NTNU
Norwegian University of
Science and Technology

Forecasting Time Charter Equivalent Oil Tanker Freight Rates

- determinant driven, route-specific Markov
regime-switching models

Eivind Hjertaker Molvik
Martin Salte Stafseng

Industrial Economics and Technology Management
Submission date: April 2018
Supervisor: Sjur Westgaard, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

Problem Description

The tanker shipping segment can be considered cyclical and volatile. It involves huge capital investments and exposes stakeholders to great amounts of risk. The total tanker fleet consists of approximately 14,500 ships which transport several billion tonnes of oil and gas each year among other things. They play a signifianct role in the transport of liquid commodities, allowing for high-volume, low-cost shipping in bulk. These tankers are generally categorized by size, e.g., VLCC, Suezmax, Aframax and Panamax, whereof some operate certain routes with a greater frequency than others. Core routes are those supporting the most essential flow between major markets. These routes are associated with everything from varying geography, geopolitics and trade flows, to different climate, currencies and distances.

In this study we will look deeper into the routes concerning oil tankers. A selection of a few specific routes will be made, and these will be examined extensively. The intricacies of these routes will explored meticulously from well-established theory, articles, related fields and expert knowledge. This will be the basis for finding determining factors for a forecasting model of the route specific oil tanker freight rate. A Markov regime-switching regression model will be pursued in the attempt of constructing a suitable generalized model, accounting for seasonality, lag and global factors among other things. The model will be rigorously benchmarked and holdout-validated to ensure operationality.

Preface

This study is written as our Master's thesis and concludes our university degrees in Industrial Economics and Technology Management at the Norwegian University of Science and Technology (NTNU).

Shipbuilding and shipping are among the oldest industries in Norway. Today, Norway has one of the largest and most comprehensive maritime sectors globally, both within traditional shipbuilding, offshore petroleum industry, shipping, as well as industrial fishing and aquaculture. In the research project "A Knowledge Based Norway" by Reve and Sasson (2012), the future prospects for 13 of Norway's most important business and industry sectors were analysed. The study concluded that the maritime sector is Norway's only global competence based industry and an industry where Norway has the industrial competitiveness to succeed internationally. This has lead us to pursue areas of research within this industry.

We would like to extend our sincere gratitude to Clarksons Platou AS for granting us access to data and publications in the Shipping Intelligence Network (SIN) and the World Fleet Register (WFR) databases (Clarksons Research Services Limited, 2017). We further want to express our gratitude to our supervisor, Professor Sjur Westgaard at the Department of Industrial Economics and Technology Management, as well as the Department of Industrial Economics and Technology Management.

Abstract

In this thesis, we address the issue of explaining and forecasting oil tanker freight rates for specific tanker routes. These freight rates are known to exhibit periods of extreme volatility. To predict the freight rates, we therefore utilise a Markov regime-switching multiple regression model with two states - a normal state, and a volatile state. This thesis hence postulates that predictions of short-term freight rates can be improved through a framework that can capture the distinctive nature of freight rates by switching between two regimes, while combining this with hypothetically superior route-specific and global determinants. We make a substantial attempt to combine market domain knowledge with statistical methods. Our approach to doing so is twofold.

Firstly, motivated by the findings in the existing literature, the observations of structural breaks, and the plethora of attempts at modelling the freight rate, we characterise the market. Six tanker routes, TD1, TD3, TD7, TD12, TC1, and TC2, are investigated. An extensive assessment of the key determining factors of the freight rates is given. A candidate predictor analysis is done, based on subsampling in combination with a selection algorithm. By using the novel approach of stability selection and LASSO penalization with a random tuning parameter, we are able to rank the factors based on their potential modelling importance.

Secondly, we develop a Markov regime-switching regression model for one-month ahead forecasting of the freight rates on the selected routes. The data is tested for structural breaks using a Chow test, and indications of multiple regimes are found. With a subset of variables for each route, we formulate two-regime regression models with switching coefficients. Seasonal changes and varying lags are accounted for, and the result is six regime models which are tailored specifically to each route.

Three objectives are evaluated on out-of-sample data for each route:

i) An evaluation of whether similar parsimonious models outperform variable rich models. Six additional parsimonious regime models are therefore created, one for each route. The parsimonious models are found to provide better predictions based on performance metrics and the Diebold-Mariano test for forecasting accuracy. Top performing variables in the parsimonious models include, amongst others, secondhand prices, import and export factors, Chinese crude imports, vessel fixtures, and the ClarkSea index.

ii) An assessment of the forecasting capabilities of the regime model on never-before-seen data. These models yield promising results, and consistently rank in the top positions in regards to forecasting when compared to a set of benchmark models.

iii) An evaluation of the benefit of incorporating route-specific variables. The route-specific regime models are compared to a generic benchmark regime model with globally universal variables. Route-specific regime models are found to provide valuable outcomes and they improve the forecast in most cases, as opposed to the generic factor-driven models.

Keywords: Markov regime-switching, Regression, Freight rate, Oil tankers, Routes, Stability selection, Forecasting

Sammendrag

Oljetankermarkedet utgjør en betydelig del av det internasjonale shippingmarkedet, og anses som en kapitalintensiv og volatil industri. I denne avhandlingen er problemstillingen knyttet til å forutsi fraktraten for oljetankere på ulike shippingruter. For å predikere fraktraten, benytter vi en Markov-svitsjende regresjonmodell med to tilstander - én normal tilstand, og én volatil tilstand. Denne avhandlingen postulerer dermed at prediksjoner av kortsiktige fraktrater kan forbedres gjennom et rammeverk som har muligheten til å fange opp den distinkte underliggende naturen til fraktraten, ved å bytte mellom to adskilte regimer, i kombinasjon med betydningsfulle rutespesifikke og generelle determinanter. Vi gjør et inngående forsøk på å kombinere domenekunnskap om markedet, med statistiske metoder. Vår framgangsmåte for å gjøre dette er todelt:

