessels. This suggests that geography is an important element in determining the freight rate. This is confirmed by multiple researchers, which investigate the routes and their effect on freight rates.

For several shipping routes, Laulajainen (2007) looked into variations in freight prices and operational profitability. The research aims to assess any potential inefficiencies in the dry bulk shipping business and confirm its geographic efficiency. Laulajainen (2007) separated the fleet into size segments: big and small Capsizes, Panamaxes, and Handysizes, with class restrictions of 150,000, 80,000, and 50,000 DWT. These class restrictions are made because large boats have distinct movement patterns from tiny ones. The investigated period consists of 1995, 1997 and 1998, because this includes high, medium and low rates. Due to the fact that high prices leave greater room for pricing negotiation than lower prices. The research is conducted manually by transforming unprocessed ship movement data into trade matrices and computing rate functions for each route using the Baltic Freight Index as an explanatory variable in the used regressions. It is concluded that an important aspect of understanding dry bulk freight prices for certain routes is the relationship between demand and available ship tonnage, weighted by sailing distance to a discharging/loading zone (Laulajainen, 2007).

Alizadeh & Talley (2011) want to determine if the choice of laycan time and the ship freight rate are related, or if the two factors are related simultaneously and take for this the geographical features into account. In single equation models, the variables are first treated as dependent variables. They are then taken into account as dependent variables in a model of simultaneous equations. The data on Panamax and Capesize trip-charter fixtures from January 2003 to July 2009 are obtained from the Shipping Intelligence Network (SIN) section of Clarkson's Research Services Ltd website. The difference between the freight rate for a certain shipping contract and the value of the Baltic freight rate index for that particular type of vessel on the fixture date is used as the dependent variable for the regression model used to investigate factors, such as the voyage routes, that affect shipping freight rates in dry bulk markets. Finally, it is concluded that voyage routes are important determinants of dry bulk shipping freight rates (Alizadeh & Talley, 2011).

More recent research by Guan et al. (2019) used a support vector machine model to forecast the dry bulk carrier route selection based on the Baltic Supermax Index (BSI) and historical decision data of various enterprises. The final unloading of the ships at the conclusion of the planning period corresponds to the start of the planning period, and the BSI of the four routes currently chosen by the fleet is an input variable. The data is collected from June 2014 to August 2015 of 320 ship sample results and additionally, the BSI of the same period is added to the data. Through comparative analysis is demonstrated that the support vector machine model suggested in this article performs better than several other widely used approaches in the field, such as K nearest neighbours and linear regression. The research assists that geography is an important element relative to an economic performance parameter, which is in this case the BSI (Guan et al., 2019).

It can be confirmed that many researchers use an economic performance parameter and want to predict it by geographical elements. Therefore, it can be concluded that geography is an important aspect of research to include in investigating freight rates.

3.3.2.3 Freight rate and the elements of the freight

Finally, the elements of the freight, such as the type of bulk and quantity transported, and their effect on freight rates will be investigated in this section. Jonnala et al. (2002) empirically investigate the key variables influencing grain ocean freight rates. The data consists of 12,296 transactions (voyage charters) between international grain dealers and ship owners are included in a time series data set ranging from 1988 to 1997. Jonnala et al. (2002) used for the estimation of model parameters OLS, ARCH and GARCH. The findings demonstrate that freight characteristics such as the ship tonnage contracted for the hauling of certain other dry bulk commodities, as well as journey distance, ship size, contract terms and flag are significant explanators of pricing for shipping grain (Jonnala et al., 2002).

The research of Adland et al. (2016) aimed to identify the factors that influence how a ship's cargo-carrying capacity is used during particular journeys specific to iron ore. The data used is from November 2008 to August 2014, consisting of monthly information on the boats departing Brazil's six main iron ore export terminals. The sample consists of 3954 boats that made 9862 separate port visits, with sizes ranging from 24,000 to 405,000 DWT and an average of 152,000 DWT. In several regression models, Adland et al. (2016) employed a variety of characteristics, including DWT, freight rate, bunker price, iron ore price, and distance. The empirical findings indicate that vessel-specific factors (DWT), with smaller boats often having lower capacity utilization, predominate the influence of general market circumstances (Adland et al., 2016).

Using a weekly dataset from October 2010 to August 2019 with shipping-related data obtained from Clarkson Shipping Intelligence Network, Michail & Melas (2021) examine and quantify the link between agricultural commodities and ocean-going freight rates. The dataset consists of trip charters and has been specified on bulk what falls into the grain category and looked specifically at trade routes for Handymax, Panamax and Supramax vessels. For this research, a Vector Error Correction Methodology is applied to the data. The empirical findings of Michail & Melas (2021) indicate that commodity prices may have a significant influence on freight rates across the majority of vessel classes. They also support the idea that vessel classes are closely interrelated in terms of freight rates, suggesting there could be a substitution effect (Michail & Melas, 2021).

It can be stated that multiple elements of freight characteristics could influence the freight rate. For that reason, it is interesting to investigate which elements in individual contracts have an effect on the freight rate. Nevertheless, before individual elements of contracts can be examined, the final element of contracts, the risk premium, will be addressed in the following section.

3.3.3. Risk premium

The risk premium, as used in the freight rate market, is the extra payment required by market participants to cover the risks involved in delivering goods. Numerous variables, such as supply and demand dynamics, fuel costs, geopolitical developments, climatic circumstances, and operational hazards, have an impact on freight rates. The risk premium is a measure of the extra expense or payment that carriers, shippers, or freight forwarders must provide in order to cover the uncertainties and potential dangers associated with moving goods (Adland & Cullinane, 2005).

Adland & Jia (2008) found evidence that the risk premium related to the possibility of a charter default is favourable and rising with spot freight rate level and period charter duration. The main use is to assess the financial risk and economic worth of time charters, for example, by lending institutions and credit rating agencies. Barrot et al. (2018) investigate how globalization is reflected in asset prices and how risk premium is connected to the cash-flow covary of investors. The results imply that foreign productivity shocks are linked to periods of high consumer prices for investors. Roels (2013) describes how the risk premium behaves when demand is modelled using an additive-multiplicative approach.

It demonstrates that the elasticities of the mean and standard deviation of demand at the riskless price can independently influence the sign of the risk premium. Chen (2011) states that an increase in the period time charter rate over the anticipated TCE spot freight rate from voyage chartering over the same time period is implied by a positive risk premium in the freight market, and vice versa.

The risk premium is not a fixed or standardized part of freight prices, which is crucial to keep in mind. It can change according to the state of the market, specific contracts, and how market players perceive risk. A risk-averse ship owner would most likely have invested in smaller vessels rather than larger ones when comparing the volatilities of different-sized vessels, as Kavussanos (1996) discovered that risk premiums were generally higher for larger vessels. This was caused by restrictions on the trades that larger vessels could engage in. It can be argued that the risk premium does not necessarily have a standard value, however, it indicates information about the perceived risk of the trip. Therefore, it can be concluded that risk premium could be considered an important element in contracts.

To summarize, the freight rate market and the corresponding elements are affected by supply and demand factors, freight prices, risk premiums, multiple macroeconomic factors, fleet availability, transport costs, and carrier rates. From this chapter, it can also be concluded that the elements in contracts containing time, geography and freight characteristics are all considered important in the dry bulk freight market. Therefore, it can be stated that the dry bulk freight market is not dependent on one variable but on many elements.

As mentioned earlier, the Baltic Dry Index could be used as an economic performance parameter in the dry bulk shipping market. For that reason, it is interesting to investigate which individual elements of contracts have an impact on the BDI and thus price dynamics. As the research would become too broad to include all types of vessels, this research will therefore look explicitly at individual spot rate contracts of Capesize vessels relative to the BDI. The next chapter discusses in detail the used methodology for this research.

4. Methodology

Having discussed the literature, it is necessary to dive into the methodology used to analyse the spot-rate contracts in order to answer the sub-questions and ultimately answer the research question. First of all, it is necessary to provide a brief outline of the existing spot-rate contract data. The data consists of spot-rate contracts for Capesize ships on certain dates, these contracts contain all the basic elements and elements such as time, geographic and freight elements, as mentioned in the literature. Most important to understand, is that (most) contracts contain the name of the vessel and a certain contracting date. A vessel may appear multiple times on various dates in the dataset and this can be recognized by the name of the ship. The appearance of vessels differs, as some vessels appear once in the data, while others appear nine times in the data. This indicates that the panel data is not homogeneous (or pooled), but heterogeneous. In other words, not all parameters are similar across different vessels, but it varies across different vessels and is therefore heterogeneous. This type of data is also referred to as pseudo-panel data or repeated cross-sectional data (Verbeek, 2008).

However, this research investigates the price dynamics relative to the elements of the spot-rate contracts, which indicates that not only the repeated cross-sectional data of the contracts is investigated. As mentioned earlier, the price dynamics in this research are examined using the Baltic Dry Index (BDI). The index data is published daily for weekdays all year round and for that reason is considered time-series data (Beck & Katz, 1995).

To summarize, the dataset has (repeated) cross-sectional aspects and time-series dimensions. Time-series data and cross-sectional data differ, in that time-series data focuses on a single variable across time, whereas cross-sectional data focuses on many variables at the same time. Basically, all elements that are in the contracts in the dataset belong to the cross-sectional classification and only the BDI, used for price dynamics, is covered by time series. A statistical method called cross-sectional time-series analysis combines the cross-sectional and time-series dimensions to analyse data. To take into consideration the heterogeneous pseudo-panel form of the data in this situation, the most frequently used panel models are the Fixed Effects and Random Effect models (Beck & Katz, 1995; Borenstein et al., 2010; Verbeek, 2008). The methodology is divided into three parts. The first part discusses the literature behind the methods, to investigate what methods researchers with similar research goals and datasets have used. The second part elaborates on the methodology used for each sub-question, this clarifies what methods and models are used for the sub-questions. The third and final part discusses the mathematical background of the used models.

4.1. Methodology Literature

Since it is evident that the dataset requires a cross-sectional time-series analysis, is it important to look at possible models. The models will be contrasted based on four factors: relevance, interpretability, accuracy, and computational time. To gain confidence and acceptance and reach useful results for the models, interpretability is crucial (Ribeiro et al., 2016). The models must thus be accurate in order to produce accurate predictions and provide insightful information since the trade-off between interpretability and accuracy is crucial (Lee & Shin, 2020).

The research of Da Silva (1975) focuses on estimating linear relationships in cross-sectional time-series data. Da Silva (1975) states that the variance component in linear relationship techniques does not take the potential serial correlation due to the time effects into account. This should be taken into account by using a certain first-order autoregressive error structure to treat cross-sectional unit heterogeneity effects. Two models are proposed to take the error into account: the Random Effects and the Fixed Effects Model (Da Silva, 1975). The Fixed Effects models are more commonly used in economics and political science because it is known as the 'gold standard' by default (Shurer & Yong, 2012). As Random Effects models are known as multilevel, hierarchical, linear and mixed models and have grown in popularity in economic science, political science and medical science (Bell & Jones, 2015). Researchers frequently assess the effectiveness and consistency of the Fixed and Random Effect models using model specification tests, such as the Hausman test, to decide which model is more appropriate (Ahn & Low, 1996).

Cross-sectional time analysis is a popular research method and is conducted to investigate many different topics. According to Shor et al. (2007) applied political science has increasingly used time-series cross-sectional data over the past 10 years. The research examines the use of a Bayesian MultiLevel model (BML), the Fixed Effects model and a regular OLS model on the time-series cross-sectional data consisting of 20 years. Shor et al. (2007) state that the BML model offers the most modelling freedom relative to the Fixed Effects model and the OLS for the intricate error patterns and contextual information unique to this type of data. It is concluded that the BLM model is performing best and is about 50% more efficient than the Fixed Effects model and nearly twice as efficient as the OLS model. The authors suggest that the choice of model is depending on the dataset and that the BLM model is most efficient for this dataset due to the fact that the data is mostly contextual (Shor et al., 2007).

Sarkar & Hong (2004) focus their research on the effective duration of callable corporate bonds. In order to calculate the effective length of callable corporate bonds, both default risk and call risk are included. The data consists of a monthly time series of long-term government bond yields and the price of long-term corporate bonds for the 36-month period from January 1994 to December 1996. As the data is primarily contextual, a cross-sectional regression analysis is employed. They conclude that the cross-sectional regression analysis performs pretty accurately to examine the model's empirical implications on the factors that affect effective length (Sarkar & Hong, 2004). This suggests a cross-sectional regression analysis works well in predicting which factors affect the effective length of contracts.

Adland et al. (2018) offer a model that takes into account owner-charterer match effects, charterer heterogeneity, and individual contract freight rate development. This research is conducted on 2863 VLCC Tanker contracts and 1789 contracts for Capesize ships ranging from 2011 to 2014. The authors develop an extension of the state-of-the-art models on microeconomic determinants of the freight rate to account for buyer/seller and relationship effects. They empirically investigate whether these effects have an impact on the freight rates for individual fixtures and whether there are differences between ship types. The Fixed Effect model is used for this research because the extra heterogeneity terms are linked with the exogenous factors added to the regression for the most optimal results. Adland et al. (2018) conclude that the qualities of charterers, owners, and their matches are also important microeconomic predictors of the freight rate level, albeit market circumstances and routes continue to be the most relevant factors. According to these authors, the Fixed Effects model can be used to investigate the individual contract elements for the VLCC Tanker and Capesize contracts and confirm that it could provide relevant insights for the freight rate (Adland et al., 2018).

The research of Gao et al. (2023) has been published recently, creating an extended space-time network that includes decisions about ship scheduling, routing, and sailing speed in order to optimize the planning horizons. The model is rather theoretical, as the dataset is rather small and thus not representative for the whole fleet. The data consists of 7 ports and a planning horizon of 10 days, two theoretical ships from the dry bulk shipping fleet are available at this time, and their starting positions, free time, and operating expenses are all indicated in the model. They also extend this for the theoretical model to 11 ships carrying only grain for 120 days. The authors conclude that the model is promising, as shown by the numerical findings, and can offer management of dry bulk shipping operations with important information (Gao et al., 2023). This shows that the planning horizon of ships is difficult to map. Looking at this retrospectively, it might be stated that the effects of fleet availability are also difficult to map.

The goal of the empirical research of Xu et al. (2011) is to ascertain how the fluctuating fleet size and the time-varying volatility of freight rates are related. The dataset consists of Capesize and Panamax time-charter data from January 1977 to October 2010 and January 1973 to October 2010. The monthly freight rate, fleet size, industrial production, and bunker price are the data sets used in the research. One-step forward conditional volatility estimates of freight rates are created using the AR-GARCH model to analyse the link between freight market risk and the change in fleet size. The findings show that the size of the fleet positively influences freight rate volatility, but that the spot rate volatility of Capesize dry bulk responds more strongly to the size of the fleet (Xu et al., 2011). From this, it can be concluded that the spot rate volatility for Capesize ships reacts strongly to the size of the fleet, according to the AR-GARCH model.

Examining other research that has a similar type of dataset or a similar research angle, it is evident that the Fixed Effects and Random Effects models always emerge very strongly for determining price dynamics. Especially the research of Adland et al. (2016) uses literally the same type of data for Capesize ships, these researchers suggest that the Fixed Effects model is very appropriate for this research. Not much research has been done about the impact of fleet availability on price dynamics. However, the research of Xu et al. (2011) showed and concluded that the AR-GARCH model provides evident results. For this part of the research, an alternative manner could be devised. In the next sub-chapter, the most appropriate models are discussed for this research per sub-question in order to answer in the end the research question.

4.2. Chosen method per sub-question

For the first sub-question: *Can insights into time elements of contracts provide information about price dynamics in dry bulk freight rates?*, the Baltic Dry Index (BDI) is used to investigate the effects of time elements in spot-rate contracts in the price dynamics in bulk freight rates. Then, it is necessary to determine which model is most appropriate to use to predict these price dynamics and which variables must be used. As indicated earlier and suggested by the literature, Fixed Effects models and Random Effects models are well suited for this type of spot-rate contract data and eventually to determine the price dynamics and for that reason, these models are used to answer this sub-question. As the last step when both models are created with the desired dependent and independent variables, a Hausman test is done to determine which model is most appropriate to investigate the element time on price dynamics. The effect on the price dynamics is in both models the examined variable and for that reason, the BDI is used in both models as a dependent variable (Verbeek, 2008).

To understand which time elements of contracts impact price dynamics, various contract elements are used as independent variables. First, is the number of days a vessel is in laycan indicated in the models, which investigates the laycan duration on the price dynamics. Second, is the number of days between the contract date and the laycan from date considered. This investigates if the period between the contract date and the vessel going into laycan has an influence on the price dynamics and shows the effect of contracting at shorter or longer notice from the laycan date. Third, is the season of the laycan from date considered in the models to investigate if there is any seasonal effect on the price dynamics of the laycan periods. Fourth, is the freight rate included in the models as an independent variable and is used as an interaction term. This shows the effects of the freight rate on the time elements from contracts used in the models and the effects on the price dynamics. Eventually, the effect of the freight rate can be used as an interaction term to make a comparison of the effect relative to other models from the other sub-questions. This results in better conclusions about the price dynamics of different elements investigated by different sub-questions.

The second sub-question: *What are the effects of the geographical elements of shipping activities on the price dynamics?*, uses the BDI to investigate the effects of the geographical elements in spot-rate contracts on price dynamics. As the contracting data and the research angle is similar to the first sub-question, it can be concluded that Fixed Effects models and Random Effects models are most appropriate in answering this sub-question too. For that reason, these models are used and both using the BDI as the dependent variable to investigate the effects of the independent variables on the price dynamics. If both models with all desired variables are created, a Hausman test is done to determine which is most appropriate for this sub-question (Verbeek, 2008).

This sub-question examines which geographic elements impact price dynamics and therefore uses the following elements as independent variables. First, the load and discharge locations of the vessels are used to investigate the effect of different locations on the price dynamics and possibly reveal certain patterns in the shipping routes. Second, to dive more into the examination of the shipping routes and their effects on the price dynamics, the number of days the vessel is sailing from load to discharge location is used in the models. Third, is the freight rate, as with the previous sub-question, included in the models as an independent variable and used as an interaction term to see the effects on the geographical elements and on the price dynamics. The freight rate is also used as an interaction term, as in the previous model, to make comparisons between other investigated elements from the other sub-questions.

The third sub-question, *How do freight elements of dry bulk influence the price dynamics?*, takes the same approach as the previous two sub-questions, as the spot-rate contract data and the price dynamics are similar. Therefore, a Fixed Effects model and Random Effects model are used to investigate the influence of various freight elements on the price dynamics. As both models are created with the BDI to examine the price dynamics as the dependent variable and with freight elements as independent variables, a Hausman test is performed to determine which model is most appropriate for this sub-question (Verbeek, 2008).

Freight elements of contracts and their effects on the price dynamics are investigated by various independent variables. First, is the type of bulk the vessel is transporting, as it is important for determining the freight rate and thus has an impact on the price dynamics. Second, as spot-rate contracts negotiate freight rates per tonne of dry bulk transported, the quantity the vessel is transporting is necessary to take into account as an independent variable. Third, as the quantity is paid per tonne transported dry bulk, it is necessary to take the freight rate into account as an independent variable in this model. This shows the impact of the freight rate on price dynamics. Also, it is used as an interaction term to investigate the effect of freight elements on the freight rate and compare it to the models from the other sub-questions.

The fourth and final sub-question: *What are the effects of fleet availability on price dynamics in the market?*, takes a different approach than the other sub-questions. As has already become evident from sparse existing literature, there is no standard way to examine fleet availability and price dynamics. For that reason, it is decided to investigate the effects of fleet availability step by step on price dynamics. Similar to the previous sub-questions, the BDI is used to determine the price dynamics for this sub-question. As the sub-question is different from the other sub-questions, this sub-question is answered in multiple steps.

As the first step, the impact of fleet availability on the BDI is examined graphically, to determine if there is any effect of the fleet availability on the price dynamics that can be observed from the raw data. In order to have insight into fleet availability, it is necessary to determine how many contracts are occupied on a certain day. For that reason, the spot-rate contract data must be transformed into a day-to-day timeline to determine the number of occupied ships per day. This makes it possible to make a scatterplot, to visualize the effect of fleet availability on price dynamics.

As the second step, in order to answer the sub-question appropriately, it is necessary to investigate the spot-rate contract data more deeply and start with modeling. For that reason, it is decided to investigate as an intermediate step, the element fleet availability first. Therefore, the number of occupied ships obtained in the day-to-day timeline is merged back into the spot-rate contract data. That way, the individual contract data have an indicator of how many ships are occupied at the contract date and thus the fleet availability. To examine the fleet availability more deeply, it is reasonable to examine the origin of fleet occupation. Therefore, it is useful to adopt a model that is capable of showing which variables impact fleet occupation. In order to investigate which variables are impactful in predicting the fleet occupation and thus the fleet availability, a Random Forest model is used.

The Random Forest model is an ensemble learning method used in machine learning, which grows and combines multiple decision trees and combines these trees in a forest. The Random Forest model is flexible in terms of usage, as it can be used for both classification and regression tasks. After the Random Forest is created, the model can be used to determine which variables have the most predictive power and are considered important in terms of predicting the dependent variable. The relevance of each variable implemented in the created Random Forest model may be assessed by computing the variable importance. This can be divided into the percentage increase in means squared error and the increase in node impurity, after receiving the Random Forest results. This gives a broad understanding of how each variable affects the model's ability to forecast the dependent variable and takes into account all interactions by creating the decision trees (Biau, 2012; Stob et al., 2007).

In this research, as the Random Forest is in the first instance used to predict the number of ships occupied (the fleet availability), it is used as a regression task. However, as the Random Forest model predicts the ship's occupation, the outcome itself is not relevant for this research, but the variable importance feature, which can be used after creating the model, is relevant for this research. The variables which are considered most important, according to the variable importance feature of the Random Forest, in predicting the fleet availability are used in the next step in answering the sub-question.

As the impact of fleet availability on the raw data is investigated and it is examined which variables are important determinants of fleet availability. It is necessary as the third and last step to investigate the effect of fleet availability on price dynamics. Since the spot-rate contract data is similar to the previous sub-questions and it is clear which variables affect fleet availability. It is decided for this research to keep the modeling approach identical to the previous sub-questions. For that reason, the final step of this sub-question uses a Fixed Effects model and Random Effects model to investigate the effects of fleet availability on price dynamics (Verbeek, 2008).

The BDI is used in both models as the dependent variable to investigate the effect on the price dynamics. In both models, the number of occupied ships is used as an independent variable in combination with the most crucial variables, according to the Random Forests. The variables obtained by the Random Forests are also used as interaction terms, to see their next to the effect on the price dynamics, also the effect on the ship's occupation. Finally, a Hausman test is done to determine which model is most appropriate to investigate the ships occupied on price dynamics. That way, the impact of the number of ships occupied with the other suggested variables can evaluate the effect of fleet availability on price dynamics. As the methods for all sub-questions are discussed, it is necessary to dive into the mathematical background of the models in the next sub-chapter.

4.3. Mathematical background

This sub-chapter elaborates on the mathematical background of the models used. First, the mathematics of the Fixed Effects model are explained, followed by the Random Effects model. Next, the math behind the Hausman test, used to determine the model choice, is discussed. At last, the mathematics behind the Random Forest model is reviewed.

4.3.1. The Fixed Effects model

The main idea of the Fixed Effects model is to time demean the data to eliminate the unit heterogeneity. The time demean is obtained by calculating the differences by subtracting the original model from the unit mean averages. The individual-specific effects are explicitly incorporated in a fixed effect model to represent the heterogeneity among the panel's entities. Dummy variables or fixed effects coefficients serve as a representation of these fixed effects. The model accounts for unobserved heterogeneity that changes among entities but is stable across time by introducing fixed effects (Da Silva, 1975; Wooldridge, 2015). This can be expressed in mathematical terms, consider a model with one explanatory variable:

$$y_{it} = \beta_1 x_{it} + \alpha_i + \upsilon_{it} \quad t = 1, 2, \dots, T$$

y_{it}= the dependent variable for the ith observation at time t
x_{it}= the vector of the independent variable for the ith observation at time t
β_1= the vector coefficient to be estimated
α_i = the unit heterogeneity, which represents the fixed effect for the ith observation
υ_{it}= the error term
t = the time period

Fortake for each i the time average for this equation:

$$\underline{y}_{it} = \beta_1 \underline{x}_{it} + \alpha_i + \underline{\upsilon}_{it}$$

Because the unit heterogeneity (α_i) is fixed over time, it is also in the time average equation. If the first equation is subtracted by the time average equation for each time period (t), this results in the following formula:

$$y_{it} - \underline{y}_i = \beta_1 (x_{it} - \underline{x}_i) + \upsilon_{it} - \underline{\upsilon}_i$$

As can be seen, the unobserved effect of the unit heterogeneity (α_i) disappeared from the equation. The time demeaned data is within the $y_{it} - \underline{y}_i$, $x_{it} - \underline{x}_i$, $\upsilon_{it} - \underline{\upsilon}_i$. For that reason, the fixed effects transformation is also called the within transformation, because it uses the time variation in y and x within each cross-sectional observation.

The objective of this model's estimation is to find the values of the coefficient β_1, which, after taking into account the fixed effects and the error term, shed light on the connections between the independent and dependent variables. The parameter estimates in this model may be obtained using a variety of estimation strategies, such as fixed effects regression or least squares estimation (Wooldridge, 2015).

4.3.2. The Random Effects model

The main idea of the Random Effects model is that it uses a quasi-demeaning transformation to eliminate serial correlation. The individual-specific effects are viewed in a Random Effect model as random variables that are presumptively unrelated to the independent variables. The conventional assumption is that these random effects will follow a certain distribution, such as the normal distribution. The model calculates the variance of the random effects as well as the fixed effects (Da Silva, 1975; Wooldridge, 2015). This can be expressed in mathematical terms, consider a model with one explanatory variable:

$$y_{it} = \beta_0 + \beta_1 x_{itj} + \alpha_i + \upsilon_{it} \qquad t = 1, 2, \ldots, T$$

y_{it}= the dependent variable for the ith observation at time t
β_0= the intercept coefficient
x_{itj}= the vector of the independent variable for the ith observation of the jth group at time t
β_1= the vector coefficient to be estimated
α_i = the unit heterogeneity, which represents the fixed effect for the ith observation
υ_{it}= the error term
t = the time period

The difference between of this equation and the Fixed Effects model is the intercept β_0. This intercept is included, such that (without losing generality) it can be assumed that the unobserved impact of α_i has zero mean. Normally, we would also include time dummies among the explanatory factors. Because it is believed that α_i is linked with one or more of the x_{itj}, eliminating α_i via fixed effects or first differencing is the objective. However, let's say it is believed that α_i is not associated with any explanatory variable over all time periods. Ineffective estimators are produced as a result of employing a transformation to remove α_i. It becomes a Random Effects model when the unobserved effect of α_i is assumed to be uncorrelated with each explanatory variable:

$$Cov(x_{itj}, \alpha_i) = 0\,, \; t = 1, 2, \ldots, T;\, j = 1, 2, \ldots, k$$

Actually, the optimal random effects assumptions demand that α_i be independent of all explanatory factors over all time periods in addition to all of the fixed effects assumptions. If the composite error term is rewritten into: $v_{it} = \alpha_i + \upsilon_{it}$, then the equation becomes:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \ldots + \beta_k x_{itk} + v_{it}$$

The v_{it} are serially correlated throughout time because α_i in the composite error for each time period. In actuality, under the following assumption of random effects:

$$Cov(v_{it}, v_{is}) = \sigma_\alpha^2 / (\sigma_\alpha^2 + \sigma_u^2) \qquad t \neq s$$

of which $\sigma_\alpha^2 = Var(\alpha_i)$ and $\sigma_u^2 = Var(\upsilon_{it})$. The Ordinary Least Squares models neglect the standard errors correlation and therefore will provide inaccurate estimates. This consists mainly of positive serial correlation in the error term. The serial correlation in the errors can be removed by using the GLS transformation. The transformation itself is straightforward, giving a value between 0 and 1 for the Θ results in the following equation:

$$\Theta = 1 - [\sigma_u^2/(\sigma_u^2 + T\sigma_u^2)]^{1/2}$$

This can be transformed to the following equation:

$$y_{it} - \Theta\ \underline{y}_i = \beta_0(1-\Theta) + \beta_1(x_{it1} - \Theta\ \underline{x}_{i1}) + \ldots + \beta_k(x_{itk} - \Theta\ \underline{x}_{ik}) + (v_{it} - \Theta\ \underline{v}_i)$$

The equation involves the quasi-demeaned data on each variable and the overbar presents the time averages. The temporal averages are subtracted from the appropriate variable using the fixed effects estimator. A portion of the time average is subtracted by the random effects transformation; the portion depends on $\sigma_\alpha^2, \sigma_u^2$ and the number of time periods, T. The biggest advantage of the Random Effects model over the Fixed Effects model is the transformation's ability to accommodate explanatory factors that remain constant throughout time. This is doable because, whether or not the explanatory factors are stable across time, the Random Effects model assumes that the unobserved impact is uncorrelated with all explanatory variables.

In summary, whenever the random effects are uncorrelated with the explanatory variables, the Random Effects model is more flexible in capturing individual-specific effects and can produce more accurate estimates than the Fixed Effects model. It does, however, make the assumption that there is no correlation between the random effects and the explanatory factors, which may not always be true. The fixed effects model may be more suitable under certain circumstances. The choice between the Fixed Effects and Random Effects model depends on various factors and can be tested. The next section is devoted to this matter (Wooldridge, 2015).

4.3.3. Fixed vs Random Effects models

The determination between the Fixed Effects or Random Effects model depends on the kind of data and research. Whenever there is a concern for unobserved heterogeneity that is unique to each entity in the data, the fixed effect model is more appropriate. The model accounts for time-invariant unobserved elements by introducing fixed effects. The Random Effect Model, on the other hand, makes the assumption that the effects that are unique to each person are random and unrelated to the independent factors. In the case that there is no concern about time-invariant unobserved components, the Random Effects model is more appropriate (Da Silva, 1975; Wooldridge, 2015).

The Fixed Effects model makes the assumption that each individual's individual-specific effects, represented by the symbol, are time-invariant and distinct. Individual-specific dummy variables or fixed effects are incorporated into the regression model to reflect these impacts. Because the fixed effects in this situation already account for the average effect for each individual, a constant term is not added.

On the other hand, the Random Effects model makes the assumption that the effects that are unique to each person are random and connected with the independent factors. These impacts are seen as random variables and included in the model through a disturbance term, which is often predicated on a set of distributional assumptions. In this instance, the average effect across all people is represented by a constant component in the model. The estimate of the overall intercept, which reflects the average impact across all people, is made possible by the inclusion of a constant term in the Random Effects model. Individual-specific effects then capture variations from this average effect (Wooldridge, 2015).

Researchers frequently assess the effectiveness and consistency of the Fixed and Random Effect models using model specification tests, such as the Hausman test, to decide which model is more appropriate. Hausman (1978) developed the Hausman test and several econometric software programmes in order to perform the test. The test takes all the random effect assumptions into account:

$$Cov(x_{itj}, \alpha_i) = 0, \; t = 1,2,\ldots,T; j = 1,2,\ldots,k$$

Unless the Hausman test finds a negative result, the Random Effects model needs to be used. Both models require the idiosyncratic shock to be unrelated to the variable of interest to be unbiased. The Random Effects model, also requires the individual heterogeneity to be uncorrelated with the variable of interest to obtain unbiased estimates. The null hypothesis state that there are no systematic differences between the Random Effects coefficients and the Fixed Effect coefficients. For that reason, if the null hypothesis is not rejected, the Random Effects model is chosen because it is more efficient. The alternative hypothesis state that there are systematic differences between the Random Effects model and the Fixed Effects model and therefore the Fixed Effects model is preferred. The Random Effects model would be biased in that case (Hausman, 1978; Wooldridge, 2015).

4.3.4. The Random Forest model

As stated previously, the Random Forest model is flexible enough to handle both classification and regression tasks, however, this research uses the model as a regression task. In short, the Random Forest model consists of a combination of multiple decision trees, which are regression trees in the case of a regression task. First the math behind the Random Forest model is discussed and it is followed by the mathematical background for the variable importance.

As a first step, a group of randomized base regression trees make up the overall framework, which is a nonparametric regression estimation $\{ r_n(x, \Theta_m, D_n), m \geq 1\}$, where the $\Theta_1, \Theta_2, \ldots$ are outputs of a randomizing variable Θ. These trees are combined to create the aggregate regression estimate:

$$r_n(X, D_n) = \mathbb{E}_\Theta[r_n(X, \Theta, D_n)]$$

in which $\mathbb{E}_{\Theta}$ is standing the expectation to the random parameter conditionally on X and the dataset D_n. The coordination and the points at which the trees are divided are determined while creating individual decision trees using the randomizing variable Θ. This result in the following expectation:

$$\mathbb{E}_{\Theta}[r_n(X, \Theta, D_n)]$$

Then, this expectation is evaluated by Monte Carlo approximation. This method for estimating variables that depend on one or more random factors, by creating m (often large) random trees. After which, it takes the mean of each tree's final individual outcomes:

$$r_n(X, D_n) = \mathbb{E}_{\Theta}[r_n(X, \Theta_m, D_n)]$$

The sum of the total of each tree's individual $r_n(X, \Theta)$ is taken by the Random Forest regression, which gives the average overall Y_i as output with the corresponding vectors X_i. The corresponding vectors X_i fall in the same cell of the random partition as X. This results in the following Random Forest regression:

$$r_n(X) = \mathbb{E}_{\Theta}[r_n(X, \Theta)] = \mathbb{E}_{\Theta}\left[\frac{\Sigma^n{}_{i=1} Y_i 1[X_i \epsilon A_n(X, \Theta)]}{\Sigma^n{}_{i=1} 1[X_i \epsilon A_n(X, \Theta)]} \; 1_{E_n(X,\Theta)}\right]$$

The rectangular cell $A_n(X, \Theta)$ of the random partition, contains X and the event $E_n(X, \Theta)$ is defined by:

$$E_n(X, \Theta) = \Sigma^n{}_{i=1} 1[X_i \epsilon A_n(X, \Theta) \neq 0]$$

According to Biau (2012), this suggests that the estimate for empty cells is set to 0. Most Random Forests algorithm use by default 500 decision trees, keeping certain computation time of the algorithm in mind. It can be decided to adjust this manually by any number or calculate the optimal number of decision trees. The algorithm calculates optimal number variables used at each split in the trees and gives the percentage variance explained by the total model. The variance explained is an indicator of how effectively the model reflects the input characteristics' influence on the variability of the target variable. In conclusion, a Random Forest predictor uses many randomized base regression trees to provide a model with good interpretability and predictive power (Biau, 2012; Peters, 2022).

As second step, the relevance of each variable in the dataset with regard to the created model may be assessed by computing the variable importance. This can be divided into the percentage increase in means squared error and the increase in node impurity, after receiving the Random Forest results. This gives a broad understanding of how each variable affects the model's ability to forecast the future and takes into account all interactions (Strobl et al., 2007).

The variable importance of the Random Forest model is used in this research. The variable importance can be expressed in two ways: the 'percentage increase in means squared error' (%incMSE) and the 'increase in node impurity' (IncNodePurity). First, in order to get the percentage increase in means squared error, the Mean Squared Error (MSE) is calculated for the original data:

$$MSE_{original} = \frac{1}{n} \Sigma^{n}{}_{i=1} (y_i + \hat{y}_i)^2$$

The number of data points is denoted by n, the actual value of the target variable for the ith point by y_i and $\hat{y}_i$ the predicted value for the ith point of the target variable. Then the target variable is randomly permuted (shuffled) while the other variables remain the same. As next step, the MSE is calculated on the same way for the permuted data:

$$MSE_{permuted} = \frac{1}{n} \Sigma^{n}{}_{i=1} \left(y_i + \hat{y}_{permuted,i}\right)^2$$

At last, the Percentage Increase in Mean Squared Error statistic can be calculated by quantifying the increase in the mean squared error (MSE) of the model:

$$\%inc\mathrm{MSE} = \frac{MSE_{permuted} - MSE_{original}}{MSE_{original}} \; x \; 100\%$$

The importance of the variable to the model increases with the size of the MSE rise, suggesting the variable is important in predicant for the dependent variable and thus for the model (Molnar, 2020; Peters, 2022).

Second, in constructing the individual decision trees in the Random Forest model, the increase in node purity assesses the rise in node impurity brought by a particular variable. It evaluates how well the variable can divide the data and produce reliable predictions. The increase in node impurity caused by each variable in a decision tree is calculated as:

$$IncNodePurity_{tree} = NI_{split} - NI_{orignal}$$

The node impurity of a tree before splitting on the variable of interest is $NI_{orignal}$ and the node impurity after splitting is NI_{split}. Eventually, the total number of trees is expressed by N_{total} and the average increase in node impurity is expressed by AI_{node}. Then the model calculates the average increase in node impurity across all trees in the Random Forest in the following way:

$$AI_{node} = \frac{1}{N_{total}} \sum_{tree=1}^{N_{total}} IncNodePurity_{tree}$$

A higher score AI_{node} implies the variable is more important in the decision-making process in the decision trees and thus an important predictor for the model (Molnar, 2020; Peters, 2022). Having explained all the methods and models, the underlying mathematics and reasoning. It is necessary to investigate these data in the following chapter.

5. Data

This cross-section analysis explicitly investigates individual elements of individual Capesize contracts and the effect on to the Baltic Dry Index. Prior to explaining the source of the data as well as the steps taken to arrive at the final dataset that is used to answer the research question and sub-questions with appropriate models. It is first necessary to elaborate on the background of the collected data. This data is obtained by the Clarksons Shipping Intelligence Network, which is part of the enterprise Clarkson PLC. The corporate office is located in London and Clarksons is a well-known supplier of integrated shipping services and solutions. It is among the biggest shipbroking and integrated shipping services businesses on the planet. Ship brokerage, chartering, research and consultancy, financial services, and risk management are just a few of the services provided by Clarksons. To aid in decision-making in the shipping and maritime industries, they provide their clients with market knowledge and analysis (Clarksons, 2023).

The individual contracts examined in this research are exported from the Clarksons Shipping Intelligence Network by Albert Veenstra (Clarksons, 2023). The day-to-day data for the Baltic Dry Index is exported from the platform Trading Economics (Trading Economics, 2023). The platform of Trading Economics provides accurate information including historical data and forecasts for over 20 million economic indicators, exchange rates, stock markets and economic indexes for 196 countries (Trading Economics, 2023). As the background of the obtained data is discussed, the next sub-chapter dives into the raw obtained data.