For det første, så utfører vi en omfattende undersøkelse av shippingmarkedet, og mer spesifikt, oljetankermarkedet. Ved å gjøre dette, legger vi grunnlaget, og utvikler den nødvendige markedsforståelsen for modelleringen som skal utføres. Motivert av funnene i eksisterende litteratur av blant annet de drivende faktorene til fraktratene, observasjonene av strukturelle brudd, og mengden av eksisterende forsøk på å modellere fraktratene, så karakteriserer vi markedet. Vi tar et ekstensivt blikk på shippingmarkedet, undersøker hvordan tilbud og etterspørsel påvirker frakratene, og ser på den underliggende sykliske naturen til disse ratene. Videre, så analyseres de ulike shippingrutene grundig - seks av de globale rutene, TD1, TD3, TD7, TD12, TC1 og TC2 blir valgt ut i denne avhandlingen. En omfattende evaluering av de påvirkende faktorene til fraktraten blir så gjort. Disse faktorene grupperes etter om de regnes som tilbudsdrivende, etterspørselsdrivende, eller økonomiske og ikke-fundamentale. Videre skilles det mellom faktorer for hver rute. Disse er avgjørende for modelleringen.

For det andre, så utvikler vi en regresjonsmodell med regime-svitsjing for å forutsi fraktraten til oljetankere på ulike ruter, én måned fram i tid. Dataen testes for strukturelle brudd ved hjelp av en Chow-test, og indikasjoner på flere regimer blir funnet. En analyse for å velge ut gunstige kandidater til forklaringsvariabler utføres, basert på å gjøre delutvalg i kombinasjon med en seleksjonsalgoritme. Her er vi ute etter å validere de påvirkende faktorene til fraktraten, som tidligere ble plukket ut på teoretisk grunnlag. Ved å benytte metoden stabilitets-seleksjon sammen med en Lasso-straff med en tilfeldig regulariseringsparameter, så klarte vi å rangere faktorene basert på deres potensial i modelleringen. Med en delmengde av variabler for hver rute, så estimerte vi regresjonsmodellen for to tilstander. Sesongvariasjoner og tidsforskjøvne variabler ble tatt høyde for, og resultatet var seks regime-modeller, hvor hver modell var tilpasset spesifikt til én oljetanker-rute.

Prediksjonsmodellen ble evaluert på testdata for hver rute, basert på tre målsettinger. i) Først, en evaluering av om modeller med betydelig færre variabler kan forutsi observasjoner bedre enn modeller med mange variabler. Dette betydde at seks ytterligere regime-modeller måtte lages, én for hver rute. Det ble funnet at disse modellene gir bedre resultater etter vurdering av ulike effektmål. ii) Videre, en vurdering av modellenes evne til å forutsi ny data, satt opp mot ulike referansemodeller. Regime-modellene produserte gunstige utfall og ble konsekvent rangert i øvre sjikt. iii) Til slutt, en evaluering av fordelen ved å modellere spesifikt for ruter. De rute-spesifikke modellene ble sammenlignet med en tilsvarende modell, som kun benyttet generiske, globale forklaringsvariabler. Modellering for ruter ble funnet å gi verdifulle resultater og forbedret prediksjonen i de fleste tilfeller.

Contents

List of Figures

List of Tables

1 Introduction

The tanker shipping sector can be considered cyclical and volatile. It involves huge capital investments and exposes stakeholders to great amounts of risk. The tanker shipping market is one of the largest sectors in the world of shipping in terms of trading volume. The total tanker fleet consists of approximately 14,500 ships which transport several billion tonnes of oil each year among other things. They play a significant role in the transport of liquid commodities, allowing for high-volume, low-cost shipping in bulk. These tankers are generally categorized by size, e.g., VLCC, Suezmax, Aframax and Panamax, whereof some operate certain routes with a greater frequency than others. Core routes are those supporting the most essential flow between major markets. These routes are associated with everything from varying geography, geopolitics and trade flows, to different climate, currencies and distances.

Alizadeh and Talley (2011) state that fluctuations in the tanker freight rates are affected by the global economic activity, and the state of the tanker shipping market, in addition to other variables related to vessel and route characteristics.

In this study, we will look deeper into the routes concerning oil tankers. A selection of a six (6) specific routes will be made, and these will be examined extensively. The intricacies of these routes will explored meticulously from well-established theory, articles, related fields and expert domain knowledge. This will be the basis for finding determining factors for forecasting models of route-specific oil tanker freight rates. Furthermore, an approach to variable selection using regularization techniques to rank variables will be used. A Markov regime-switching regression model will then be pursued in the attempt of constructing a suitable generalised model, accounting for seasonality, lag, and route-specific and global factors among other things. We also wish to maintain a human element in the process - this specifically applies to the variable selection and modelling whereby a lot *could* be automated, but not without sacrificing domain knowledge. The model will be rigorously benchmarked and holdout-validated to ensure operationality.

The purpose of this thesis may be placed in the context of two research areas:

i) We look to determine the market dynamics and key determinants for the oil tanker freight market. This will be achieved by examining the freight rate mechanism and shipping cycles; studying the oil tanker market characteristics; assessing six key tanker routes which are different in terms of vessel size and type of oil trade (crude oil or oil products); revealing the fundamental factors on a macro-level which affects the freight rates across these *6* tanker routes; and performing rigorous variable selection to statistically rank the best fitting determining factors.

ii) We investigate whether it is beneficial to create route-specific, one-month ahead, prediction models based on two regimes. This will be conducted by incorporating the best factors from (i) in a Markov two-regime regression model, and benchmarking the results against generic models.

1.1 Chapter Breakdown

The remainder of this thesis is structured as follows:

Chapter 2. Literature review: As an initial measure we scrutinize the existing literature to get an overview of what has been done before, and how we can build on this work. We will consider different forecasting methods that have been applied to the shipping industry, as well as other markets. An investigation into the presence of regimes in the tanker market will also be made. Additionally, we take an extensive look at studies which examine the factors which impact the freight rate, with a focus on identifying route specific variables.