5.1. The raw data

The raw spot-rate contract data for Capesize ships consists of one Excel export from the Clarksons Shipping Intelligence Network per calendar year. Unfortunately, not all historical data can be exported, only the years 2020, 2021 and 2022 were available on the platform. For that reason, contracts from those years will be included in the research. Every observation, which consists of one row in the dataset, is a bulk spot contract for a Capesize ship. The year 2020 consists of 1381 contracts, the year 2021 consists of 1479 contracts and the year 2022 consists of 1328 contracts. All the datasets consist of 16 columns, which imply every contract consists of 16 variables. The following variables (columns) are stated in the raw data output:

Table 3. Column description raw Clarkson Data

Variable name	Description
Date	The date of the contract.
Name	The name of the Capesize ship, TBN (To Be Announced) if the name is not indicated by Clarkson.
Built	The year the ship is built.
DWT	The Death Weight Tonnes of the ship.
Quantity	The quantity transported of the dry bulk.
Type	The type of dry bulk transported.
Charterer	The chartering company.
Laycan From	The date when the laycan period starts.
Laycan To	The date when the laycan period ends.
Delivery	The location where the delivery is dropped.
Load	The location (port) where the ship is loaded.
Discharge	The location (port) where the ship is discharged.
Redel	The location (port) of the redel.
Rate	The rate of the freight is in US dollars, depending on the unit.
Unit	The unit the rate is expressed in, $/Tonne and $/Day, which expresses the rate per tonne (quantity) transported of the rate per day the vessel is hired.
Owner	The owner of the ship.

Source: Clarkson 2023

It is good to mention in advance that although some variables are included by default in the format by Clarkson, in reality, they are almost never filled in. The variables Delivery, Redel and Owner have mostly empty observations, but none of these variables are relevant to this research. The variables Name, Built and DWT are not filled in regularly and thus have many missing values. It is a major pity that the Name and DWT were not consistently filled in for the research, as empty names are filled in by default as To Be Nominated (TBN). This makes it more difficult to distinguish different boats listed in the data as TBN. However, prior to examining the data preparations, the Baltic Dry Index data will be examined first.

The Baltic Dry Index data retrieved from Trading Economics consists of day-to-day prices ranging from January 2020 to December 2022. More data could be obtained, but since the Spot Bulk contracts run to this period, the same period was obtained. The data consists of 744 observations, which indicates that there are 744 days in the dataset. This is due to the fact that the financial market closes at weekends and therefore no Baltic Dry Index is published on weekend days. Each row consists of five variables, the price (BDI), the opening price, the highest price, the lowest price and the percentage change from the previous day (Trading Economics, 2023). Having clarified what the raw datasets look like, the next sub-chapter will explain how the data preparation was done in R-Studio to obtain the final dataset with all the necessary variables to make the appropriate models for the research question and sub-questions.

5.2. The data preparation and manipulation

In order to obtain the final dataset with the necessary variables for modelling, it is necessary to make one dataset containing all the necessary information. The first step in data preparations starts with loading raw data into R-Studio from bulk spot-rate contracts of Capesize ships from the years 2020, 2021 and 2022. Then, before making it into one dataset, it is necessary to investigate and check if all three raw bulk spot-rate datasets for Capesize ships are formatted in the same way. It turns out that the columns contain the exact same information but in a different order for the year 2022. After putting them into the same format, the three raw datasets of combined into one bulk spot dataset with 4188 observations.

As the next step, the Baltic Dry Index data is loaded into R-Studio. For this research, the opening, closing, highest and lowest prices of the BDI are not relevant. The price of the BDI column is the most relevant for this research and the percentage change could be interesting to take into account. For that reason, these two columns are matched to the combined dataset consisting of the three years bulk spot contracts of Capesize ships, in order to create one dataset containing all information. The BDI and the percentage change are matched on the contract date, which is the Date column of the bulk spot dataset. As a result, the total basic dataset consists of 18 variables, but some of these elements need to be modified and added to achieve the final dataset.

5.2.1. Preparing time-related variables

First, preparations necessary for the element time are discussed. There are three variables which are linked to the element time: Date, Laycan From and Laycan To. The Date variable is considered as the start of the contract and the Laycan From date as the start of the laycan period. To indicate the days between the contract date and the start of the laycan period, the variable Days Contract Laycan From is created. By calculating how many days the laycan period lasts, the new variable Days Laycan is created and is expressed in full days. Also, diving further into the laycan dates, there is a variable created which indicates the date between the Laycan From and Laycan To implemented as the Start date. In addition, there are categorical variables indicating the seasons added for the Contract date, Laycan From date, Laycan To date and Start date. Finally, two binary variables are created that indicate whether there is a seasonal difference between the Contract date and the Laycan From date and the Laycan From date and Laycan To date. These data modifications created nine variables, making the dataset consisting of 27 variables.

5.2.2. Preparing geographically related variables

Subsequently, the preparations related to the element of geography are covered. For that, all load and discharge locations of the contracts are investigated and the data showed that there are over 146 different locations within the dataset. Geographically, load and discharge locations can be used to determine how long a dry bulk Capesize ship sails over the route. However, how long it takes for Capesize ships to sail from A to B cannot be found in any open-source dataset on the internet. These datasets can be purchased from private companies for exceptionally large sums of money. However, this is not in the budget for this research. For that reason, it is decided in cooperation with Albert Veenstra that all routes need to be searched in a Maritime Distance Calculator. The routes needed to be filled in by hand using a Maritime Distance Calculator to determine how long it takes for a Capesize ship to sail from load to discharge location.

For that reason, it is decided in cooperation with Albert Veenstra that these locations needed a categorization to make it possible to search it by hand. Albert Veenstra suggested the following 13 categories: North America West Coast, North America East Coast, Caribbean (inc. US Gulf), South America West, South America East, ARA range, North Europe, Mediterranean, Africa, Middle East, Indian Subcontinent, Far East, South Pacific. All load and discharge locations are categorized in these 13 categories in the new variable Load category and Discharge category, adding two variables to the dataset making a total of 29 variables. As these variables are added, it is possible to investigate the sailing time of the routes.

5.2.3. Preparing shipping time-related variables

There are a few things important to note before diving into the output of the Maritime Distance Calculator. First of all, the data consists of Capesize ships, which means that these boats do not pass through the Panama or Suez Canal, as mentioned in the literature. In determining the sailing time, the speed at which the Capesize ships sail is a very important factor, as it significantly affects this variable.

In the Maritime Distance Calculator of Searates (Searates, 2023), a speed must be entered to calculate transit time. The Maritime Page (2023) indicates an average speed of Capesize ships of 13.2 knots, while S&P Global (2021) suggests that the average speed laden is 12 knots and with ballast 13 knots.

The research of Cepowski (2019) states that the average speed for Capesize ships is 14.47 knots. Adland & Jia (2017) state that the average speed for Capesize ships is reduced from 14.5 knots to 12 knots, due to a significant decrease in fuel consumption. It can be concluded that there are many different sources of sailing speed. All sources suggest fairly similar speeds but do not indicate exactly the same speed. Therefore, this study assumes a speed of 13 knots as the sailing speed for the Capesize ships.

As the sailing speed is decided, it has to be decided which routes should be searched by hand in the Maritime Distance Calculator. In the end, it is decided to search a route by hand if at least 10 boats have sailed it. This was broken down into two parts, first, the load and discharge ports that had more than 10 ships sailing are investigated. Then the remaining contracts by load category and discharge category, which routes had been sailed by more than 10 ships sailing are investigated. A total of 26 routes have more than 10 vessels from load and discharge ports and 23 routes have more than 10 vessels from load category to discharge category. For the load category and discharge category, a port had to be chosen to be the departure and arrival port. For this purpose, the most centrally located port for each category was often chosen. These 49 routes were searched by hand in the Maritime Distance Calculator of Searates (Searates, 2023). The shipping routes and the number of occurrences can be found in Appendix B.

Searates' Maritime Distance Calculator calculates distance and transit times expressed in days and hours. The calculator does not take into account the average days a ship is in port to load and unload. The United Nations Conference on Trade And Development (UNCTAD) has been researching port times for years and states that on average Dry Bulk carriers spend between 2020 and 2023 spend 2.07 to 2.18 days in ports and Capesize ships on average 2.3 days (UNCTAD, 2022). Therefore, in addition to the calculated transit time from the Maritime Distance Calculator, 2.3 days are added. The total hours were converted into days and are included as days in the dataset as the variable Days Shipping. Using the approach (suggested by Albert Veenstra) of taking the most common routes for load, load category, discharge and discharge category, there are a total of 359 routes without Days Shipping. This is a limitation, but fortunately, it totals 8.7% of the total dataset. These 49 routes, ports and shipping times can be found in Appendix B.

Since the shipping time is known, it is possible to set an end date for a ship. Based on this the end date is calculated, adding the shipping days to the Start date. As previously stated, the middle date between the laycan to and laycan from date is used as the Start Date. Along with the variable Shipping Days, the variable End Date was also added to the dataset, making a total of 31 variables in the dataset. In summary, the data preparations and modifications increased the number of variables from 18 basic variables with 13 informative variables for the research, reaching a total of 31 variables.

5.3. The final dataset

The final dataset consists of data from January 2020 to December 2022 and contains 4188 observations and 31 variables. This suggests that there are a number of elements available for each contract to investigate the price dynamics. For that reason, the content of the dataset needs to be examined first. The most important is to check if all the rows have a value for the variable Price, which is the BDI. Otherwise, the price dynamics are not possible to investigate for the sub-questions and the research question. In total 50 contracts have a date during a weekend day and thus have no value for Price (BDI). These contracts are omitted for this research and reduce the observations to 4138. The data consists of date variables, numerical variables and categorical variables. These variables all describe certain elements of the data, such as the time elements, geographical elements and ship characteristics. Since price dynamics and economic elements such as the freight rate are important for all sub-questions, this is examined first and adjusted if necessary.

5.3.1. Date elements of the dataset

The date variables are first inspected. The variables Laycan From and Laycan To have both minimum values starting in 2002, which is fairly unexpected. After some investigation, it is decided that these dates are typos and these observations are therefore deleted from the dataset. It is also conspicuous that 229 contracts have negative values as Days Laycan, this suggests the Laycan To date is previous to the Laycan From date, therefore these contracts are deleted from the dataset. After these computations, 3907 observations are left in the dataset. The table displays the summary statistics of the date variables:

Table 4. Summary statistics date variables

Variable	N = 3907	Mean	Min	Max
Date	3907	28-06-2021	02-01-2020	30-12-2022
Laycan From	3847	15-07-2021	28-12-2019	01-04-2023
Laycan To	3907	20-07-2021	28-12-2019	01-02-2029
Start Date	3907	17-07-2021	28-12-2019	08-02-2025
End Date	3563	29-07-2021	11-01-2020	19-03-2025

As can be seen, there are some dates missing in the dataset. The variable Date is self-explanatory and contains no outliers. Inspecting all mean, min and max dates, there seem to be no more outliers in the variables. Looking at the number of variables in the dataset, especially for the End Date there are 344 missing values, however, this was expected due to the way of calculating the routes by hand as discussed in the previous section. The next section investigates the price dynamics and economic elements, as the outliers and missing values need to be deleted first in order to make the models work appropriately.

5.3.2. Price dynamics and economic elements of the dataset

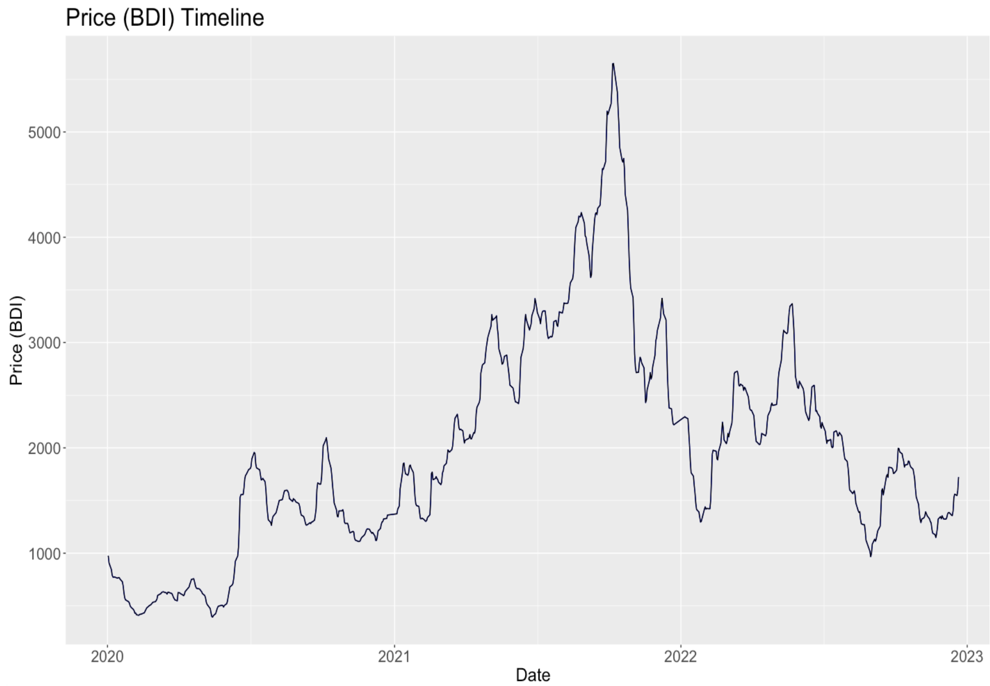
In order to observe price dynamics, the variable Price, which represents the Baltic Dry Index for that contract day and other economic variables such as Quantity, Unit, Rate and Percentage Price Change are investigated. First, it is important to note that there are 3716 contracts on a $/Tonne basis and 181 contracts on a $/Day basis and for the remaining 205 contracts nothing is specified. Furthermore, the difference between the two is not important for this study, only for the Rate is there a major difference. Therefore, the Rate is included for both contract forms in the data inspection. Thereby, it is useful to notice that the quantity is filled only for the Tonne contracts and not for all Day contracts.

Table 5. Summary statistics economic elements of the dataset

Variable	N = 3907	Mean	Min	Max
Price (BDI)	3907	2.006	393	5650
% Price Change	3907	2.966	0	22.550
Quantity	3907	166,439	75,000	750,000
Rate $/Tonne	3716	13.692	1.500	63.500
Rate $/Day	181	27,559	1,800	150,000

First of all, the quantity was examined and about that it can be said that Capesize ships transport between 75,000 to 750,000 and on average 166,439 per trip. This suggests there is an outlier, as the maximum value of the quantity cannot be 750,000. In terms of Price dynamics, it can be seen that the average Baltic Dry Index is 2,006 for the three years of data in the dataset. However, it appears to have been fairly volatile during the years as the min and max values are pretty far apart. As can be seen in the Percentage Price Change which is on average 2.966%. This is displayed in the following figure:

Figure 8. Price (BDI) timeline



As can be seen, there was a positive trend from early 2021 to a very high peak achieved in the fall. A similar trend but to a less intense degree can be seen in 2020. In 2022 the peak seems to be seen earlier and the peak for the first half of the year occurred. The average Rate for $/Tonne is 13.692 and the Rate for $/Day is 27,559. Especially the $/Tonne Rate seems fairly volatile and for that reason, the following figures are created to see if the trend movements are similar to the BDI:

Figure 9. $/Tonne rate timeline

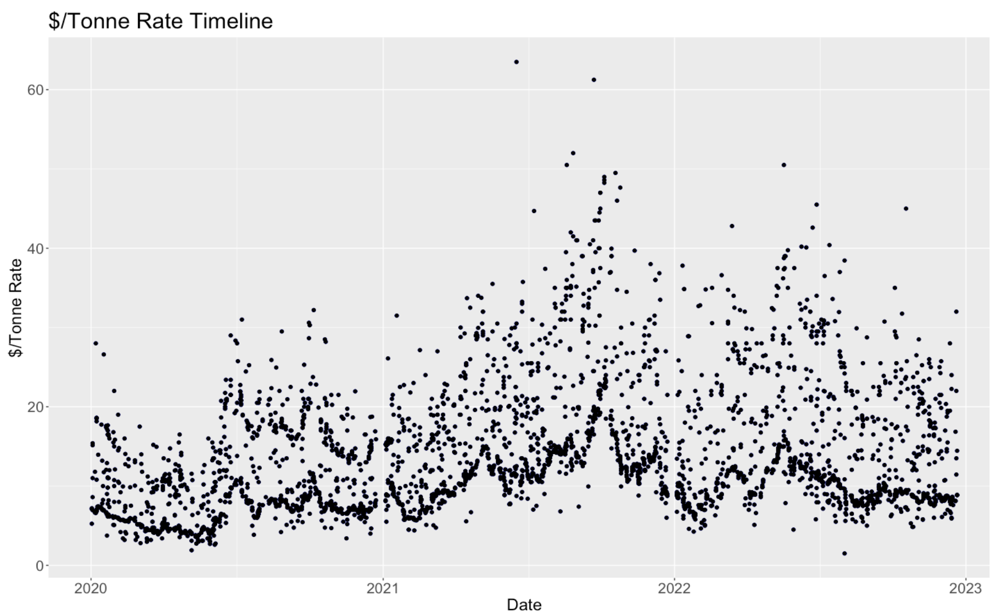
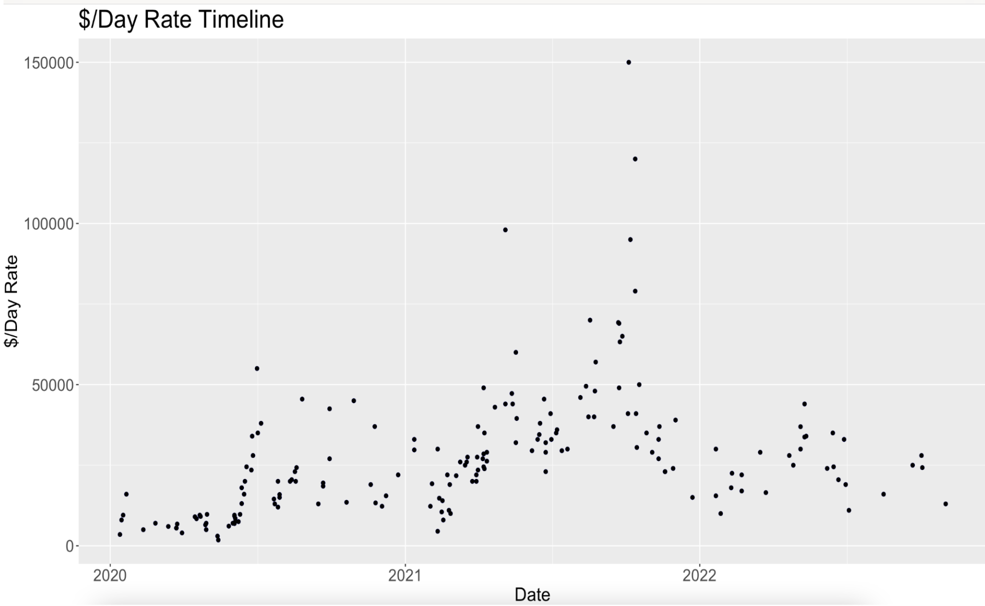


Figure 10. $/Day rate timeline



As mentioned earlier in the literature, the BDI and freight rates are interdependent and correlated in theory. It is conspicuous that both timelines show very similar trends to those found in the BDI and from that it can be concluded that there is correlation. However, the effects of the relatively big shocks in the BDI are less severe for the Rates. This suggests that the Rate is less volatile than the BDI and thus indicates that the Rates are less influenced by external factors relative to the BDI. This does not correspond one on one with the theoretical frame previously mentioned in the literature context and review, but there is correlation. For that reason, it can be concluded that the Rate is an interesting variable to keep an eye on while investigating the BDI when creating models.

However, 3716 contracts, consisting of 95.2% of the total contracts in the dataset, have freight rates paid in $/Tonne. The fact that only 181 contracts have a freight rate paid per $/Day in the dataset, will result in a shortage of data points for the models to show good results. Therefore, it is chosen to focus on the contracts which have freight rates paid per $/Tonne for this research in the final dataset. For that reason, the research continues with 3716 observations and 31 variables. This results in the adjusted summary statistics:

Table 6. Adjusted summary statistics economic elements of the dataset

Variable	N = 3716	Mean	Min	Max
Price (BDI)	3716	1,999	393	5650
% Price Change	3716	2.967	0	22.550
Quantity	3711	166,131	75,000	212,500

As can be seen, only using the $/Tonne contracts decreased the average BDI from 2,006 to 1,999. The average % price change increased by 0.001%, suggesting the price changes are a little bit more volatile without the $/Day contracts. The average value for quantity decreased from 166,439 to 166,131, this is due to the fact that 5 very high outliers are deleted. The $/Day contract had mostly missing values for the quantity, so this affected the average quantity very little.

5.3.3. Time elements of the dataset

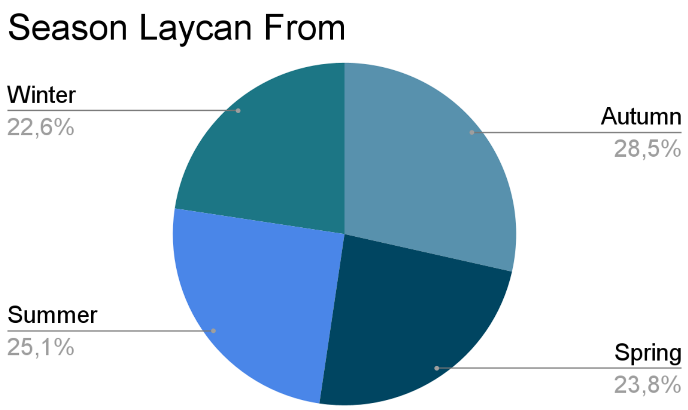
Having examined the date, price dynamics and economic variables, is it necessary to look at the time variables such as the number of days the ship is in laycan, the days between the contract and the start of the laycan period, the days shipping and the seasons. These are summarized in the following table and figures:

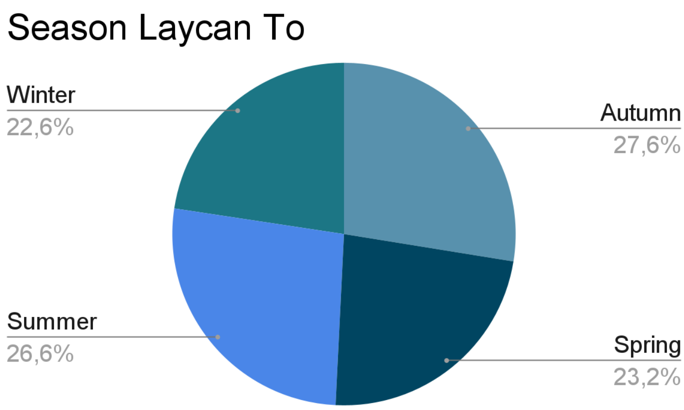
Table 7. Summary statistics time variables

Variable	N = 3716	Mean	Min	Max
Days Laycan	3716	5.377	0	2908
Days Contract Laycan From	3680	17.42	0	385
Days Shipping	3459	20.98	2.3	50.2
Different seasons laycan period = 1	3716	0.027	0	1
Different seasons date and laycan from = 1	3716	0.205	0	1

Figure 11. Distribution of seasonality Laycan From

Figure 12. Distribution of seasonality Laycan To

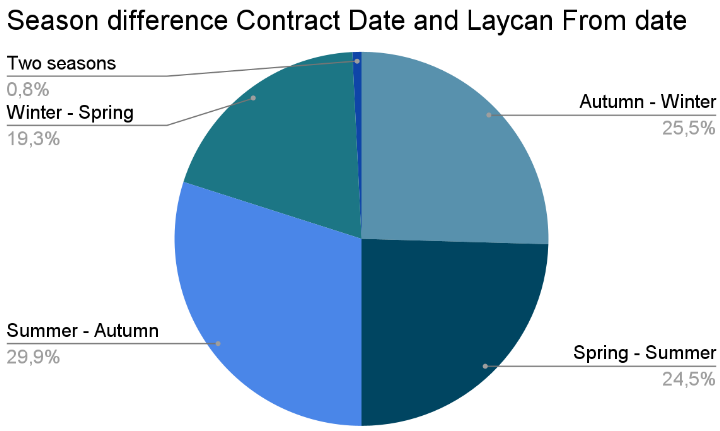




As indicated earlier, not all contracts are entered into the system with equal care and therefore there are some missings in the time variables. The data show that the average laycan period lasts about 5.4 days.

It can also be seen from the low mean value of the different season laycan period that the laycan period almost always is in the same season, only in only 2.7% there is a switch of seasons. This is also confirmed by Figures 11 and 12, showing that most laycan from and laycan to dates are during the summer and autumn. Notably, about 20.5% of all contracts have a different season from the contract date and the laycan from date, this suggests that the laycan from date is often in another season as the contract date season. The changes are presented in the following figure:

Figure 13. Changes from seasons from Contract Date to Laycan From date



As can be seen, most changes between the seasons of the Contract Date and the Laycan From date, 29.9%, happens between summer and autumn. So this figure indicates that most contracts, which are not in the same season, are closed during summer and have a starting laycan period in autumn. This occurs only 19.3% for contracts that do not take place in the same season for contracts concluded in winter and whose laycan period starts in spring and occurs least of all changes. It is notable that only 0.8% of the contracts have two season difference between the Contract Date and the Laycan From date, which indicates that it barely happens. It is important to keep in mind that 2943 contracts have the same season for the Contract Date and the Laycan From date and this figure only represents 20.5% of the data.

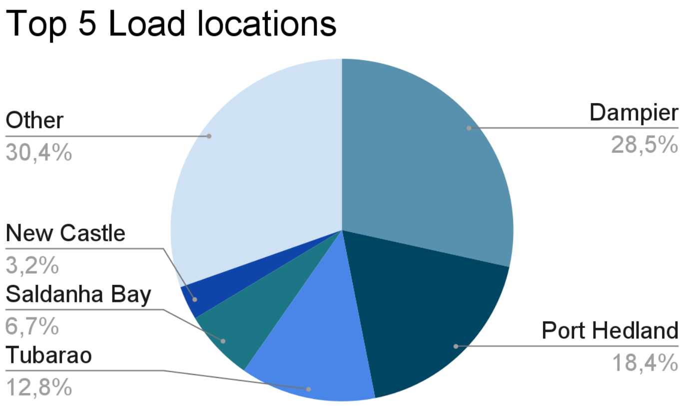
The average number of days between the contract dates and the first laycan day is 17.4 days, suggesting that on average there are more than two weeks before the contract and the actual loading of the vessel happens. It is conspicuous that the minimal value is 0, suggesting that there are ships that immediately start loading on the contract day. The mean value for the variable Days Shipping suggests that the average sailing time is 21.0 days in the contracts. In general, the summary statistics look logical for the date and time elements.

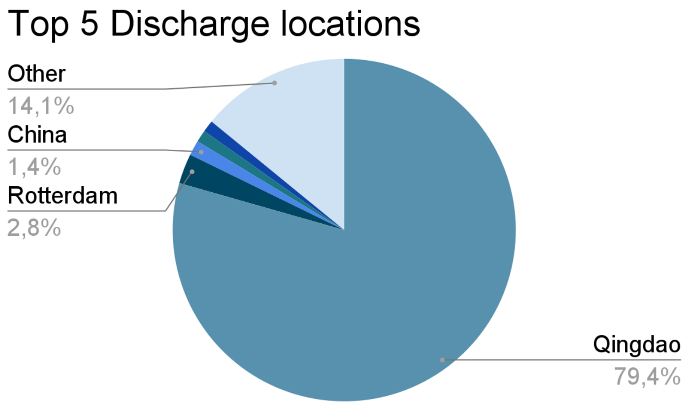
5.3.4. Location and geographic elements of the dataset

Next, it is necessary to look at the different locations of the load and discharges. As mentioned earlier, the load and discharge consist of 146 different locations and are relatively distributed. The following figures provides a summary of the five most common load and discharge locations:

Figure 14. Top 5 load locations

Figure 15. Top 5 discharge locations

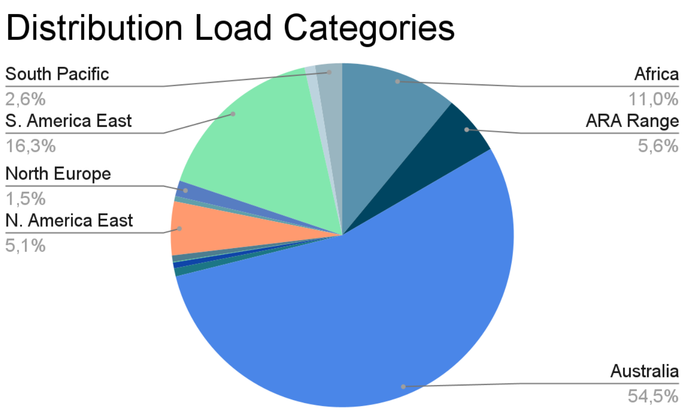


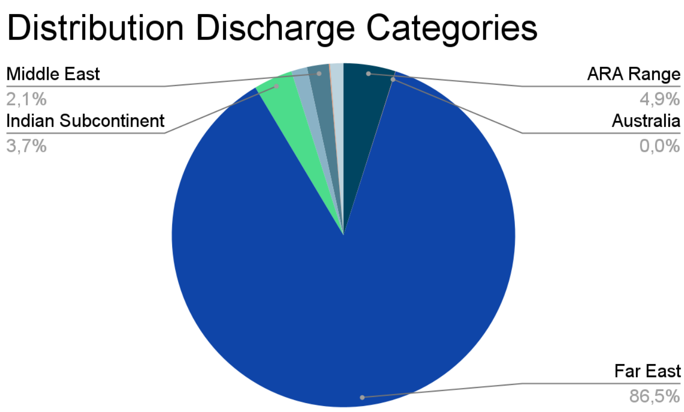


It can be clearly seen that the load locations are fairly distributed, but this is not the case with the discharge locations. Qingdao greatly dominates the discharge locations, as 79.4% of all ships are discharged at the ports of Qingdao. The other discharge locations consist of a very small proportion, suggesting that the remaining locations are widely distributed. Especially because discharge sites are widely distributed, categorization could be very useful. The following figures show the distribution of the categorised load and discharge locations:

Figure 16. Distribution of load categories

Figure 17. Distribution of discharge categories





The results of the data manipulation in transforming the load and discharge locations into the 13 categories have negatively affected the data spread in the Discharge Categories. As can be seen in the figure of the discharge categories, 86.5% of all discharges are located in the Far East. The Indian Subcontinent, which is relatively close to the Far East, has the third most contracts discharging. The figure on the left presents the Load Categories, which are positively impacted by the data transformation. As can be seen, more than half of the contracts start from Australia, however, there are also a number of other locations that have a reasonable number of contracts. In the following section, the variables committed to the ship's characteristics are discussed.

5.3.5. Ship and bulk characteristics of the dataset

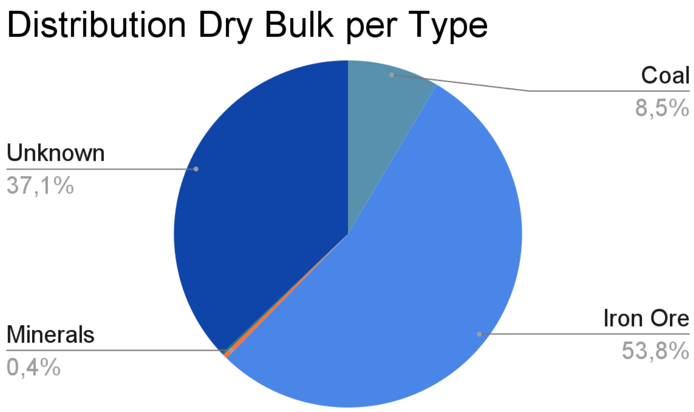
At last, the variables related to the characteristics of the ships and bulk are inspected. Starting with mapping out the number of distinct vessels in the dataset by inspecting the variable Name. The variable Name contains 2573 TBN, which suggests that 69.2% of all contracts cannot distinguish different ships. In this regard, there are 632 ships in the dataset of which many ships occur multiple times ranging from 9 to 1 times. This suggests that these ships took multiple contracts between January 2020 and December 2022 in the dataset. In addition, properties such as the year of construction and DWT of the vessels are investigated:

Table 8. Summary statistics ship characteristics

Variable	N = 3716	Mean	Min	Max
Built	1145	2011	1999	2022
DWT	1145	180,942	108,943	233,584

Both Built and DWT have many missing values, this is because all ships that do not have their name on the contract also automatically do not fill in the columns for the year of construction and DWT. The average DWT of the Capesize ships also matches the previously average weight of Capesize ships in the literature context. Thereby, the 2011 average year of construction shows that the oldest vessel is from 1999, this shows that the vessels in the dataset are relatively new. Furthermore, there are no peculiarities in the data regarding characteristics. As the last characteristic, the type of bulk is inspected. It is immediately apparent that this is not entered very accurately in the system of Clarksons and so many contracts state 'Unknown' as the type of dry bulk. This is summarized in the following figure:

Figure 18. Distribution dry bulk per type



As can be seen, 53.8% of all dry bulk contracts transport Iron Ore with Capesize ships. In addition, it is important to remember that the Unknown category is also large, where as much as 37.1% of the contracts are not properly completed. It is assumed that the distribution of the Unknown category would be distributed the same as the rest of the dataset but is kept as 'unknown' in the dataset.

A lot of data adjustments and modifications took place to arrive at the final dataset. The final dataset consists of 3716 contracts and 31 variables. This has been a very long process in this research to prepare the final dataset. In doing so, the research question and sub-questions are kept in mind by modifying the data in such a way that relevant output can be extracted from it. Having the dataset prepared, manipulated and discussed, the next chapter is devoted to the methods in this research.

6. Results

This chapter discusses the results per sub-question. The methodology of the first three sub-questions is fairly similar and for that reason the approach of describing the results is comparable. The methodology of the fourth sub-question deviates from the other sub-questions and therefore describing the results is somewhat different.

6.1. Sub-question 1 - *Can insights into time elements of contracts provide information about price dynamics in dry bulk freight rates?*

In order to investigate this sub-question, a Fixed Effects and Random Effects model are first created and then a Hausman Test is done to see which model is most appropriate. The output of the Fixed Effects and Random Effects is the following:

Table 9. Output Fixed Effects and Random Effects sub-question 1

	Dependent variable:	
	Baltic Dry Index	
	Fixed Effects	Random Effects
	(1)	(2)
Days Laycan	−9.169 (7.227)	−0.286 (0.293)
Days between contract and Laycan From	−13.161*** (4.922)	−8.118*** (1.364)
Season Laycan From - Spring	−359.544*** (94.041)	−435.614*** (39.840)
Season Laycan From - Summer	−193.416** (96.828)	−119.071*** (38.053)
Season Laycan From - Winter	−527.766*** (97.569)	−500.465*** (40.564)
Rate	84.284*** (4.876)	68.840*** (1.851)
Constant		1,448.336*** (43.012)
Observations	3,680	3,680
R^2	0.441	0.355
Adjusted R^2	−3.180	0.354
Note:		*p<0.1; **p<0.05; ***p<0.01

First when looking at the number of observations of the two models, it can be seen that 3,680 out of 3,716 observations are used. This is because both models cannot deal with missing values in the data and therefore leave out all observations that are missing values. In this case, 37 observations are missing the number of days between the start of the contract and the laycan from the date. As can be seen, the effects are the same for both models and the same variables are considered significant, but the magnitude differs. Thereby, it is worth mentioning that as discussed in the methodology, the Fixed Effects model does not has a constant and the Random Effects model does have a constant.

To find out which model is the most appropriate for this sub-question, a Hausman test is done after running both models. The Hausman test provided a P-value of 0.014, which indicates that the null hypothesis, that there is no systematic difference between the Random Effects and Fixed Effects coefficients, needs to be rejected at 5%. This suggests that the Fixed Effects model is the most appropriate to use for this sub-question.

The coefficient of ‘Days Laycan’, suggests that the effect of an extra day of laycan has a negative, but not significant effect on price dynamics. The coefficient of ‘Days between contract and Laycan From’ is -13.2 and is considered significant at 1%. This suggests that if the days between the contract and the laycan from increase by one day, the BDI decreases by 13.2, keeping other variables constant. Suggesting the number of days between the contract date and the laycan from date negatively influences the price dynamics.

The effect for the variable 'Rate' is considered significant at a 1% level. The positive coefficient of 84.3, suggests that if the rate increases by 1$, the Baltic Dry Index increases by 84.3, keeping other variables constant. This finding suggests that the freight rate prices positively affect the price dynamics. However, it is important to remember that this variable is also an interaction term to compare the effect of the freight rate on the price dynamics relative to other models.

For the variable ‘Season Laycan From’, the season autumn is used as the reference season. Interestingly, all seasons have a negative, but significant at 1%, effect. This suggests that the BDI is highest if the laycan from date (if the laycan period starts) is during autumn. As can be seen from the coefficient ‘Season Laycan From - Summer’, whenever the laycan period starts during the summer, the DBI is on average 193.4 lower than during Autumn, keeping other variables constant. Thereby it is also notable that about 53.6% of all shipping activity takes place during autumn and summer. The rest of the shipping activity occurs during spring and winter, both having a significantly bigger negative effect on the BDI than the other half of the year. The variables for Spring and Winter are -359.5 and -527.8, suggesting the BDI is significantly lower when the laycan period starts during those seasons compared to Autumn. This suggests there is seasonality and possibly a six-month cycle influencing the price dynamics according to this model.

To summarise, the Hausman test implies that Fixed Effects model is the most appropriate for investigating the effects of time elements on price dynamics. The output for the days between the contract and the laycan from date suggests that a longer period between the contract and the laycan period has a negative impact on the price dynamics. However, it does emerge that seasons have a major effect on price dynamics, suggesting there could be a six-month cycle. It can be stated on the model including the time elements, that the rate has a positive average impact on the price dynamics.