Chapter 3. World of Shipping: To be able to explain the market, and consequently develop a regime forecasting model, it is imperative to understand the market dynamics. We therefore embark on a comprehensive look at the freight market. We will use fundamental shipping theory as a basis to select the appropriate factors, as well as suitable model.

Chapter 4. Tanker Trade and Routes: In this chapter, we establish the core shipping routes we will examine throughout this thesis. We also look at the the dynamics of the supply and demand for oil which concerns countries and regions that matters to the routes.

Chapter 5. Factors Affecting Tanker Freight Rates: The key determining factors for the freight rates are reviewed in this chapter. The chapter is divided into three sections, which focus on supply, demand, and economic and non-fundamental factors. We describe the reason of interest for a factor, and the hypothetical significance to the rates.

Chapter 6. Data Analysis: In this chapter, we look at the data that we will be working with. The data is based on the previous chapter, whereby determining factors were considered. We look at where the data is gathered from, how the data is processed and descriptive statistics. We also motivate the regime model we are going to develop, by investigating stylized facts of the freight rates' behavior. Lastly, we study the underlying foundation of our independent variables, as this is useful in the interpretation of the results, as well as model development.

Chapter 7. Methodology: In the methodology chapter, we establish how we will develop our model. We set out to build a tanker freight rate forecasting model for individual routes, based on regime-switching and regression. In this chapter, we provide our methods for dealing with non-stationary variables, lags, seasonality, and assessment of the models. Our approach to variable selection is also presented here, which ranks factors based on the method of stability selection with randomized Lasso.

Chapter 8. Results: Our findings are presented based on the methodology in the previous chapter. Six route models (in addition to six parsimonious route models) will be considered. We will evaluate the determinants that where ranked during our variable selection procedure, how well the expected impact of the model's coefficient are, how seasonality is modelled, and assess the regimes. We further look at whether we are able to forecast with the proposed framework that theoretically should be able to explain the rates, and how good forecasts are. The regime models are compared with benchmark models on out-of-sample data, and further evaluated with performance measures.

Chapter 9. Discussion: In this chapter, we will discuss potential problems, remark on certain pitfalls, errors, and potential improvements.

Chapter 10. Conclusion: Lastly, we will provide a conclusive chapter to summarise our work.

2 Literature Review

A large number of empirical studies have been devoted to analysing the tanker shipping industry in terms of freight rates, chartering decisions and policies, transportation strategies, and fleet deployments and operations (Alizadeh and Talley, 2011). Our study is concerned with the former, the freight rates.

In the following, we will present our findings of empirical literature in the context of (i) freight rate literature on a general level, (ii) literature on factor driven models, and (iii) regime-switching and route-specific literature. Finally (iv), rationales for our methodology approach is outlined in the light of literature, along with our research contributions.

Note that the literature review that follows does not include detailed explanations of shipping terminology and related theory. We may therefore recommend readers who are unfamiliar with shipping to consider reading through Chapters 3 and 4 beforehand.

Literature on Modelling Shipping Freight Markets

Freight rate dynamics have been reviewed and modelled throughout decades with a vast variety of approaches and techniques. Nevertheless, few articles prove to generate promising out-of-sample forecasts. Broadly, shipping literature throughout history can be classified into two schools of thought; traditional structural models, and modern time series models.

Tinbergen (1934) and Koopmans (1939) are among the early efforts to describe the well-known cyclical nature of shipping through fundamental supply-demand theory, i.e. structural models. Koopmans (1939) argues that cyclicality is a result of the time-lag between ship capacity (supply) meeting demand, which hence triggers future expectations of the market. Koopmans (1939) was further among the first to propose theories for tanker forecasting. Today, the framework of classic static equilibrium (structural) models set out by Tinbergen (1934)[1] and Koopmans (1939) is still broadly valid, but today's literature is far more sophisticated in regards to both real-time and lagged effects of supply and demand (see, for instance, Zannetos (1964); Hawdon (1978); Norman and Wergeland (1981); Strandenes (1984); Beenstock and Vergottis (1993); Adland (2003); Randers and Göluke (2007)).

In the 1990s, advances within the field of econometrics developed rapidly, leading maritime economist to break new ground in the understanding of freight rate dynamics. Studies by Beenstock and Vergottis (1993) probably represents the inception of modern time series models in freight rate modelling. To date, studies have been conducted on several types of advanced Vector Autoregressive (VAR) models (Veenstra and Franses (1997); Tsioumas et al. (2017)), (Generalized) Autoregressive Conditional Heteroskedasticity ((G)ARCH) models (Jing et al. (2008); Abouarghoub and Mariscal (2011)), Markov Regime-Switching (MRS) models (Abouarghoub et al. (2014); Abouarghoub and Mariscal (2011)), Autoregressive (Integrated) Moving Average (AR(I)MA) models (Abouarghoub et al. (2017)), and Stochastic Volatility (SV) models (Benth et al. (2015); Benth and Koekebakker (2016)), to mention some.

The introduction of advanced time series models has enabled maritime economist to explore both time-varying structure and non-linear dynamics of freight rates, as well as co-integration between variables in factor models. For instance, a widely researched topic of freight rate dynamics is the stationarity property (see, for instance, Koekebakker et al. (2006); Adland and Cullinane (2006)). Koekebakker et al. (2006) are examining and testing whether freight rates in the dry bulk and tanker markets are stationary, and compare their findings to several other maritime research articles (see, further, Chapter 6).

Literature on Factor-Driven Tanker Models

Literature on the driving factors of freight rates can roughly be divided into two research fields; micro-economic models and macro-economic models. Micro-economic factors are concerned with factors that are specific in regards to a single ship or contractual terms, while macro-economic factors are factors that are determining for several vessels, and that are able to describe characteristics of either parts of, or the whole market (see, further, Chapter 5).

Major advances in empirical maritime research are evident when comparing literature a few decades ago to today's modern articles. For instance, Velonias (1995) bases most of his literature study on articles concerning traditional structural models, and performs a very simplistic regression model himself. While Fan et al. (2013) are, to our knowledge, the first to apply machine learning/artificial intelligence and Wavelet Neural Networks (WNN) to tanker forecasting. They attempted to forecast the BDTI index by incorporating factors such as the oil price (Brent), a volatility index (VIX), an oil trade index (Amex oil index) and stock indices (MSCI World Transportation, Dow Jones, S&P Global 1200), and managed to outperform an ARIMA model out-of-sample for longer-term forecasting.

[1] Fun fact: Tinbergen (1934) investigated, among other, the effect of the price of bunkers (fuel) on freight rates – which at the time was the price of *coal*! and not heavy fuel oil or marine gas oil as today

Zannetos (1964) empirically investigates tanker freight rates by regression analyses with numerous variables, both macroeconomic and microeconomic. Earlier efforts of factor analysis further belongs to the researches of Hawdon (1978), Velonias (1995), Strandenes (1984), and Beenstock and Vergottis (1989), who find factors affecting the interaction between supply and demand in the tanker market to include, for instance, world gross domestic production, country-specific industrial production, vessel tonnage capacity, delivery and demolition of vessels, oil prices, and commodity trade. However, and not surprisingly, considering the time these researches were conducted, there exists a vast amount of potential statistical violations and erroneous results in some of the models[2].

The use of macroeconomic variables to model and forecast tanker freight rates are also seen in modern studies by, e.g., Dikos et al. (2006) and Randers and Göluke (2007), who use system dynamic (SD) techniques to model freight rates. Randers and Göluke (2007) argue that the history of the world's shipping markets can be explained as the interaction of two cycle loops with different durations; a 20-year capacity adjustment loop and a 4-year capacity utilisation loop. They consider supply and demand to be endogenously and exogenously related to freight rates, respectively. They suggest that an endogenous supply model is sufficient in describing past freight history, while excluding exogenous noise. Macroeconomic factors such as newbuilding orders, average building time, vessels average life, demolition of vessels, and fleet utilisation are included in their study.

Other studies, by Kavussanos and Alizadeh (2002); Kavussanos (2003); Adland and Cullinane (2005); Adland and Cullinane (2006); and Lyridis et al. (2004), use univariate and multivariate time series models in an attempt of describing dynamics of freight rates, and further to forecast freight rate volatilities and levels. A variety of aggregate data and macroeconomic variables are utilised in the researches.

Later on, Alizadeh and Talley (2011) have been looking at the possible factors affecting the tanker freight rates, including the delivery time of chartered ships (the laycan period). They state that macroeconomic determinants have been investigated previously within the research field of both structural and time series models, and therefore focus their efforts on microeconomic factors (see, also, Tamvakis (1995); Tamvakis and Thanopoulou (2000)). Using contract data for the period of 2006 to 2009, the authors find evidence that the freight rate affects the laycan period and vice versa. Furthermore, they also find a significant link between the freight rate and hull size, fixture deadweight utilization ratio, vessel age and voyage routes. A similar connection is found between these factors and the laycan period, as well as the Baltic Dirty Tanker Index (BDTI) and its volatility.

A more recent study, by Lyridis et al. (2017) states that their framework of measuring risk with the FORESIM[3] simulation technique is the first attempt to express future tanker market risk in relation to current market fundamentals by combining Artificial Neural Networks (ANN) and stochastic models. They investigate both internal and external parameters affecting tanker risk. Fundamental factors included are, among others, oil price, fleet capacity, orderbook, demolition price, and OPEC production. Their research yielded promising out-of-sample results on market risk. FORESIM was also conducted in the research by Zacharioudakis and Lyridis (2011), who attempt to express future tanker freight levels (on the Ras Tanura – Rotterdam route) in numerous states of OPEC oil production levels. Results showed that ANNs were adequately capable of simulating future freight rates.

Literature on Regime-Switching

In the classic maritime economic literature, the short-term supply curve is characterised by two distinct regimes; one elastic part and one inelastic part for lower and higher freight rates, respectively. Hence, volatility in freight rates are conditional on the freight rate level. This characteristic shape of the shipping supply curve was first proposed by Koopmans (1939), and is today widely acknowledged by market practitioners, as well as empirical researchers (see, further, Chapter 3 and Chapter 6, and for instance, Zannetos (1964); Norman and Wergeland (1981); Stopford (2009); Alizadeh and Talley (2011); Strandenes (2002)).

We have come across almost a dozen articles incorporating regimes and non-linear characteristics of tanker freight rates in modelling. However, most of them are concerned with various risk models in terms of freight return volatility, and do not attempt to forecast freight rate levels. Nevertheless, many of the articles are providing empirical evidence for the existence of distinctive regimes in the freight rates. Important to note, and as literature suggest, *regime-switching models provide a better understanding of volatility clusters in the lower and higher volatility regimes for the distinctive nature of the freight market, regards to, e.g., magnitude and duration.*

Abouarghoub et al. (2014) are modelling a two-state MRS distinctive volatility model for tanker freight returns (BDTI index) in a GARCH framework applied to Value at Risk (VaR) measures. They address some

[2] Violations of ordinary least square regression (OLS) assumptions. See, further, Chapter 7.

[3] FORESIM: an innovative simulation technique combining stochastic models and artificial neural networks

important dynamics of the freight rate, and the results are profound. In line with maritime theory, they find that the low volatility regime accounts for a higher percentage share of total observations and has a longer duration on average. In other words, long memory is more pronounced in lower freight rate states than in higher freight rate states. Moreover, authors describe that the distinct regimes are characterised with a higher tendency to shift from the higher volatility structure to the lower volatility structure, as opposed to the tendency of shifting the opposite direction. They themselves motivate a two-state conditional regime framework by referring to the studies of Kavussanos and Alizadeh (2002); Alizadeh and Nomikos (2011); Abouarghoub and Mariscal (2011); and Abouarghoub et al. (2012). Abouarghoub et al. (2014) finalise their study by suggesting that: «further research should be conducted to examine the extent of the impact of vessel size and shipping routes on conditional volatility during different phases of the supply curve», which matches parts of the motivation and aims of this study.