6.2. Sub-question 2 - *What are the effects of the geographical elements of shipping activities on the price dynamics?*

In order to investigate this sub-question, a Fixed Effects and Random Effects model are first created and then a Hausman Test is done to see which model is most appropriate. The output of the Fixed Effects and Random Effects is summarized in the following table:

Table 10. Output Fixed Effects and Random Effects sub-question 2

	Dependent variable:	
	Baltic Dry Index	
	Fixed Effects	Random Effects
	(1)	(2)
Load.CategoryAfrica	−867.357*** (154.624)	−800.256*** (49.559)
Load.CategoryARA Range	−1,055.555*** (207.489)	−719.440*** (59.027)
Load.CategoryFar East	−343.837 (851.915)	−80.647 (204.215)
Load.CategoryMiddle East	−266.194 (433.794)	−823.012*** (184.131)
Load.CategoryNorth America East Coast	−1,824.584*** (246.940)	−1,789.075*** (76.270)
Load.CategoryNorth America West Coast		−357.180* (190.029)
Load.CategoryNorth Europe	−1,425.515*** (288.439)	−673.128*** (106.996)
Load.CategorySouth America East	−950.477*** (186.640)	−930.557*** (58.481)
Load.CategorySouth America West	−883.405*** (286.474)	−1,070.408*** (140.069)
Load.CategorySouth Pacific	565.299** (277.893)	298.659*** (73.528)
Discharge.CategoryARA Range	1,480.182*** (224.622)	1,131.076*** (72.096)
Discharge.CategoryIndian Subcontinent	−437.690 (318.986)	−109.648 (75.164)
Discharge.CategoryMediterranean	1,384.857*** (307.424)	584.238*** (127.276)
Discharge.CategoryMiddle East	−5.458 (351.235)	106.356 (99.944)
Discharge.CategoryOther		−983.199 (642.615)
Discharge.CategorySouth Pacific	418.699 (929.273)	461.453*** (111.350)
Days.Shipping	−32.822*** (6.938)	−37.931*** (2.201)
Rate	158.630*** (4.726)	155.592*** (1.998)
Constant		1,017.577*** (37.329)
Observations	3,459	3,459
R^2	0.730	0.653
Adjusted R^2	−1.075	0.651
Note:		*p<0.1; **p<0.05; ***p<0.01

First, when looking at the number of observations of the two models, it can be seen that 3,459 out of 3,716 observations are used. This indicates that 257 observations are dropped, due to the fact that there are 257 observations missing for the 'Days Shipping'. This is due to the manual calculations, which are done for routes which are used at least 10 times per location or categorized location in the dataset.

The significance levels differ for the 'Load Category Middle East', and 'Discharge Category South Pacific', whereas these variables are not significant for the Fixed Effects model and are considered significant for the Random Effects model. To find out which model is the most appropriate for this sub-question, a Hausman test is done on both models. The Hausman test provides a P-value of 0.001, which indicates that the null hypothesis does need to be rejected at 1%. This suggests that there is a systematic difference between the Random Effects and Fixed Effects coefficients and thus that the Fixed Effects model is the most appropriate to use for this sub-question.

First the ‘Days Shipping’ and the ‘Rate’ are investigated. The negative coefficient -32.8 of ‘Days Shipping’, suggests that a one-day increase in sailing from one location to another decreases the BDI by 32.8, keeping other variables constant, this effect is significant at 1%. This suggests that longer routes have a negative effect on the price dynamics. The variable ‘Rate’ has a positive significant coefficient at 1% of 158.6. Suggesting a positive effect of an increase in the freight rate on the price dynamics. Whenever the freight rate increases by 1$, the BDI increases by 158.6, keeping other variables constant. It is notable that the positive effect of this interaction term is almost twice as high as in the previous model including the time elements.

The coefficients for the load and discharge categories are only interpreted if they are significant for at least 10%, which in this case are also significant at the 5% level. It is important to keep in mind that the load and discharge categories need a reference category. In this case, there is not one load and discharge category that occurs about equally often. For that reason, the two most frequently occurring categories are picked as references. The load category is Australia used as a reference, as 54.5% of ships load from that category and the discharge category is the Far East, with 86.5% of ships discharging from that category, used as a reference. This is in line with the literature context, as it stated that these routes are very important to global trading and therefore often used.

The coefficients for loading in the Far East and Middle East are not considered significant and therefore it only can be stated that loading from these locations, compared to loading in Australia, negatively impacts the price dynamics. It is notable that all coefficients for all other categories except South Pacific are all negative, suggesting that loading from those locations negatively affects price dynamics compared to loading from Australia, keeping other variables constant, these effects are significant at 5%. As stated in the literature context, Australia is rich in natural resources such as iron ore and coal and one of the leading exporters of these two bulk goods. Also, the shipping routes from Australia to the Far East are relatively favourable and well-established as Australia supplies much of these goods to the Far East. For that reason, there are many long-term trade agreements for Capesize ships transporting iron ore and coal to the Far East and through these arrangements are less sensitive to price dynamics.

Only the South Pacific coefficient of 565.3 is positive, suggesting that loading from the South Pacific increases BDI by 565.3 compared to loading from Australia, significant at 5%, holding all other variables constant. It should take into account that only a total of 93 contracts, representing 2.6% of all loading locations, were loaded from the South Pacific. Overall, it can be stated about the load categories that especially all other locations, except the South Pacific, have a negative impact on the price dynamics.

The coefficients of 1,480.2 and 1,384.9 for the ARA Range and the Mediterranean are the only two discharge categories which are considered significant. This suggests that whenever the ship is discharged in the ARA Range or in the Mediterranean, the BDI increases by 1,480.2 and 1,384.9 compared to discharging at the Far East, holding other variables constant, this effect is statistically significant at 1%. It should be taken into account that there is a total of 180 contracts discharged at the ARA Range and 55 contracts discharged at the Mediterranean. This suggests that the discharging at the ARA Range and Mediterranean, consisting of 4.8% and 1.5% of total contracts, positively impacts the price dynamics. The other discharge categories are not considered significant, however, it can be stated that discharging in the South Pacific positively affects the price dynamics and discharging at the Indian Subcontinent and the Middle East negatively affects the price dynamics, compared to discharging in the Far East.

In summary, the Hausman tests pointed out that the Fixed Effects model was the most appropriate model to use with the geographical elements as independent variables and thus for this sub-question. From the Fixed Effects model, it can be stated that ships loading in Australia or in the South Pacific positively impact the price dynamics relative to loading from all other load categories. The discharging category ARA Range and the Mediterranean positively impact the price dynamics compared to the Far East according to the model. The other discharge categories are not considered significant, which makes it irrelevant to interpret the coefficients. The coefficient for the rate is considered relatively high in this model. Keeping in mind that the effect of the rate in this model is almost twice as high, compared to the previous model including the time elements. From that, it can be stated that including geographical elements in the model has a large impact on the freight rate and the price dynamics.

6.3. Sub-question 3 - *How do freight elements of dry bulk influence the price dynamics?*

In order to investigate this sub-question, a Fixed Effects and Random Effects model are first created and then a Hausman Test is done to see which model is most appropriate. The output of the Fixed Effects and Random Effects is summarized in the following table:

Table 11. Output Fixed Effects and Random Effects sub-question 3

	Dependent variable:	
	Baltic Dry Index	
	Fixed Effects	Random Effects
	(1)	(2)
Type - Iron Ore	−471.925** (201.962)	−232.744*** (59.045)
Type - Minerals	−1,475.302** (621.007)	−1,205.616*** (233.869)
Type - TCT	−1,652.124 (1,215.261)	−682.371** (336.255)
Type - Unknown	−644.185*** (204.361)	−276.939*** (59.492)
Quantity	0.011*** (0.004)	0.002* (0.001)
Rate	87.529*** (4.831)	72.133*** (1.847)
Constant		890.990*** (191.658)
Observations	3,711	3,711
R^2	0.402	0.317
Adjusted R^2	−3.470	0.316
Note:		*p<0.1; **p<0.05; ***p<0.01

First, when looking at the number of observations of the two models, it can be seen that 3,711 out of 3,716 observations are used. This can be explained by the 5 missing observations for the 'Quantity', which result in the total used observations of 3,711. As can be seen, the effects are the same for both models, but the magnitude differs. It is striking that in both models almost all variables are considered significant, except the 'Type - TCT' in the Fixed Effects model.

Before the variables and their coefficients can be interpreted, it is necessary to examine which model works best using the Hausman test. The Hausman test provides a P-value of 0.004, which suggests that there is a significant difference between the Random Effects coefficients and the Fixed Effects coefficients. The null hypothesis must be rejected at 1% significance and therefore the Fixed Effects model is most appropriate for this sub-question.

As can be seen, the variable 'Type' is categorical and it is important to keep in mind that 'Coal' is used as reference category. All coefficients for the other 'Types' are negative and also relatively high, this suggests that Coal is mostly transported whenever the BDI is high. This is notable, as coal transportation is only 8.5% of all contracts, while iron ore consists of 53.8% of all contracts. The coefficient of -471.9 of 'Iron Ore' suggests that the BDI is on average 471.9 lower whenever a ship transports Iron Ore compared to transporting Coal. The coefficients for 'Minerals' and 'Unknown' of -1,475.3 and -644.2

suggest that the BDI is on average 1,473.3 and 644.2 lower whenever a ship transports Minerals or if the cargo is Unknown, compared to transporting Coal. These effects are significant at 5%. It is important to keep in mind that only 15 contracts (0.4%) transport Minerals.

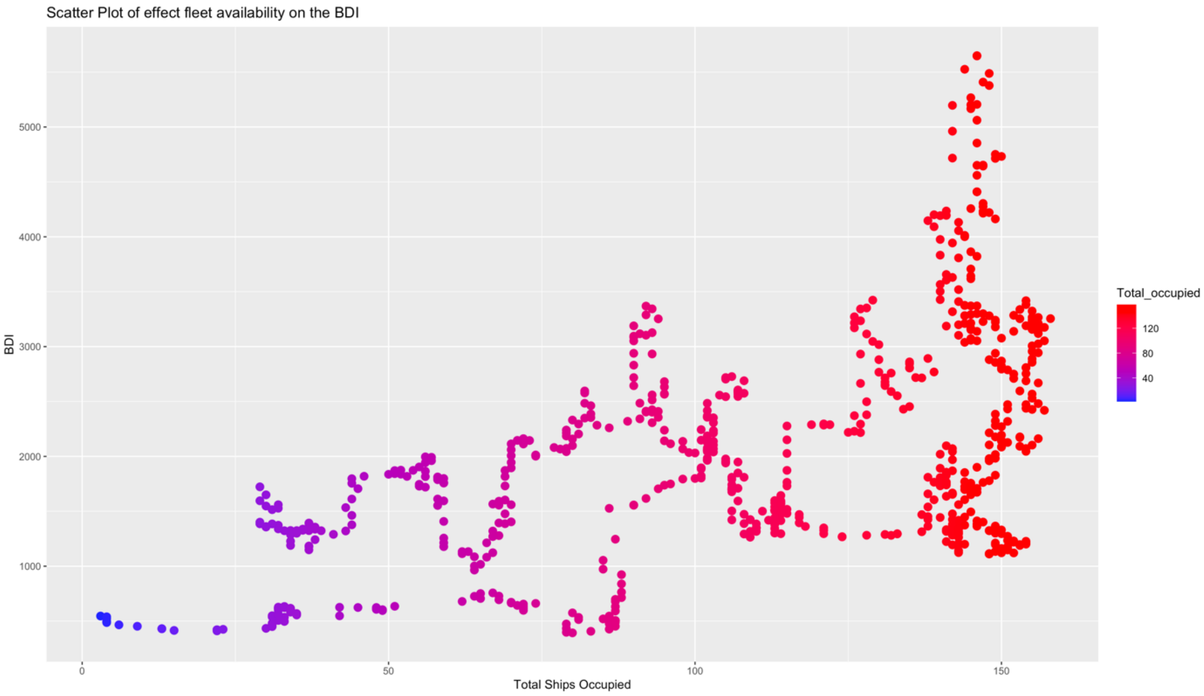
The coefficient for 'Quantity' of 0.011 is significant at 1% but is relatively small. This suggests that if the quantity increases by 1, the BDI increases by 0.011, keeping other variables constant. This suggests that shipping a higher quantity has a positive effect on the price dynamics. However, it is important to keep in mind that the mean value of the quantity is 180,942 and it should be taken into account that the maximum capacity of a ship can not be exceeded. The 'Rate' has a positive coefficient of 87.5 in this model. This indicates that a 1$ increase in the freight rate, increases the BDI by 87.5, keeping all other variables constant, this effect is significant at 1%. This coefficient is close to the freight rate of the model including time elements and relatively small compared to the freight rate of the model including geographical elements.

In summary, the Fixed Effect model shows that the type of bulk has a significant effect on the price dynamics. All the negative coefficients for all other types than Coal suggest that these negatively impact the price dynamics. Also, it can be stated that the quantity of shipping positively impacts the price dynamics. All three Fixed Effect models included the freight rate as the independent variable. From these models, it can be stated that the ship characteristics and time elements have approximately the same effects on the rate and price dynamics. The effects of the geographical elements on the rate and price dynamics are considered greater relative to ship characteristics and time elements. The freight rate always has a positive impact on Baltic Dry Index and thus on the price dynamics, which is in line with the literature.

6.4. Sub-question 4 - *What are the effects of fleet availability on price dynamics in the market?*

The first step in order to investigate this sub-question, was creating a graphical representation of the ship's occupation relative to the Baltic Dry Index. It should be taken into account here that there are many missing names in the dataset, but it is not clear which 'TBN' ships sail multiple times, and which do not. It was decided not to include these ships for this sub-question but only those whose Name was available in the dataset. As a result, a total of 609 different boats were included in this sub-question. As mentioned earlier in the literature context, the total Capesize fleet consists of 1593 ships, which means that these three years of contracts thus comprise 38.2% of the total fleet. The impact of the number of ships occupied on BDI is expressed in the following figure:

Figure 19. Scatter plot of fleet availability on the BDI



As can be seen, a positive relationship between the total number of ships occupied and the BDI. Especially the points on the left side show little variation and suggest that the BDI is relatively low if relatively few boats are occupied. It is notable that the points more to the right, which indicate that the number of total ships occupied is higher, have a lot more variation in the BDI. This suggests that there is an impact of fleet availability on price dynamics, as the BDI is more volatile during high-occupied times. For that reason, it is necessary to investigate where this volatility comes from.

For this purpose, a Random Forest model was created to investigate which variables have an impact on the total ship occupation. For that reason, the total ships occupied is used as the dependent variable and all other variables which could be relevant, such as the type, DWT, quantity, rate, days laycan, days between contract, seasonality, days shipping, load categories, discharge categories, are used as independent variables. The following figures show the results of the Random Forest model:

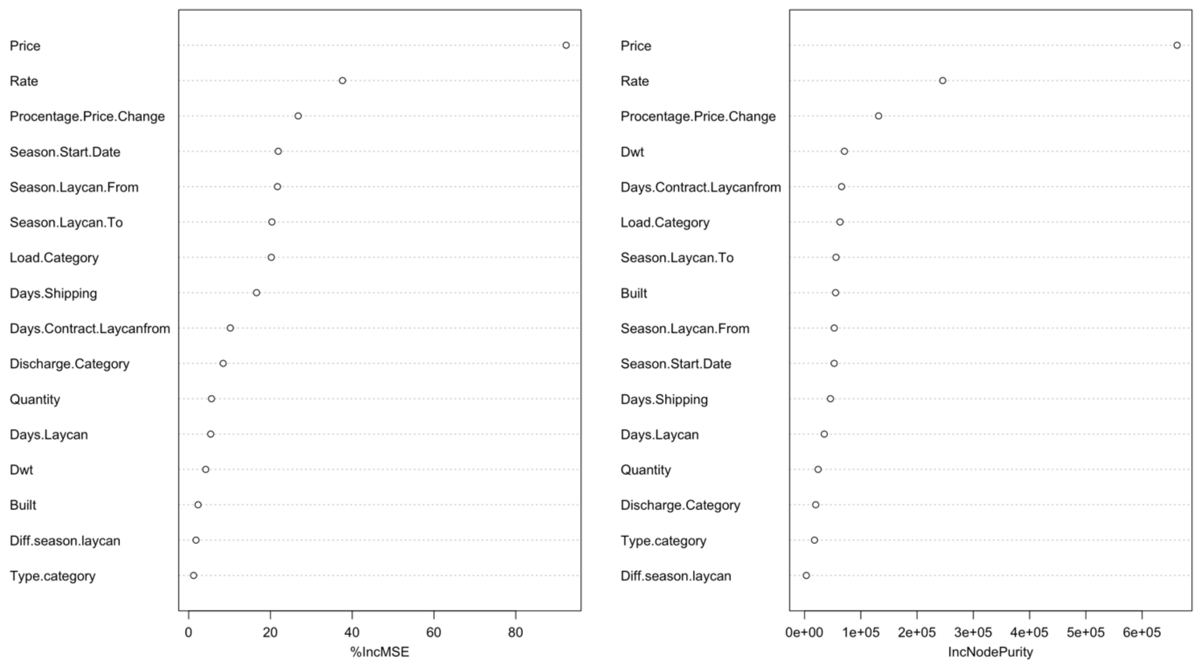
Figure 20. The Random Forest model on total ships occupied

```
Call:
 randomForest(formula = Total_occupied ~ Price + Built + Dwt +      Quantity + Rate + Days.Laycan +
 Type.category + Days.Contract.Laycanfrom +      Load.Category + Discharge.Category + Days.Shipping
 + Procentage.Price.Change +      Season.Start.Date + Season.Laycan.From + Season.Laycan.To +
 Diff.season.laycan, data = Q4.DF.occupied, importance = TRUE,      na.action = na.exclude)
               Type of random forest: regression
                     Number of trees: 500
No. of variables tried at each split: 5

          Mean of squared residuals: 549.1677
                    % Var explained: 65.69
```

The Random Forest model is created, and the output is presented in the figure above. The model is based on 500 trees due to its viability in terms of accuracy and permutation time. The model discovered that the biggest variances were produced by doing Random Forests on five variables that were randomly picked as candidates at each split. The variation described by Random Forests for ship occupation is 65.69%, which is an indicator of how effectively the model reflects the input characteristics' influence on the variability of the total ships occupied. From this, it can be concluded that the independent variables can predict the occupation quite well. Therefore, as a next step the variable importance plot, showing the %incMSE and IncNodePurity, need to be investigated, to see which variables give the most predictive power to the model.

Figure 21. The Variable Importance Plot on total ships occupied



It is immediately notable that the (Price) BDI is at both the %incMSE and IncNodePurity at the top of the list. The highest %incMSE implies that BDI is the most important variable in predicting how many ships are occupied and the IncNodePurity implies that BDI is also the most important for deciding the splits in decision trees in the models, which eventually are formed to the Random Forest model. This result suggests that the trend visible in the previous figure with the impact of the number of ships occupied on BDI is valid.

Furthermore, it becomes clear that five variables mattered when making the splits in the decision trees. This can be seen in the IncNodePurity on the right side of the figure, as the top three variables have a significantly higher score than the other variables. This suggests that the BDI, the (freight) rate and the percentage change in the BDI are the most important variables to determine the splits in the decision trees, according to this Random Forest model. The DWT and the days between the contract and the laycan from date are also important determinants according to the IncNodePurity of the Random Forest model. Strikingly, the exact same three variables: BDI, (freight) rate and percentage change in the BDI, have the highest %incMSE and are thus also most important for predicting the ship's occupation. Also, seasonality seems to be important in determining how many ships are occupied.