IAME (2014) presents a study on MRS for 9 dirty tanker routes which are part of the BDTI index. The BDTI is used in the models as an explanatory variable, serving as a market condition benchmark. The study finds evidence of a positive correlation between the size of tanker vessels and their four statistical moments[4].

Abouarghoub et al. (2012) model average TCE freight rates across 5 tanker classes, namely VLCC, Suezmax, Aframax, dirty MR, and all tankers. They take on a MRS approach, and are testing for the occurrence of structural breaks between 1990 and 2011 by using a Chow test (see Chapter 7). 3 structural breaks result from their analysis. Even though maritime economists agree that shipping cycles are endogenously driven in the short-term, a general consensus is also that exogenous factors channel structural breaks in the long-run.

Alizadeh and Talley (2011) study the dynamics of the term structure and volatility of freight rates for 3 sizes of bulk carriers and 3 sizes of tankers. They find asymmetry to be apparent in the freight rates when the market is in backwardation versus contango, indicating that the rate of increase in volatility increases (decreases) as the degree of the former (latter) state increases.

Abouarghoub and Mariscal (2011) investigate 5 routes that are part of the BDTI index, namely Worldscale[5] rates of TD3, TD4, TD5, TD7, and TD9 (see, further, Chapters 3 and 4), in a two-regime MRS GARCH framework in order to model VaR. They find evidence of tanker freight rates exhibiting such properties, and market shocks to increase and have a lasting effect on volatility. According to the authors, this is the first attempt to investigate conditional freight volatility regimes in the tanker market. Additionally, they find volatility in the larger tanker segments to be more sensitive to market shocks in comparison to smaller segments, which provides support to our motivation of looking at different vessel sizes across different routes.

Kavussanos and Alizadeh (2002) apply MRS to both seasonality testing and forecasting. They reject the existence of stochastic seasonality. Deterministic seasonality is evident though, and is indicating decreases in rates from January to April and increases in rates in November through December - although some variation in seasonality is observed depending on vessel size and market condition (see, further, Chapter 8).

Overall, Mr. Abouarghoub is probably the single largest contributor to research on regime models in the tanker space. Interested readers may therefore find Abouarghoub's collective work paper (partial fulfilment of PHD) as a good source to information, both in regards to his contribution on regime models in shipping, but also in general empirical terms (Abouarghoub, 2013).

Besides shipping literature, studies on non-linearity and regime-switching can also be found in other research fields. Dafas (2004) attempts to model the crude oil spot price with a mean-reverting jump-diffusion model, along with a Markov-switching approach. Escribano (2002) implements a model for the electricity price which accounts for seasonality, mean-reversion, conditional heteroskedasticity and jumps. Weron et al. (2004) propose various models to electricity spot prices. They analyse and model the logarithm of the deseasonalised average daily spot prices, with a focus on regime-switching. Kosater and Mosler (2006) compare regime-switching models against an AR(1)-process in the electricity market, and find the prior to be slightly better.

Methodology Rationale and Thesis Contribution

From the aforementioned literature, we find research on freight rates in the tanker market to be quite extensive in general terms, yet somewhat limited in the area of our research scope. How does our methodology approach combined with our specification of routes and associated factors contribute to the field of freight rate prediction? It is done by providing results that are fairly easily interpretable, are readily available, and relevant. They are usable by market participants with limited empirical modelling experience, while the results at the same time ensures to capture certain complexity of the freight rate dynamics. Furthermore, com-

[4]The four statistical moments refer to: mean, variance, skewness, and kurtosis

[5]In Section 3.1, we briefly outline why we do not think Worldscale rates are appropriate to model as opposed to TCE rates.

pared to earlier empirical literature, we are testing a selection of determinants (*approx. 170*) that is substantially more comprehensive than ever done before, to our knowledge. Hence, there is some lack of prior empirical evidence on the tanker market that could support our choices for some of the determinants. Nevertheless, tanker market commentary, «expert-knowledge» and literature on other markets, does provide support. And, in the selection of the *statistically* most appropriate factors, we will be utilising a sophisticated variable selection technique based on stability selection with randomized LASSO (see, further, Chapter 7).

In order to select appropriate factors, we need, in combination with empirical literature, a thorough understanding of market fundamentals. In Chapters 3 – 4 we delve into shipping theory to understand the most relevant conditions and factors that are prevalent to the oil tanker market. We use this knowledge to identify and examine the most important behavioural properties of the tanker freight rates. Maritime Economics by Stopford (2009) is the basis for fundamental shipping theory[6]. Our selection of hypothetical driving factors is introduced in Chapter 5, and supportive empirical literature is referenced to selected factors where available.

Our motivation of modelling freight rates across different vessel sizes and routes in a Markov regime-switching framework is supported by Abouarghoub et al. (2014) and others, as aforementioned references describe. Furthermore, Abouarghoub et al. (2014) point out that incorporating seasonality effects would be an interesting and recommended extension of their research – which we aim to incorporate in our models. As described above, literature on shipping seasonality suggest that tanker freight rates exhibit deterministic seasonality (see, for instance, Kavussanos and Alizadeh (2002)). We do not test the existence of deterministic seasonality explicitly ourselves, but we rather build on findings of existing literature in our methodology (see, further, Chapter 7).

Market-wise, we further motivate why a two-regime model could be appropriate to model shipping rates throughout the theory chapters in advance of outlining the specific methodology that is conducted in this thesis, as it is useful to have knowledge of some fundamental market theory before commencing into regime switching. Furthermore, we set out to prove or disprove some theoretical characteristics of freight rates regards to having regime distinct properties. We describe these characteristic along with supporting literature in more detail in Chapter 6.

This thesis hence postulates that predictions of short-term freight rates can be improved through a framework that can capture the distinctive nature of freight rates by switching between two distinctive regimes, while combining this with hypothetical superior route-specific and general-specific determinants.

[6]Maritime Economics by Stopford (2009) is broadly used across universities lecturing in shipping economics.

3 World of Shipping

Maritime transport of goods and raw materials has been essential to world economic trade and growth for millenniums. Shipping enables development by bringing together capital goods, intermediates and consumers via extensive cross-border transport networks, further promoting global (e.g. WTO) and regional (e.g. EFTA, NAFTA, AFTA) economic integration. Import and export on the scale necessary for the modern world would not be possible without the international shipping industry. The world of shipping accounts for around 85% of world trade, corresponding to 1.49 tonnes per world capita, and is thus the largest mode of trade in the world. Total seaborne trade has almost doubled since the start of the 21st century. In 2016, total seaborne trade amounted to 11.1 billion tonnes of cargo or 55.1 trillion tonne-miles[7] (Clarksons Research Services Limited, 2017). Increasing industrialisation, economic liberalisation as well as advances in technology will support continued growth in shipping trade. In the forthcoming chapter, we will highlight key fundamentals of shipping, which serves as both useful background information and a base for understanding and selection of variables and econometric models.

Shipping Segments

The merchant fleet comprises above 50,000 vessels (>1k dwt). These transport all types of cargo and are put into categories accordingly. The three major ones being *dry bulk carriers*, *oil tankers* and *container ships*. The two former segments are traditionally concerned with Tramp Shipping (on-demand services), while the latter segment is often referred to as Liner Shipping (predefined routes and schedules services). Within the tanker segment, we also find gas (LNG and LPG), specialised and chemical vessels. Other shipping segments include, e.g., offshore, car and passenger vessels. Vessels in smaller segments are normally built for a more specific purpose or voyage than for a general market (e.g. ferries, cruise ships, well boats).

Dry bulk commodities include iron ore, coal, grain and other minor bulk [8], and are normally transported in large ship hulls in bulk form. The dry bulk segment transports the highest volume of cargo each year. Oil tankers are ships designed for liquid bulk transport of crude oil and oil products. Container ships typically carry intermodal container units which are possible to carry on different modes of transport. Containers could hold everything from dry storage items to specialised refrigerated cargo. Table 3.1 provides a brief overview of these segments' standing in terms of seaborne trade and fleet size.

3.1 Four Shipping Markets

In shipping, there are four overarching markets (Stopford, 2009):

Newbuilding	Sales and Purchase
Demolition	Freight market

Cash flows between these markets are driving the overall shipping market. The freight market is the main driver in shipowners' positive cash flow. This does however vary based on the current market situation. In certain periods, income from for example demolition could constitute the majority of positive cash flow. Moreover, shipowners usually have a margin policy for acceptable freight rate below or above lay-up cost (break-even). This could, however, vary based on a given market situation, so that in periods income from, e.g., demolition, could constitute the majority of positive cash flow. Shipowners do usually have a margin policy, dependent on market situation, for acceptable freight rate below or above lay-up cost. In the newbuilding market, vessels are built at shipyards on the order of shipowners. The S&P (or secondhand) market is where existing vessels are sold and bought. A vessel's life ends in the demolition market, where vessels are scrapped. The shipowner is then compensated based on the value of the vessel's steel, as well as other materials and components.

Stakeholders

The key stakeholders within the four shipping markets are:

Shipowners	Charterers
Shipbrokers	Shipyards
Capital markets	

Shipowners are the companies owning and operating the ships, and their fleet may constitute of ships operating across different shipping segments.

The charterers are the companies that hire the ships. These companies have specific requirements for ship specifications (e.g., speed, fuel consumption, deck and volume capacity), contract length and off hire terms. As noted by Alizadeh and Talley (2011), the behavior of shipping freight rates and the timing of shipping contracts affect the transportation costs of charterers and the operating cash flows of shipowners. The shipbroker's task is to find suitable ships for a given job, with regards to ship specification, contract length and

[7] Tonne-miles is a shipping measure of true seaborne demand; tonnes of cargo carried multiplied by the distance it travels. The tonne-miles measure also indicates the amount of transportation the fleet is capable and/or willingly to supply. Equilibrium freight rates are thus found in the intersection of tonne-miles supply and demand (see Section 3.2).

[8] Steel products, forest and agricultural products, fertilisers, cement etc.

Table 3.1: Shipping segments. Overview of seaborne trade and fleet size of the major shipping segments. Fleet size includes vessels above 10k.dwt (Clarksons Research Services Limited, 2017)

Shipping Segment	Seaborne Trade (10 year CAGR %) m.tonnes	b.tonne-miles	Fleet Size Number of vessels (#)	Tonnage capacity (dwt)	Percentage of total fleet (%)
Oil Tanker	3,017 (1.25%)	12,620 (3.27%)	4,980	546,943,878	30%
Dry Bulk	4,895 (4.15%)	27,203 (3.85%)	11,111	817,018,864	45%
Container	1,727 (4.69%)	8,580 (4.37%)	2,445	247,639,353	14%
All merchant segments	11,130 (3.17%)	55,112 (3.26%)	26,381	1,817,057,440	100%

price. Shipbrokers act as a link between the shipowners and operating companies, and participate in the negotiations of the contractual terms. Shipbrokers are hired both by shipowners and charterers. Additionally, a shipbroker may participate in the negotiations of a ship's laytime, accounting services and in the case of disagreements.

The shipyards perform the construction of the ships. Additionally, they offer repairs, modifications and maintenance. The stipulation of the newbuilding prices is an important part of the business of the shipyards.

In the capital markets, capital is raised and restructured, as well as liquidated in various market places. Stakeholders in these markets are many, such as corporate and commercial banks, traders and investors.