Summing up, it can be stated that the BDI is the most important variable, followed by the freight rate and the percentage change in the BDI. It is rather striking that the number of days of shipping is not an important determinant of the occupation according to the model. The DWT and the days between the contract date and the laycan from date are considered important to determine splits in the trees and thus important to model the ship's occupation. The seasonality, on the other hand, is mostly important to predict the ship's occupation itself, however, it would make the most sense to put the seasonality of the contract date, as the ship's occupation is based on the contract date. For that reason, the variables total occupied, rate, days between contract date and laycan from, DWT and the seasons of the contract date are included in a Fixed Effects and Random Effects model, with the BDI as the dependent variable, to determine the effect of fleet availability on the price dynamics. Here it is specifically chosen not to include percentage change in the models, as it would not make economic sense to model the percentage change in price against price.

Table 12. Output Fixed Effects and Random Effects sub-question 4

	Dependent variable:	
	Baltic Dry Index	
	Fixed Effects	Random Effects
	(1)	(2)
Total Occupied	10.034*** (0.862)	10.522*** (0.593)
Rate	81.141*** (4.637)	65.278*** (2.933)
Days Contract Laycan From	−13.555*** (3.931)	−10.829*** (2.890)
Dwt	0.004 (0.011)	−0.005** (0.002)
Season Spring contract date	−103.345 (87.209)	−201.528*** (62.379)
Season Summer contract date	−93.429 (87.258)	−129.545** (62.617)
Season Winter contract date	−383.298*** (91.170)	−374.761*** (66.877)
Constant		1,155.402*** (445.642)
Observations	1,043	1,043
R^2	0.603	0.559
Adjusted R^2	0.048	0.556
Note:		*p<0.1; **p<0.05; ***p<0.01

Firstly, looking at the observations, it can be seen that a total of 1,043 observations were included in the model. As stated previously, the dataset consists only of contracts of which the Name of the ship is known, because otherwise, it was not possible to map out the occupation of the fleet. For that reason, only 609 ships are included in the model, which contain of 1,043 contracts. It is notable that all coefficients for the Random Effects model are considered significant, which is not the case for the Fixed Effects model. In order to determine which model is the most appropriate, the Hausman test is performed. The Hausman test provides a P-value of 0.002, suggesting that the null hypothesis, that there are no systematic differences between the Random Effects and Fixed Effects coefficients, must be rejected at 1% significance. For that reason, the Fixed Effects model is most appropriate to use as the Random Effects model would be biased.

The coefficient 10.0 for the variable 'Total Occupied' is considered significant at 1% and suggests that if one extra ship is occupied, the Baltic Dry Index increases with 10, keeping the other variables constant. This suggests that a higher ship occupation, which means less fleet availability, results in a higher BDI. In terms of price dynamics, the variable 'Rate' has also a positive coefficient of 81.1. This suggests that if the freight rate increases with 1$, the average BDI is 81.1 higher, keeping the other variables constant, this effect is significant at 1%. These findings suggest that the ship occupation and the freight rate prices positively affect the price dynamics.

The variable 'Days Contract Laycan From' has a coefficient of -13.6, which suggests that the number of days between the contract and the laycan from date has a negative effect on the price dynamics at a significance level of 1%. If the number of days increases with one, the BDI decreases with 13.6, keeping the other variables constant. This suggests that if the number of days between the contract date and the laycan from date is longer, the BDI is on average lower. The ship characteristic DWT has no significant effect, but does suggest that it has a positive effect on the BDI.

Similar to the model from sub-question 1, autumn is used as a reference for the seasons. Again, all the effects are negative, suggesting the BDI is highest in autumn according to this model too. Remarkably, only the coefficient for Winter of -383.3 is considered significant and has the largest impact of all seasons. This suggests that for contracting the BDI is 383.3 during winter, compared to contracting during autumn.

To summarise, the first effect of the ship occupation on the BDI is investigated graphically. This figure clearly showed that there is a positive trend in the ship occupation and the BDI. The figure also showed that the volatility in the BDI increases if the number of occupied ships increases. This suggests that there is an impact of fleet availability on price dynamics, as the BDI is more volatile during high-occupied times. The variation of 65.69% of the Random Forest, indicates that the model reflects the input characteristics' influence on the variability of the total ships occupied quite well.

As a next step is the Variable Importance Plot created, which suggested the five most important variables in determining ship occupation. These variables are included in the Fixed Effects and Random Effects model to investigate the effects of fleet availability on price dynamics. The Hausman test showed that the Fixed Effects model was the most appropriate, as with the other sub-questions. The Fixed Effect model shows that the total number of ships occupied and the freight rate are both positively affecting the BDI and thus the price dynamics. It is notable that the effect of the interaction term rate is similar to the models including time elements and freight elements and thus relatively smaller than the model including geographical elements. Having discussed and examined all the results, it is time to draw conclusions in the next chapter.

7. Conclusion

This chapter is divided into five sub-chapters. First conclusions are drawn for the sub-questions and eventually, answers are given for each sub-question. Ultimately, this also allows the main question to be answered and an overall conclusion for the research to be drawn.

7.1. Conclusion sub-question 1

The first sub-question investigates the effects of time elements in individual spot-rate contracts on the price dynamics with a Fixed Effects model. From the output of the model, it can be concluded that the days laycan have no significant effect on the price dynamics. Also, from the model it can be concluded that the days between the contract date and the laycan from date and the seasons during the laycan from date do have significant effects on the price dynamics.

It can be concluded that the days between the contract date and the laycan from date negatively affect the BDI. This suggests that contracts with a shorter period between the contract date and the laycan from date are closed while the BDI is higher on average. On the contrary, this suggests that contracts are more likely to be closed further in the future whenever the BDI is lower. This makes sense from an economic point of view, as the ships contracting on a shorter time notice probably have more urge to ship bulk from and to a desired location and thus have less margin to wait for a favourable BDI.

Whenever the BDI is higher, it frequently indicates a competitive market with fewer available vessels than there is cargo demand. In these circumstances, shipping companies and owners can be more picky with charter bids and may choose to lock in short-term contracts so they can choose to close more lucrative transactions in the near future. In other words, the shipping companies have more power to select the contracts they prefer, which are in that case short-term contracts. This outcome is also in line with the research of Alizadeh & Talley (2011).

Concerning the seasonality in the laycan from dates, it can be concluded that the BDI is highest during autumn and lowest during winter. The output of the Fixed Effects model suggests that there is seasonality and possibly a six-month cycle influencing the price dynamics. This effect is also reflected in the BDI data from 2020 to 2023 (figure 8), where the year 2021 has the biggest impact on this effect. It can be stated from this that the conclusion of the model is consistent with the trend in the data. Also, this suggests that contracts whose laycan period begins in winter and spring are, on average, closed during a lower BDI. This is not in line with the research of Poblacion (2015), as stated in the literature chapter. Poblacion (2015) conclude that the freight rates are higher during winter and spring according to the TCE and WS data of five routes defined by Baltic.

Kavussanos & Alizadeh (2001) concluded that seasonal oscillations are more pronounced and sharper during market expansion times than during market contraction periods. They state that the differences diminish as contract duration grows, suggesting longer contract durations result in more stability in freight rates. This is seasonal effect and effect on the contract duration is in line with the conclusions of this research.

From an economic perspective, it is conspicuous that during the winter and spring, only 46.4% of the contracts have their laycan period starting. Logically, a lower BDI would lead theoretically to a higher demand for contracting, however, this is not the case. In total 20.5% of the contracts have a switch in seasons from the contracting date to the laycan from date. The average number of days between the contract date and the laycan date is 17.4, however, from this, it can be concluded that it is quite well distributed as the changes from winter to spring and spring to summer are occurring 19.3% and 24.5% in the data. From an economic perspective, it would make more sense if the biggest portion of contracts closed during winter and starting during spring, as the BDI is on average lower during winter, however, this is not the case. From this output, it can be concluded that the time elements do have an effect on the price dynamics, however, the effects of these price dynamics do not seem to make much effect on the timing contracting decision-making.

At last, it can be concluded that the freight rate in the Fixed Effects model has a positive effect on the BDI, however, this effect can be best implemented as an interaction. This is also suggested in the literature, which can be interpreted as a confirmation of the model's reliability in investigating the effects of time elements on price dynamics as freight rate and BDI are correlated. In this model, the freight rate has a relatively average impact on the price dynamics, and it can be concluded that the time elements do have a significant positive impact on the price dynamics.

Finally, the sub-question can be answered: *Can insights into time elements of contracts provide information about price dynamics in dry bulk freight rates?* Yes, insight into laycan periods of contracts does provide information about price dynamics. It can be concluded that it is important to keep time elements such as the contract date and the laycan date, in a seasonality perspective in mind and the days between the contract date and the laycan from date. These elements do influence the price dynamics significantly as contracting on shorter notice and contracting during autumn and summer result in higher prices relative to contracting on longer notice during spring and winter.

7.2. Conclusion sub-question 2

The second sub-question investigated the effects of geographical elements of shipping activities on price dynamics. This is investigated with a Fixed Effects model, using the Baltic Dry Index as the dependent variable and the load locations, discharge locations and days shipping as independent variables.

It can be concluded that the days shipping has a negative effect on the BDI. Suggesting longer routes have a negative effect on the BDI and thus a negative effect on the price dynamics. This is understandable, as it is economically speaking more attractive to plan and contract longer routes whenever the BDI is lower, as this effect in BDI difference levies more for longer than shorter routes. For that reason, the conclusion made from the model is understandable from an economic perspective and is in line with the conclusions from Zhang & Zeng (2015), which conclude that a greater price discovery function is produced by longer voyages.

Concerning the loading location, all coefficients are considered negative and only loading from the South Pacific has a positive coefficient. From this, it can be concluded that contracts loading from Australia, which is used as references, are on average closed while the BDI is higher relative to all other locations except the South Pacific. In total 2.6% of the contracts for ships loading from the South Pacific are on average closed on a much higher BDI relative to all other locations. From this, it can be concluded that ships departing from the South Pacific are on average closed during higher BDI and thus have a large impact on price dynamics. It is notable, that 54.5% of the ships loading from Australia and that the BDI is relatively high compared to other locations on the contract date. This could be explained by the fixed trading routes and long-term agreements between Australia and the Far East, which could result in less consideration being given to price dynamics. This is in line with the findings of the research of Grammenos (2013) and Over (2017). From the loading perspective, it can be concluded that the Far East requires the import goods and therefore accept to pay higher prices for the contracts.

This effect is not represented in the result of the Fixed Effect model for the discharge locations. Since 86.5% of the contracts are discharged in the Far East, it can be concluded that the data is not very well divided to model the effect of all discharge locations. However, this distribution is in line with the literature and therefore considered representative of the Capesize shipping industry. As the dataset is considered representative, it is certifiable to conclude that the output of the model is representative. From this, it can be concluded that discharging in the Far East does not necessarily influence the price dynamics.

Only discharging in the ARA Range and the Mediterranean are considered significant and have a relatively big positive effect on the price dynamics. It makes sense that the BDI is higher if the ships are discharging in these locations compared to discharging in the Far East, due to the fact that costs such as port handling costs and labor costs are higher in these locations. Also, the costs to ship to Europe from Australia are relatively higher as the distance is longer relative to shipping to the Far East. For that reason, it is concluded that due to the fact that these costs are higher anyway, the freight rate is a lesser part of the total cost and therefore ships still sail to this the ARA Range and the Mediterranean.

From the Fixed Effects model, with the geographical elements as independent variables, it can be concluded that the freight rate has a major impact on the price dynamics. As with the previous sub-question, the rate is considered an interaction between the geographical elements and the BDI. It is striking that the rate in this model has almost twice the effect on BDI relative to the model including the time elements. From this, it can be concluded that the interrelationship is larger for the freight rate of the geographical elements relative to the time elements. In conclusion, the freight rate is an important factor in contracts to determine the route.

According to the literature, geographical elements have an effect on the price dynamics, causing some deviation in route usage. This is reflected and confirmed by the Fixed Effects model, therefore, the sub-question: *What are the effects of the geographical elements of shipping activities on the price dynamics?*, can be answered. Overall, the effects of the geographical distribution of shipping activities impact the price dynamics both positively and negatively. It can be concluded that longer routes negatively impact the price dynamics. The loading location positively impacts the price dynamics as ships are loaded from Australia and the South Pacific compared to all other locations. It also can be concluded that fixed trading routes and long-term agreements are noticeable from the data and do influence the price dynamics. The discharge location is positively impacted whenever ships are discharged in the ARA Range and the Mediterranean compared to the Far East. To conclude, including geographical elements in the model has a relatively large impact on the freight rate and the price dynamics.

7.3. Conclusion sub-question 3

The third sub-question investigates how freight elements of spot-rate contracts influence the price dynamics with a Fixed Effects model. In the model, the type of the bulk, the quantity transported and the freight rate are used as independent variables and the BDI as dependent variable.

According to the model, a higher quantity transported increases the BDI and therefore it can be concluded that the dry bulk quantity shipped has a positive effect on the price dynamics. This makes sense, as the BDI is high, there is typically a requirement to move big amounts of products effectively since it signals a strong demand for bulk commodities. By chartering bigger Capesize vessels, there is an advantage of economies of scale, which have cheaper transportation costs per tonne of cargo than smaller vessels. This makes sense, as Cooke et al. (2014) state that freight rate is typically calculated pro rate based on the loaded cargo. Thus, when the BDI is high and the charter prices are high, chartering larger ships becomes more appealing than chartering smaller ships.

Regarding the type of bulk, coal is used as the reference category in the model. As all coefficients for the other 'Types' are negative, it can be concluded that Coal is mostly transported whenever the BDI is high. As stated in the literature, a higher BDI correlates with higher worldwide economic activity. The demand for raw commodities like coal to power manufacturing, electricity generation, and other energy-intensive businesses rises during periods of rapid economic expansion (S&P Global Market Intelligence, 2021; Over, 2017). Therefore, it is understandable that the demand for coal transportation increases with a higher BDI and it can be concluded that the output of the model makes sense.

The BDI is on average lower for the transportation of iron ore compared to coal transportation, but the BDI is still higher relative to the other types of bulk. As stated earlier in the literature, whenever the economic landscape is nourishing, industrial output is rising and therefore the demand for iron ore increases. Also, the demand for the raw commodity iron ore is also affected by the iron ore prices (S&P Global Market Intelligence, 2021; Over, 2017). It can be concluded that the effect of iron ore on the BDI and thus on the price dynamics has a similar trend as coal but to a lesser extent. The trend for minerals, TCT and if the type is unknown is different, as it influences the BDI negatively compared to coal. From this, it can be concluded that the price dynamics are negatively affected in transporting minerals, TCT or if the type is unknown. The unknown category was assumed to be distributed the same as the other categories and from the output of the model, it can be concluded that this assumption holds.

The freight rate has a positive effect on the BDI according to the Fixed Effects model with the freight elements implemented as independent variables. As with the model from the previous sub-questions, the freight rate can be best implemented as an interaction. Literature suggests that the BDI is correlated with the freight rate and has an impact on the transportation of dry bulk goods, and this is confirmed by the model. It can be concluded that the interaction effect of the model with the freight rate on the bulk characteristics is fairly the same as for the interaction effect of the model with the rate on the time elements.

Eventually, the third sub-question can be answered: *How do freight elements of dry bulk influence the price dynamics?*. The type of dry bulk freight and the quantity transported both have an impact on the price dynamics. It can be concluded that coal transportation is dependent on the BDI due to the underlying correlation to the worldwide economic activity and thus influences the price dynamics. For iron ore, it can be concluded that it follows the same trend but to a lesser extent, as iron ore transportation is also affected by the iron ore price. Also, it can be concluded that the quantity of shipping positively impacts the price dynamics. At last, it can be concluded that there is an effect of freight elements on the price dynamics and that this effect on the freight rate is comparable to the time elements.

7.4. Conclusion sub-question 4

The last sub-question investigated the effects of fleet availability on the price dynamics in the market. This sub-question is examined in three steps, as the first step, the data is transformed and visualized, in the second step a Random Forest is created to determine the variable importance and in the third step a Fixed Effects model is created on the most important variables.

It can be concluded from the scatter plot that the BDI varies significantly more when the total number of ships occupied is larger. The fact that the BDI is more variable during periods of high occupancy shows that there is an effect of fleet availability on the price dynamics. For that reason, the Random Forest model investigated which variables caused this effect.

The Random Forest model, based on 500 trees, is created with the total number of occupied ships as the dependent variable and all relevant variables as independent variables. The algorithm found it is optimal to use five variables to make the splits in the trees. Based on the %incMSE and the IncNodePurity four variables: the days between the contract date and the laycan from date, the DWT, the season of the contract date and the freight rate are implemented as the independent variables in the Fixed Effects model. The Random Forest model also indicated that BDI was the most important variable to predict the total number of ships occupied. Therefore, it can be justified and concluded that the ship occupation is a good indicator to determine the BDI and thus the price dynamics.

At last, a Fixed Effects model is created with the total ships occupied, rate, days between the contract date and laycan from date, DWT, and the seasons of the contract date as independent variables and the BDI as the dependent variable. From the results, it can be concluded that the total number of occupied ships has a significant positive effect on the BDI and thus on the price dynamics. This indicates that more ships' occupation results in a higher BDI and makes sense from an economic point of view. The freight rate also has a significant positive effect on the BDI. It is suggested that a higher ship occupation decreases the available supply and whenever the demand keeps the same, the freight rate increase. This is in line with the sea transport demand and supply function on freight rate (figure 7) of Stopford (2009). This decrease of available supply increases rivalry among charterers to obtain Capesize ships for the bulk transportation requirements and thus shipowners can expect higher charter prices during these times. As indicated earlier in the literature, the freight rate and the BDI are interrelated and an increase in freight rate increases the BDI. This is also confirmed by the Fixed Effect model. From this, it can be concluded that the model is in line with the literature and makes sense from an economic point of view.

The Fixed Effects model shows that the days between the contract date and the laycan from date have a negative effect on the BDI and thus negatively influence the price dynamics. This effect is comparable to the effect in the first sub-question and therefore the same can be concluded. Ships contracting on shorter time notice have more urge to ship bulk from and to a desired location, thus having less margin to wait for a favourable BDI and for that reason the BDI is higher whenever the number of days in between is lower. Also, whenever the BDI is high, charterers could have the chance to renegotiate shorter-term contracts more frequently, which enables them to benefit from rate volatility and obtain better contracting terms.

The seasonality effects of the contract date in this sub-question are in line with the seasonality effects of the laycan from date from sub-question 1. Autumn is used as a reference season and from the model, it can be concluded that closing contracts in autumn increases the BDI and closing contracts during the winter decreases the BDI. The effects for Spring and Summer are not considered significant in this model. To conclude, a lower BDI during the winter suggests there is less trade activity and demand for dry bulk shipping services, while during autumn the opposite effect takes place.

Finally, the last sub-question: *What are the effects of fleet availability on price dynamics in the market?*, can be answered to conclude the final part of the research. First of all, the effects of the fleet availability are according to the models interrelated to the Baltic Dry Index. The fleet availability is in the models expressed as the ship occupation, which is the contrary of the fleet availability. For that reason, the positive effect of ships' occupation on the BDI, has to be reversed in order to answer the sub-question appropriately. To sum up, the effects of the fleet availability are visible from the visualization, confirmed by the Random Forest model and calculated by the Fixed Effect models. From this, it can be concluded that a higher fleet availability decreases the BDI and thus has a negative relation to the price dynamics. This final conclusion from the models is in line with the literature and therefore it can be assumed that the model is giving realistic output.