The Freight Market

Contract Types for Seaborne Freight

In pursuance of understanding the dynamics of specific tanker freight rates, it is essential to know how freight rates are determined for various segments, vessel sizes, cargo types and trading routes. There are four commonly used contract types (charter-parties) for freight agreements between shipowners and charterers (Stopford, 2009). A brief description of each of them is given here, and summarized in Table 3.2. Contract types are usually set apart based on what expenses and responsibilities that are covered by the charterer. For a single voyage, costs are associated with operational (opex) and voyage-specific (voyex) expenditures. Opex include mostly administration, lubricating oil, maintenance and repairs, docking, crew and insurance costs, while voyex mainly consist of fuel, port and canal costs that are directly related to a specific voyage.

On a *voyage charter*, vessels are hired to carry a specific cargo from a load port(s) to a discharge port(s) at an agreed freight rate. The owner bears all costs in a voyage charter contract. Payment is normally per tonne of cargo (Clarksons, 2017b). Voyage charter rates are not necessarily comparable across different trading routes, since voyex costs may vary from route to route. A single voyage charter is often referred to as a spot fixture or spot voyage charter. If a voyage charter is extended to include a series of shipments for a fixed period of time, it is called a *contract of affreightment (COA)*. The charterer then pays the shipowner a fixed amount per tonne of cargo transported. The shipowner is also free to choose the vessel(s) under this contract type. For example, a COA could be an agreement of the delivery of one shipment each month for a period of one year. A *bareboat charter* type of contract gives the charterer full operational control of a vessel for a specified period of time. The charterer is therefore responsible for both opex and voyex. The freight rate is paid on a per day basis. Financial investors are common shipowners that charter their vessels out on such agreements, since the shipowner does not need any operational shipping expertise. On a *time charter*, vessels are chartered for a fixed period at a set hire rate. Shipowners pay for opex, while charterers cover bunkers (fuel) cost and other voyex costs. In return, shipowners are paid an agreed freight fee per day (or month or year), less potential off-hire time. The charterer is instructing the crew, paid by the shipowner, about where, when and what to load and deliver. Time charters could be either spot- or term charter parties. The difference is the duration of the contracts; contracts with duration less than three months are usually defined as spot contracts. Term contracts could vary, but are often sorted into two categories, medium- and long-term contracts, which have duration one month to a year and longer than a year, respectively (Clarksons, 2017b). If time allows, other charters can be done in between a voyage-, bareboat- or time charter.

Spot contracts are usually written only a few days before start of operations, and are re-negotiated frequently. Today's spot rate is therefore not necessarily equal to tomorrow's spot rate. Shipbrokers are important actors in negotiating these spot rates on behalf of both shipowners and operators. Ideally, spot rates are determined by the current supply and demand equilibrium for certain shipping services, while term/COA rates in addition are determined by shipowners' and charterers' long-term expectations about the future (Stopford, 2009). Based on term rates' dependency of spot rates and the liquidity that spot rates offer, spot rates are more volatile than term rates (Kavussanos, 2003). Volatility may also change dependent on what

Table 3.2: Charter contract types. A brief overview of the characteristics of common charter parties.

Charter party	Party paying for Opex	Party paying for Voyex	Freight rate
Voyage charter, COA	Shipowner	Shipowner	$ per tonne (WS flat rate)
Time charter	Shipowner	Charterer	$ per day
Bareboat charter	Charterer	Charterer	$ per day

state the market is in. Forward Freight Agreements (FFA) are derivatives instruments used to hedge freight rates against future market conditions, based on specific routes or freight indices. FFA contracts are traded over-the-counter. These contracts are common in both dry bulk and tanker markets.

Worldscale - a Tanker Industry Measure

Oil tankers are commonly fixed on spot voyage charters, where charterers could be oil and gas companies, oil traders and oil refineries (for storage purposes). Voyage charter freight rates in the tanker industry are generally settled on the basis of the Worldscale index. The index is based on the freight rate of transporting a tonne of cargo using a standard vessel on a *round voyage*. This rate ($ per tonne) is designated as the *Worldscale 100 (WS 100)* or *flat rate*. The idea is that the flat rate should reflect a voyage break-even level for the standard vessel[9], regardless of the voyage for which the ship is chartered. Characteristics (e.g., voyex and distance) for different routes vary[10], and thus it exists a unique WS 100 for an array of oil routes. Today, a schedule of about 300,000 flat rates are available to market participants (Worldscale Association, 2018). These flat rates are revised and updated annually by a Worldscale panel[11], and the standard vessel used in calculations is updated from time to time.

The purpose of the flat rates is to enable market participants to easily compare earnings on alternative voyages. Economies of scale is evident when comparing worldscale rates for different vessel sizes on the same trading route (i.e. a VLCC will typically trade at a lower worldscale than an Aframax). Large vessels therefore need a lower worldscale rate than smaller vessels to make a profit (Pagonis, 2009). For instance, if an agreement for transporting crude oil from the Middle Eastern Gulf (MEG) to the U.S. is made at WS 70, and WS 100 corresponds to $20.00 per tonne, then the shipowner receives a freight rate of $14.00 per tonne. Whether this portrays a profit depends on the vessel used.

Time Charter Equivalent

Theory above suggests that differences in the charter-parties make it difficult to compare freight rates directly. Freight rates are therefore usually quoted on a *Time Charter Equivalent (TCE)* basis to facilitate comparisons across periods (e.g. due to yearly revisions of Worldscale), trading routes and shipowner's performance despite changes in the mix of charter types. For instance, given a vessel on a $15,000 per day time charter with $10,000 per day in voyex costs, a similar vessel on a voyage charter will have to earn $25,000 per day to earn a TCE rate of $15,000/day; voyex would otherwise be paid by the charterer under a time charter contract. TCE rates are thus a measure of the gross freight income less voyex, divided by the duration of the voyage (Clarksons, 2017b).