7.5. Conclusion Research Question

The dry bulk shipping market has grown tremendously over the years. This research investigated time elements, geographical elements, freight elements and the fleet availability in individual spot-rate contracts of Capesize ships on price dynamics. These elements were divided into four sub-questions, all giving answers to a different element of the contract. All these elements of spot-rate contracts are investigated with a cross-sectional time-series analysis to investigate the dry bulk shipping market and its effects on the price dynamics. All elements are investigated with a Fixed Effects model and having answered all the sub-questions, it is possible to answer the main question for this research: *Which elements of individual spot-rate contracts in the dry bulk market can be used to conduct a market analysis on price dynamics?*

In order to conduct a market analysis on price dynamics, time elements, geographical elements and freight elements of individual spot-rate contracts for Capesize ships are insightful to investigate the dry bulk market. The individual spot-rate contracts are also used in this research to outline the overall fleet availability for Capesize ships and provided also relevant insights on the price dynamics. As all elements can be insightful, it is decided to state a conclusion for each element and how it can be used to conduct a market analysis on price dynamics.

Time elements: The contract date and laycan date are considered crucial in understanding the seasonality perspective in the data. Contracting on shorter notice and during autumn and summer results in higher prices, whereas contracting on longer notice during spring and winter may lead to lower prices. For that reason, time elements of contracts must be taken into consideration to conduct a market analysis of the price dynamics.

Geographical elements: The duration of the route has a negative impact and thus longer routes result in lower prices. Specific loading and discharge locations positively affect the prices, such as loading from Australia and the South Pacific compared to the other locations and discharging in the ARA Range and the Mediterranean result in higher prices, compared to the Far East. Therefore, geographical elements play a significant role in influencing price dynamics and must be taken into account in conducting a market analysis.

Freight elements: Different types of cargo such as coal and iron ore, have a varying degree of dependence on the Baltic Dry Index and thus on the prices. Coal transportation is highly dependent on the BDI due to the correlation to the worldwide economic activity, this effect is also detected for iron ore, but to a lesser extent than coal. The quantity of shipping positively affects the price dynamics, indicating that higher prices result in higher prices. As a result, it can be argued that freight elements should be examined to analyse the market and the effects on the price dynamics.

Fleet availability: Fleet availability is inversely related to the Baltic Dry Index. Higher fleet availability leads to a decrease in the Baltic Dry Index and has a negative relation to price dynamics. In other words, when the fleet availability is higher, it results in lower prices. As fleet availability is considered important, the variables influencing fleet availability need to be taken into consideration. The season of the contract date, the DWT and the days between the contract and the laycan from date are important determinants in the ship’s occupation and thus for the fleet availability. For that reason, fleet availability is an important element to take into account when conducting a market analysis of the price dynamics.

To summarize, many individual elements such as the contract dates, laycan dates, the geographic distribution of shipping operations, cargo types, quantity of shipping, and fleet availability should all be taken into account when conducting market analysis on pricing dynamics in the dry bulk market. Understanding the connections between these components can help to get important insights into how the market's prices change for Capesize ships in the dry bulk shipping sector.

In order to do a comprehensive market analysis on price dynamics in the dry bulk shipping market, besides examining these individual spot-rate contract elements, several additional factors need to be taken into account. It is important to take into account that dry bulk shipping affects many other markets and makes a huge impact on the global economy. In both developed and developing countries, it supplies industries, provides energy resources, promotes international trade, and propels economic growth. Thereby, there is also a trade-off effect of the dry bulk market on the global economy, as the global economy and the dry bulk trade market are interdependent. For that reason, it is advisable to examine in addition to the individual elements of spot-rate contracts, the microeconomic and macroeconomic aspects of the dry bulk shipping market to have a more complete market analysis.

In creating a more complete market analysis, microeconomic and macroeconomic aspects such as the other dry bulk vessel sizes and their exchangeability for transporting dry bulk, the supply and demand function of dry bulk and the supply and demand of the transportation of dry bulk and geopolitical events are some examples which must be taken into consideration. From this research, it also emerges that these aspects in return have an impact on the individual elements in the spot-rate contracts for Capesize vessels. From this, it can be concluded that time, geographical and freight elements from the individual spot-rate contracts can be used to conduct a market analysis on price dynamics, but to obtain a more comprehensive market analysis, more microeconomic and macroeconomic aspects need to be considered in addition while conducting a market analysis on price dynamics in the dry bulk shipping market.

8. Limitations

The research provides some relevant insights into elements of individual contracts and their effects on price dynamics. However, this chapter discusses some limitations to the research and suggestions for further research.

First of all, there are some limitations due to existing data and the lack of data availability. The main component of the data was obtained from the brokerage platform Clarkson. The spot-rate contract data is not open-source on the internet and must be obtained through a broker, such as Clarkson. Clarkson did issue contract data for the years 2020, 2021 and 2022. Albert Veenstra contacted Clarkson if it was possible to obtain more years of data, however, Clarkson answered that this was not possible. The Erasmus University of Rotterdam has only a contract with one brokerage platform (Clarkson), and therefore the decision of Clarkson limited this research to the usage of only three years of data. It is an extra limitation that in the three years of data, there was a huge spike in 2021 in the Baltic Dry Index. This makes the models less representative by having only 3 years of data to straighten out a spike as this.

Examining the datasets from Clarkson's database, it is striking that the contracts are not filled in carefully. Most datasets contain many missing values or incorrect values on a row and column basis. This generally makes it difficult to create models to investigate the data. Especially the 'name' column of the boats is not filled in carefully, which makes it impossible to know exactly how many different boats are included in each category by the dataset of Clarkson. Without knowing that, it is also difficult to draw general conclusions for the entire fleet. In the used dataset of the spot-rate contracts for Capesize ships, 609 distinctive Capesize ships are specified. As stated in the literature context, there are 1593 boats in the Capesize fleet, but it is impossible to figure out, due to the large number of TBNs in the dataset, what percentage of the total fleet is covered by spot-rate contracts and by Clarkson in total. The lack of completed data is a limitation of this research, which is beyond my control.

This research focused only on the $/Tonne contracts, due to a lack of sufficient data on $/Day contracts for Capesize ships. This might have a major influence on the thoroughness and correctness of the market analysis of pricing trends in the dry bulk market. The scope and depth of the market study on price developments are constrained due to the insufficient data on $/Day contracts of Capesize ships. To analyse the interaction between cargo volume-based pricing and time charter rates, as well as how various factors impact prices in different contract categories, it is imperative to collect data from both $/Tonne and $/Day contracts.

Another limitation that is partly within my control is that data was only used for the Capesize ships. This was specifically chosen because this dataset was more complete than the Panama/Panamax, Handymax and Handysize ships. By researching all the ship categories, it would be possible to make conclusions per category and compare the results. However, this is deliberately chosen not to do for two reasons. First, is the lack of data in the other ship categories, which would result in fewer observations and thus lead to worse-performing models. Second, is time constraints, which would make it unfeasible for this study to examine and compare all types of ships with each other. This could result in more comprehensive research, allowing conclusions to be drawn by vessel size and possibly declaring and detecting interchangeability between vessel sizes.

It's important to remember that the BDI is impacted by several factors, including the availability and demand for Capesize ships, as well as other classes of dry bulk carriers and the global shipping industry as a whole. As a result, a variety of variables, such as world economic circumstances, geopolitical developments, and shifts in commodity demand, can affect variations in the BDI and are not fully represented in the models. It is a limitation of the models, as not all factors can be accounted for in the real world.

There are in total 146 load and discharge locations within the dataset. These variables seem to have been filled in with care and appear to be mostly correct. However, since there is no ‘arrival date’ within the dataset, it is not possible to determine directly how long the ship has travelled between the load and discharge location. This consequently had to be calculated by hand via a Maritime Distance Calculator, due to the fact. that there is no open-source data for the distance and duration for Capesize ships to travel from location A to location B. Apart from the fact that this is a lot of work and takes a very long time, it is a limitation that it is vulnerable to human mistakes. Eventually, the routes had to be calculated by hand, however, it has not been possible to complete all routes from load to discharge locations, due to time reasons. As both load and discharge have 146 distinctive locations, the total number of possible routes comes to a total of 146! (146 factorial).

For that reason, a categorisation is made based on continent parts (suggested by Albert Veenstra) to reduce manual work in the Maritime Freight Rate calculator. As a result, the number of shipping days for most contracts is eventually searched by location or by categorization. It is decided that whenever the route is used at least 10 times in the dataset from location to location and after that for the other contracts from categorization to categorization, to search the shipping days by hand in the Maritime Distance Calculator. This is a limitation of the study, as the results would be better if all distinctive routes had the correct value and not a value based on the categorization or a missing value.

Along with this is the next limitation, as already discussed in the data chapter, Capesize ships do not all sail at the same speed. As stated in the literature context, the speed at which ships sail depends on a lot of factors. These factors and the final speed of ships cannot be ascertained from the data from Clarkson, but the Maritime Distance Calculator needs a sailing speed to calculate the sailing time of the route. Because of this, an average speed was taken from literature for Capesize ships and not route-specific, time-specific and cargo-specific characteristics were implemented in the speed of the vessels, which would adjust the sail duration and thus the end date. Thereby, there is no clear departure date of the vessels indicated in the data provided by Clarkson. As a result, the start and end dates are both based on theory and not practise, which is considered a limitation of the research.

Connecting to this limitation is the fact that no consideration is made for the start date where the boat comes from. For example, if the boat is already in the right port to load and it is not busy, then the boat will be able to load relatively at the beginning of the laycan period and thus leave earlier than a boat that has to travel far first. This is not clear from the data provided by Clarkson. This is also a major limitation in studying fleet availability and possible sailing patterns.

The sailing patterns should also consider that this is bulk only. Here, bulk goods are moved from A to B to often make goods in manufacturing countries, like China for example. The models in this research do not take into account that there are import and export countries of these products. As an example, 70% of the ships in the dataset go to Qingdao to unload, which is realistic in the real world, but that does mean the model is biased by the fact that the data is not normally distributed. It could also be interesting to see where all these products go again in the flow of goods from China by container ships, for example, this would give a clearer picture of overall shipping activity distribution.

In conclusion, despite the fact that the research offers insightful knowledge about certain contract components and their influence on price dynamics in the dry bulk market, there are a number of important limitations that must be noted. These restrictions are due to the lack of data, the poor quality of the data, and the study's width, all of which have an impact on how accurate and complete the market research is. It is essential to evaluate the research findings carefully in light of these constraints. By addressing these issues and doing more research using more thorough data, such as $/Tonne and $/Day contracts for different ship classifications, it may be possible to produce a more thorough and accurate market study on the pricing trends in the dry bulk market. Additionally, taking into account more extensive aspects like import-export flows and precise sailing speed characteristics will improve our comprehension of market dynamics and shipping patterns. Despite its limitations, this research offers a useful beginning point and lays the groundwork for further research in this area.

Appendix A – Additional information

A.1. Type of dry bulk ships

This section describes the differen types of dry bulk ships. As stated in the main text, deadweight tonnes (DWT) describes the weight of cargo, fuel, water and other supplies that the vessel can carry. The DWT is frequently used to describe the size of dry bulk carriers and according to their DWT ranges, they are divided into major groups. The following categorization is frequently used: Handysize, Handymax/Supramax, Panamax, Post-Panamax, Capesize and Vloc and the DWT range from 10,000 to over 400,000. From this, the following categories are discussed in this part of the appendix.

A.1.1. Handysize

The smallest vessels are the Handysize vessels, which carry between 10,000 and 40,000 DWT. Smaller amounts of grain and bauxite, as well as a variety of minor bulks, are typically carried by Handysize vessels on short-haul trading routes. Some of them also have the tools necessary to load and unload cement and wood. These vessels can operate well in smaller ports with length and draft restrictions (Plomaritou & Papadopoulos, 2017).

A.1.2. Handymax/Supramax

Handymax/Supramax vessels can carry between 40,000 and 65,000 DWT. Larger Handymax vessels often follow the trading routes and preferences of Panamax vessels, but because they have ship gear, they can also carry a wider range of commodities. This is mostly due to the fact that they conduct business in ports devoid of sophisticated terminals (Over, 2017). The Handymax industry participates in a significant number of globally distributed trades, primarily transporting grains and small quantities. In order to load and unload goods in countries and ports with poor infrastructure, ships under 60,000 DWT are frequently equipped with onboard cranes (Kapetanis et al., 2014). The Handysize and Handymax vessels are considered somewhat smaller and are often categorized as Handyclass vessels.

A.1.3. Panamax and Post-Panamax

As the names suggest, the Panamax and Post-Panamax vessels are defined by certain conditions to transit through the Panama Canal. The maximum width, also called the beam, of the ship is 32,2 meters. Panamax vessels have a DWT capacity between 65,000 and 85,000, while Post-Panamax vessels range from 85,000 and 100,000 DWT (Plomaritou & Papadopoulos, 2017). The Post-Panamax vessels came after the expansion of the Panama Canal, which made it possible to carry more DWT through the locks. These ships transport a variety of dry bulk products, including coal, grain, iron, and to a lesser degree, minor bulks. With various freight rates for each trading route, these ships frequently transact on medium- and long-haul routes that traverse through the Atlantic Basin, Pacific Ocean, and Indian Ocean (Over, 2017; USDA, 2020).

A.1.4. Capesize and Very Large Bulk/Ore Carrier

As the industry is evolving, the size of the vessels is evolving too. The conventional definition of Capesize bulk carriers in terms of deadweight tonnage has changed from 80,000 to above 100,000 (Darie et al., 2013). Since they are too large and too wide to travel via the Panama Canal, they must travel through the Cape of Good Hope. Due to the fact that these vessels have to sail past Cape of Good Hope, the name Capesize was born. These ships are divided into Small Capes that can carry from 100,00 to 130,000 DWT, Normal Capes that can carry from 130,000 to 200,000 DWT, and Large Capes, also known as Very Large Bulk/Ore Carriers which can carry over 200,000 DWT and can carry up to 400.000 DWT (Plomaritou & Papadopoulos, 2017). The Capesize industry concentrates on the long-distance trading routes for coal and iron ore and as a result, 75% of iron ore is carried by Capesize ships (Comtoi & Lacoste, 2021). Only a relatively small number of ports worldwide have the facilities to handle Capesize vessels because of their size. For that reason, the large Capesize vessels often serve the route Brazil-China and therefore they are also known as Chinamax or Valemax vessels (Over, 2017).

To summarize, DWT is used to describe the size of the carriers and according to that it is divided into six major groups ranging from 10,000 to 400,000+ DWT. Now that it is clear which major categories of vessels there are, what they transport and where they operate globally, it is necessary to look at the sale and purchase market of the vessels.

A.2. Additional information on the ship sale and purchase market

The maritime industry's sale and purchase market of dry bulk vessels entails deals in which shipowners, operators, or investors buy or sell dry bulk carriers for a variety of reasons, including fleet development, fleet replacement, or investment goals. Three parties make up the sale and buy market: the new building market, the used market, and the scrapping market, which are further separated into several types of vessels (Stopford, 2009; Over, 2017).

There are various variables that might impact the sale and purchase market of dry bulk vessels, such as supply and demand patterns, freight costs, market attitude and world economic situations. Alizadeh & Nomikos (2003) found evidence that price fluctuations can be used to anticipate trading volume, which implies that bigger capital gains stimulate more market activity. Additionally, they contend that increased trading activity lowers market volatility since volume has a negative impact on the volatility of price changes. The unusual foundations of the shipping market, such as thin trading, which posits that an increase in trading activity results in price transparency and stability, may be used to explain this (Alizadeh & Nomikos, 2003; Lun & Quaddus, 2009). In sum, the sale and purchase market consist of three different parties, the new building market, the second-hand market and the scrapping market.

A.2.1. New building market

The requirement for more sea transport capacity is reflected in the demand for new vessels (Wright, 1991). Investors in shipping should consider their expectations for future freight rates before placing an order for a ship because the time it takes for a new ship to enter service in the freight market is often several years. Increased demand for maritime transportation would result in higher freight rates and a faster rate of investment in new ships. However, the built time of these new ships could be seen as a barrier to excessive profits, because a higher demand will result in higher prices of new constructing vessels (Dikos, 2004).

Ship operators expand their fleets by acquiring new ships to enhance the availability of sea transport during a time of high freight rates. Shipbuilders would raise the price for new building vessels in response to the increase in demand for new vessels following the hike in freight rates. As a result, freight rates may be used to determine the cost of new constructing boats (Leach, 2004; Lun & Quaddus, 2009).

The cost of creating a new ship becomes important in terms of return on investment even though buying a new ship involves a sizable financial commitment. Low new building prices are typically preferred by fleet owners, however, fleet owners are more likely to demand new ships when the freight rate is high and which results in high prices. For that reason, investors choose to invest in new ships with the intention of reselling them for profit to fleet owners when they believe the price is low (McConville, 1999). The new vessel demand is impacted by the new construction of the pricing structure of the newly built vessel industry (Tsolakis et al., 2003).

A.2.2. Second-hand market

For second-hand vessels, there are a few exceptions that they have been sold with an active time charter contract, as these are typically sold without any mortgages or long-term charter commitments from the previous owner. This market's high price volatility is dependent on both the overall demand and supply for vessels as well as other shipping markets (Stopford, 2009). The revenue generated by the freight market is the main source of funding for the bulk shipping industry. The market for used ships emerges as a substitute supply of ships during a freight boom even though it takes a few years to build a new vessel. Beenstock (1985) proposed that both new and used ships might be used interchangeably because they are the same assets with the exception of their age (Tsolakis et al., 2003).

Adland & Koekebakker (2004) evaluated the Efficient Market Hypothesis for second-hand dry bulk ship market and the empirical findings reveal that trading rules are typically not able to generate surplus value over the buy-and-hold benchmark after taking transaction costs and price slippage into account. The efficiency of trading strategies based on a blend of technical trading principles and fundamental research in the sale and buy market for used dry bulk ships is examined by Alizadeh and Nomikos

(2007). They discovered data showing that trading strategies based on earnings-price ratios significantly outperform buy-and-hold strategies in the market for used ships, particularly in the market for bigger vessels, as a result of heightened volatility in these markets (Alizadeh & Nomikos. 2007).

The second-hand market makes reallocation of ships among ship operators possible and additionally, by lowering market departure costs where ship owners can sell their old ships when they leave the business. The second-hand ship market helps to increase the efficiency of the overall shipping market, by enabling potential investors to purchase old ships and enter the maritime sector (Lun & Quaddus, 2009).

A.2.3. Scrapping market

As stated above, the second-hand vessel sale and purchase market are cyclical and highly competitive due to price movements. The market for vessels which have outlived their useful lives and for that reason sold in scrapyards is known as the demolition market. Vessels that are past their operating age, are sold to scrapyards or intermediates that eventually sell the metals it holds (McConville, 1999). The vessel scrapping price represents the lowest cost of a used vessel, similar to the cost of used vessels, the cost of scrap vessels typically tracks changes in the freight market. The demand for scrap is mostly influenced by the price and demand for steel in the nations that produce it (Over, 2017).

Ship owners calculate and forecast the future operational profitability of the ship and compare it to its own financial situation before deciding whether to scrap a ship. Most often, older ships are typically supplied to the scrap market based on the relatively high scrap value (Farthing and Brownrigg, 1997). The expectation of ship owners regarding potential trading opportunities influences the choice to scrap if the ship's profitability is negative, the ship will be scrapped. The demand for ships in the wrecking industry is influenced by the scrap price and influences the total fleet size (Lun & Quaddus, 2009).

In sum, the sale and purchase market consist of three different parties, the new building market, the second-hand market and the scrapping market. The sales and purchase conditions are impacted by various factors such as the freight rate, the prices of new ships, the investors' landscape, the second-hand market, the scrap market and other world economic factors.

Appendix B – Additional data

Table B.1. Different size vessels on major routes

	Major Routes		
	Iron Ore	**Coal**	**Grain**
Capesize	Brazil to Western Europe	East Australia to Far East	Argentina to Near East
	Brazil to Far East	East Australia to Japan	Argentina to East Europe
	West Australia to Western Europe	East Australia to Western Europe	River Plate to Near East
	West Australia to Far East	South Africa to Western Europe	River Plate to East Europe
		South Africa to Far East	
Panamax	Brazil to Western Europe	North America to Japan	North America to Far East
	Brazil to Far East	North America to Western Europe	North America to Western Europe
	Australia to Western Europe	East Australia to Far East	North America to Near East
	Australia to Far East	East Australia to Japan	
		East Australia to Western Europe	
Handyclass	India to Far East	South Africa to Far East	Australia to Far East
	Canada to United States	South Africa to Europe	Australia to Japan
	Canada to Japan		Australia to Middle East
	Liberia to Western Europe		North America to Africa
	Mauritania to Western Europe		North America to West Europe

Source: Grammenos 2013 page 323

Table B.2. Load to Discharge routes with frequency

Load location	Discharge location	Frequency
Dampier	Qingdao	684
Port Hedland	Qingdao	455
Tubarao	Qingdao	396
Saldanha Bay	Qingdao	112
Wes Australia	Qingdao	81
CSN	Qingdao	59
Sudeste	Qingdao	43
Teluk Rubiah	Qingdao	41
Seven Island	Qingdao	37
West Australia	Qingdao	37
Hedland	Qingdao	36
New Castle	Boryeong	23
Seven Island	Oita	17
Porto Sudeste	Qingdao	17
NewCastle	Hadong	15
Saldahna Bay	Rotterdam	15
Itaguai	Qingdao	14
NewCastle	Younghueng	13
ACU	Qingdao	12
Port Cartier	Qingdao	11
West Australia	Singapore - Japan	11
Newcastle	Dangjin	10
NewCastle	Mailiao	10
Pacific	Singapore - Japan	10

Table B.3. Load Category to Discharge Category routes with frequency

Load Category	Discharge Category	Frequency
Australia	Far East	97
Africa	Far East	81
North America East Coast	Far East	44
ARA Range	Far East	23
South America East	ARA Range	21
Australia	Other	20
North America East Coast	ARA Range	20
North Europe	ARA Range	20
Africa	ARA Range	18
South America East	Far East	18
Africa	Indian Subcontinent	17
Middle East	Far East	14
South America East	Middle East	14
South Pacific	Far East	14
Far East	Far East	13
South America East	Mediterranean	13
Australia	Middle East	12
North America East Coast	Indian Subcontinent	11
North Europe	Middle East	11
Australia	Indian Subcontinent	10
Australia	South Pacific	10
North America West Coast	Far East	10
North Europe	Far East	10

Table B.4. Ports used for categorization

Category	Port used
Australia	Melbourne
Africa	Capetown
North America East Coast	Cartier
ARA Range	Rotterdam
South America East	San Nicolas
North Europe	Narvik
Middle east	Sohar
South pacific	Tanjung Bungah
Far East	Qingdao
Other	NA
Indian Subcontinent	Mormugao
Mediterraenean	Taranto
North America West Coast	Vancouver

References

Ådland, A. O., & Koekebakker, S. (2004). Market efficiency in the second-hand market for bulk ships. *Maritime Economics & Logistics*, *6*, 1-15.

Adland, R., & Cullinane, K. (2005). A Time-Varying Risk Premium in the Term Structure of Bulk Shipping Freight Rates. *Journal of Transport Economics and Policy.*

Adland, R., & Jia, H. (2008). Charter market default risk: A conceptual approach. *Transportation Research Part E-logistics and Transportation Review.*

Adland, R., Jia, H., & Strandenes, S. P. (2016). An Empirical Analysis of Vessel Capacity Utilization. In *International Forum on Shipping, Ports and Airports (IFSPA) 2015: Empowering Excellence in Maritime and Air Logistics: Innovation Management and TechnologyHong Kong Polytechnic University*.

Adland, R., Cariou, P., & Wolff, F. C. (2016). The influence of charterers and owners on bulk shipping freight rates. *Transportation Research Part E: Logistics and Transportation Review*, *86*, 69-82.

Adland, R., & Jia, H. (2017). Simulating physical basis risks in the Capesize freight market. *Maritime Economics & Logistics*, *19*, 196-210.

Adland, R., Jia, H., & Strandenes, S. P. (2018). The determinants of vessel capacity utilization: The case of Brazilian iron ore exports. *Transportation Research Part A: Policy and Practice*, *110*, 191-201.

Ahn, S. C., & Low, S. (1996). A reformulation of the Hausman test for regression models with pooled cross-section-time-series data. *Journal of Econometrics*, *71*(1-2), 309-319.

Alizadeh, A. H., & Nomikos, N. K. (2003). The price-volume relationship in the sale and purchase market for dry bulk vessels. *Maritime Policy & Management*, *30*(4), 321-337.

Alizadeh, A. H., & Nomikos, N. K. (2007). Investment timing and trading strategies in the sale and purchase market for ships. *Transportation Research Part B: Methodological*, *41*(1), 126-143.

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

Alizadeh, A. H., & Talley, W. K. (2011). Microeconomic determinants of dry bulk shipping freight rates and contract times. *Transportation*, *38*, 561-579.

Alizadeh, A. H., & Talley, W. K. (2011). Vessel and voyage determinants of tanker freight rates and contract times. *Transport Policy*, *18*(5), 665-675.

Axarloglou, K., Visvikis, I., & Zarkos, S. (2013). The time dimension and value of flexibility in resource allocation: The case of the maritime industry. *Transportation Research Part E: Logistics and Transportation Review*, *52*, 35-48.

Bakshi, G., Panayotov, G., & Skoulakis, G. (2010). The Baltic Dry Index as a predictor of global stock returns, commodity returns, and global economic activity. *Commodity Returns, and Global Economic Activity (October 1, 2010)*.

Barrot, J., Loualiche, E., & Sauvagnat, J. (2018). The Globalization Risk Premium. *International Trade eJournal.*

Beck, N., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American political science review*, *89*(3), 634-647.

Beenstock, M. (1985). A theory of ship prices. *Maritime Policy and Management*, *12*(3), 215-225.

Beenstock, M., & Vergottis, A. (1993). *Econometric modelling of world shipping*. Springer Science & Business Media.

Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, *3*(1), 133-153.

Biau, G. (2012). Analysis of a random forests model. *The Journal of Machine Learning Research*, *13*(1), 1063-1095.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research synthesis methods*, *1*(2), 97-111.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Cepowski, T. (2019). Regression formulas for the estimation of engine total power for tankers, container ships and bulk carriers on the basis of cargo capacity and design speed. *Polish Maritime Research*, *26*(1 (101)), 82-94.

Chatfield, C. (2001). Prediction intervals for time-series forecasting. *Principles of forecasting: A handbook for researchers and practitioners*, 475-494.

Chen, S. (2011). Modelling and forecasting in the dry bulk shipping market. *Unpublished Ph. D dissertation, Delft University of Technology.*

Clarksons (2023). Home page. https://www.clarksons.com/

Coghlin, T., Baker, A., Kenny, J., & Kimball, J. (2014). *Time charters*. CRC Press.

COMTOI, C., & Lacoste, R. (2021). Dry bulk shipping logistics. *Maritime Logistics: A Guide to Contemporary Shipping and Port Management*, 241.

Cooke, J., Young, T., Ashcroft, M., Taylor, A., Kimball, J., Martowski, D., ... & Sturley, M. (2014). *Voyage charters*. CRC Press.

Cryer, J. D., & Kellet, N. (1991). *Time series analysis*. Royal Victorian Institute for the Blind. Tertiary Resource Service.

Cullinane, K. P. B., Mason, K. J., & Cape, M. (1999). A comparison of models for forecasting the Baltic freight index: Box-Jenkins revisited. *International journal of maritime economics*, *1*, 15-39.

Darie, I., Roerup, J., & Wolf, V. (2013, October). Ultimate strength of a cape size bulk carrier under combined global and local loads. In *Proc 12th Symp Practical Design of Ships and Other Floating Structures* (pp. 1173-1180).

Da Silva, J. G. C. (1975). *The analysis of cross-sectional time series data* (Doctoral dissertation, North Carolina State University).

DeFries, J. C., & Fulker, D. W. (1985). Multiple regression analysis of twin data. *Behavior genetics*, *15*(5), 467-473.

Diakodimitris, E. N. (2019). Evaluation of Forecasting Approaches for Crude Oil Tanker Freight Rates.

Dikos, G. (2004). New building prices: demand inelastic or perfectly competitive?. *Maritime Economics & Logistics*, *6*(4), 312-321.

Dobre, F. (2016). Cargo Contract. *Risk in Contemporary Economy*, 70-74.

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

Eriksen, I. E. (1983). The demand for bulk ship services.

Farthing, B., & Brownrigg, M. (1997). *Farthing on international shipping*. Informa Pub.

Fevre, C. (2018). A review of demand prospects for LNG as a marine transport fuel.

Fischer, T., Krauss, C., & Treichel, A. (2018). *Machine learning for time series forecasting-a simulation study* (No. 02/2018). FAU Discussion Papers in Economics.

Gao, J., Wang, J., & Liang, J. (2023). A unified operation decision model for dry bulk shipping fleet: ship scheduling, routing, and sailing speed optimization. *Optimization and Engineering*, 1-24.

German, H. (2005). Commodities and commodity derivatives: Modeling and pricing for agriculturals, metals and energy.

Geman, H., & Smith, W. O. (2012). Shipping markets and freight rates: an analysis of the Baltic Dry Index. *The Journal of Alternative Investments*, *15*(1), 98-109.

Grammenos, C. (Ed.). (2013). *The handbook of maritime economics and business*. Taylor & Francis.

Guan, F., Shen, X., Wu, L., Yu, Y., Sun, D., & Yang, Y. (2019). Fleet route selection prediction problem based on support vector machine. *Advances in Mechanical Engineering*, *11*(4), 1687814019836857.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251-1271.

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *The Econometrics Journal*, *3*(2), 148-161.

Holguín-Veras, J., Sánchez, I., González-Calderón, C., Sarmiento, I., & Thorson, E. (2011). Time-Dependent Effects on Parameters of Freight Demand Models: Empirical Investigation. *Transportation research record*, *2224*(1), 42-50.

Jacks, D., & Stuermer, M. (2021). Dry Bulk Shipping and the Evolution of Maritime Transport Costs, 1850-2020. Federal Reserve Bank of Dallas, Working Papers.

Jonnala, S., Fuller, S., & Bessler, D. (2002). A GARCH approach to modelling ocean grain freight rates. *International Journal of Maritime Economics*, *4*, 103-125.

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

Kanamoto, K., Murong, L., Nakashima, M., & Shibasaki, R. (2021). Can maritime big data be applied to shipping industry analysis? Focussing on commodities and vessel sizes of dry bulk carriers. *Maritime Economics & Logistics*, *23*, 211-236.

Kang, D., & Woo, S. (2017). Liner shipping networks, port characteristics and the impact on port performance. Maritime Economics & Logistics

Kapetanis, G. N., Gkonis, K. G., & Psaraftis, H. N. (2014). Estimating the operational effects of a bunker Levy: The case of Handymax Bulk Carriers. In *Transport Research Arena (TRA) 5th Conference: Transport Solutions from Research to DeploymentEuropean CommissionConference of European Directors of Roads (CEDR) European Road Transport Research Advisory Council (ERTRAC) WATERBORNE^TP European Rail Research Advisory Council (ERRAC) Institut Francais des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux (IFSTTAR) Ministère de l'Écologie, du Développement Durable et de l'Énergie*.

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

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

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

Koekebakker, S., Adland, R., & Sødal, S. (2007). Pricing freight rate options. *Transportation Research Part E: Logistics and Transportation Review*, *43*(5), 535-548.

Köhn, S., & Thanopoulou, H. (2011). A gam assessment of quality premia in the dry bulk time–charter market. *Transportation Research Part E: Logistics and Transportation Review*, *47*(5), 709-721.

Koopmans, T. C. (1939). *Tanker freight rates and tankship builiding: An analysis of cyclical fluctuations* (No. 27). De erven F. Bohn nv.

Laulajainen, R. (2007). Dry bulk shipping market inefficiency, the wide perspective. *journal of Transport Geography*, *15*(3), 217-224.

Leach, P. T. (2004). If it moves, charter it. *Traffic World*.

Limao, N., & Venables, A. J. (2001). Infrastructure, geographical disadvantage, transport costs, and trade. *The world bank economic review*, *15*(3), 451-479.

Lun, Y. V., & Quaddus, M. A. (2009). An empirical model of the bulk shipping market. *International Journal of Shipping and Transport Logistics*, *1*(1), 37-54.

Lyridis, D., Manos, N., & Zacharioudakis, P. (2014). Modeling the Dry Bulk Shipping Market Using Macroeconomic Factors in Addition To Shipping Market Parameters Via Artificial Neural Networks. *International Journal of Transport Economics*

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PloS one*, *13*(3), e0194889.

Maritime Page (2023). Capesize Bulk Carriers: A complete overview of giant ships. https://maritimepage.com/capesize-bulk-carriers-a-complete-overview-of-giant-ships/

McConville, J. (1999). *Economics of maritime transport: theory and practice*. Witherby.

Michail, N. A. (2020). World economic growth and seaborne trade volume: quantifying the relationship. *Transportation Research Interdisciplinary Perspectives*, *4*, 100108.

Michail, N. A., & Melas, K. D. (2020). Quantifying the relationship between seaborne trade and shipping freight rates: A Bayesian vector autoregressive approach. *Maritime Transport Research*, *1*, 100001.

Michail, N. A., & Melas, K. D. (2021). Market interactions between agricultural commodities and the dry bulk shipping market. *The Asian Journal of Shipping and Logistics*, *37*(1), 73-81.

Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.

Over, P. C. (2017). Dry cargo chartering. *London: Institute of Chartered Shipbrokers*.

Peters. WDJ. (2022). The impact of offering different payment methods on marketplaces and platforms from a Payment Service Provider perspective.

Plomaritou, E., & Papadopoulos, A. (2017). *Shipbroking and chartering practice*. Taylor & Francis.

Pirrong, S. (1993). Contracting Practices in Bulk Shipping Markets: A Transactions Cost Explanation. *The Journal of Law and Economics.*

Poblacion, J. (2015). The stochastic seasonal behavior of freight rate dynamics. *Maritime Economics & Logistics*, *17*, 142-162.

Regli, F., & Nomikos, N. K. (2019). The eye in the sky–Freight rate effects of tanker supply. *Transportation Research Part E: Logistics and Transportation Review*, *125*, 402-424.

Roels, G. (2013). Risk Premiums in the Price-Setting Newsvendor Model. *IO: Theory eJournal.*

Sarkar, S., & Hong, G. (2004). Effective duration of callable corporate bonds: Theory and evidence. *Journal of banking & finance*, *28*(3), 499-521.

Searates (2023). Maritime Distance Calculator - Online Freight Shipping & Transit Time Calculator https://www.searates.com/services/distances-time/

Shibasaki, R., Azuma, T., Yoshida, T., Teranishi, H., & Abe, M. (2017). Global route choice and its modelling of dry bulk carriers based on vessel movement database: Focusing on the Suez Canal. *Research in transportation business & management*, *25*, 51-65.

Shor, B., Bafumi, J., Keele, L., & Park, D. (2007). A Bayesian multilevel modeling approach to time-series cross-sectional data. *Political Analysis*, *15*(2), 165-181.

S&P Global Market Intelligence (2021). Dry bulk ships speed up to cash in on bullish market; CO2 emissions a concern.https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/agriculture/091621-dry-bulk-ships-speed-up-to-cash-in-on-bullish-market-co2-emissions-a-concern

S&P Global Market Intelligence (30-11-2021). Shipping market outlook 2022 Container vs Dry Bulk.https://www.spglobal.com/marketintelligence/en/mi/research-analysis/shipping-market-outlook-2022-container-vs-dry-bulk-.html

Stopford, M. (2009). Maritime Economics.

Strobl, C., Boulesteix, A. L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC bioinformatics*, *8*(1), 1-21.

Tan, G., Cheong, C., & Zurbruegg, R. (2019). National culture and individual trading behavior. Journal of Banking & Finance.

Thorburn, T. (1960). Supply and demand of water transport: studies in cost and revenue structures of ships, ports and transport buyers with respect to their effects on supply and demand of water transport goods.

Trading Economics (2023). The Baltic Dry Index. https://tradingeconomics.com/commodity/baltic

Triantafillou, N. (2020). EVALUATION OF UNIVARIATE AND MULTIVARIATE FORECASTING METHODS FOR THE DRY BULK FREIGHT RATES.

Truett, L. J., & Truett, D. B. (1998). *Managerial economics: analysis, problems, cases*. Thomson South-Western.

Tsolakis, S. D., Cridland, C., & Haralambides, H. E. (2003). Econometric modelling of second-hand ship prices. *Maritime Economics & Logistics*, *5*(4), 347-377.

UNCTAD (2010). Trade and development report 2010. https://unctad.org/system/files/official-document/tdr2010_en.pdf

UNCTAD (2019). Trade and development report 2019: Financing a global green new deal. https://unctad.org/system/files/official-document/tdr2019_en.pdf

USDA (2020). United States Department of Agriculture 2020: Bulk Vessel Fleet. *United States Department of Agriculture: Agricultural Marketing Service.* https://agtransport.usda.gov/stories/s/Bulk-Vessel-Fleet-Size-and-Rates/bwaz-8sgs/

Veenstra, A. W. (1999). *Quantitative analysis of shipping markets = Kwantitatieve analyse van Scheepvaartmarkten*.

Veenstra, A., & van Dalen, J. (2008). Price indices for ocean charter contracts. In *The Second International Index Measures Congress, Washington, Digital proceedings*.

Verbeek, M. (2008). Pseudo-panels and repeated cross-sections. In *The Econometrics of Panel Data: Fundamentals and recent developments in theory and practice* (pp. 369-383). Berlin, Heidelberg: Springer Berlin Heidelberg.

Wada, Y., Hamada, K., Hirata, N., Seki, K., & Yamada, S. (2018). A system dynamics model for shipbuilding demand forecasting. *Journal of Marine Science and Technology*, *23*, 236-252.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.

Wright, G. (1991). Freight rates in the tramp shipping market. *International Journal of Transport Economics/Rivista internazionale di economia dei trasporti*, 47-54.

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

Yunianto, I. T., Lazuardi, S. D., & Hadi, F. (2018). Freight calculation model: a case study of coal distribution. In *IOP Conference Series: Earth and Environmental Science* (Vol. 135, No. 1, p. 012013). IOP Publishing.

Zannetos, Z. S. (1966). The Theory of Oil Tankship Rates: An Economic Analysis of Tankship Operations (No. 4).

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