The Baltic Exchange is the world's leading source of benchmark indices in shipping. Every weekday they publish updates on indices that are meant to track the development of different shipping segments, which are based on the professional assessment of independent shipbrokers from all over the world (The Baltic Exchange, 2018). The shipbrokers use the latest fixing prices in their assessments. In tanker shipping, we have the Baltic Dirty Tanker Index (BDTI) and the Baltic Clean Tanker Index (BCTI) that track the development of dirty and clean oil trade, respectively[12]. The indices represent a weighted average of TCE freight rates for a set of major crude oil and oil products routes (see Appendix A.6).

3.2 Freight Rate Mechanism

The freight rate mechanism is the mechanism which links cycle theory with supply-demand theory. When

[9] At present, the standard vessel is a 75,000 dwt tanker. In addition to deducting voyex costs, a fixed hire cost of $12,000 per day is included for the calculation of the WS 100; so that in reality WS 100 refers to the standard vessel earning $12,000 per day on a time charter basis - even though the Worldscale Association state that it is not the intention to produce rates providing a certain level of income or margin of profit.

[10] Worldscale Association (2018): "From 2016 additional costs in complying with emission regulations are provided within the Worldscale rates. These allowances have been calculated using a breakdown of voyage distances within and outside sulphur emission control areas (ECA)".

[11] We note that it is important to be aware that WS 100 is not directly comparable year over year, which addresses the importance of not using WS data in modelling; motivates why TCE rates should be used instead

[12] Dirty refers to oil trade of crude and heavy (black) oils, while clean refers to product tankers carrying oil products such as gasoline, diesel fuel and jet fuel (see section 3.5)

the market is tightly balanced, even small changes could cause substantial impacts. It is the combination of a volatile demand and a significant time-lag for the supply to adjust itself that creates the framework for the shipping cycle.

Charterers and shipowners negotiate to establish a freight rate that reflects the balance of available ships and services/cargo in the market (Stopford, 2009). The balance will continuously adjust the freight rate. A scenario where supply is low will push freight rates higher, with the consequence of more suppliers being willing to enter the market. Suppliers must therefore either buy new ships in the newbuilding market, or preferably existing ships in the S&P market, to take advantage of the presently high freight rates. Thus, the second-hand prices are bid up to a level where newbuildings seem to pay off. The subsequent delivery of newbuildings will increase the fleet size, and possibly cause an overcapacity of ships. The pressure on freight rates increases, and ships must be laid-up, sold or scrapped to restore market balance.

Supply and Demand Curves

Supply and demand curves in shipping markets do typically have the shape seen in Figure 3.1 (Alizadeh and Nomikos, 2011). The intersection of the supply and demand curves is the theoretical freight rate equilibrium. Furthermore, market participant's perception regarding the current market situation does also affect freight rate negotiations.

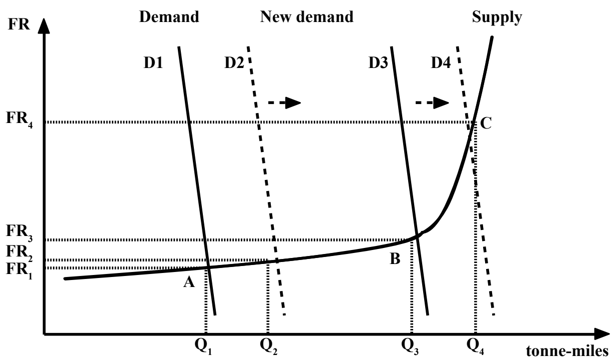


Figure 3.1: Supply-Demand curves in the freight market. Equilibrium freight rate levels can be seen for different shifts in demand (A, B, C). $Q_{\#}$ on the x-axis represents the tonnage capacity, in tonne-miles, that shipowners are willing to supply at a given freight rate level FR (y-axis).

The supply curve indicates the amount of transportation (tonne-miles) that the fleet is willing to supply at a given freight rate (Alizadeh and Nomikos, 2011). The curve is J-shaped, and is composed of each ship's individual supply curve. New and more modern ships will contribute to the lower left part of the supply curve, since these are able to operate more cost-efficiently compared to older ships. The curve indicates that the supply is price elastic up to a certain point, where the fleet sails at close to maximum capacity. Here, it cannot react to short-term increase in demand, and thus becomes inelastic. In the elastic part, supply exceeds demand, and only the most cost-efficient vessels are hired - leaving freight rates lower than in the inelastic part. In a situation where the freight rate falls below the acceptable freight rate (opex cost) for a given ship, shipowners have to decide whether the ship should be operated with a loss, laid-up, scrapped or sold.

The demand curve describes charterers' required amount of shipping supply at a given freight rate. An operating company, or oil company, has much higher earnings and costs from normal operations than the cost of shipping, hence being more affected by underlying demand for goods and products. Furthermore, operators are dependent on continuous delivery of supplies, due to high alternative costs concerned with, e.g., storage. Consequently, the demand curve is inelastic over its entire area. In the short-run, the demand has the ability to shift quite extensively in comparison to the supply.

3.3 Shipping Cycles

Shipping markets are known to be cyclical, not structural in nature (Tønne, 2016). A cycle can be summarized as a long-winded, almost constant imbalance in the market. This imbalance is mostly a consequence of the various stakeholders interacting - it is both created and attempted stabilised by the various stakeholders. Consequently, it is important to understand how this market dynamic behaves to identify the most applicable model. Cycle and supply-demand theory is central in this understanding.

Cycle Characteristics

The history reveals that significant upturns imply that a large number of shipowners are betting on the same upturning wave, which can trigger long periods of recession. When the market eventually recovers and seems profitable, shipowners again tend to overorder newbuildings, so that the cyclical movement continues. Traditionally, shipowners and shipping investors with this kind of strategy have been labelled with a certain degree of short-mindedness.

Stopford (2009) distinguishes between three cycle periods when describing cycle characteristics:

Long-term   Short-term   Seasonal

Shipping cycles are periodic, and not symmetric. That is to say, cycles can last for various lengths of time, and be very different from previous cycles. Cycles will occur as long as there exists fluctuations in the balance between supply and demand. The short-term cycle especially, acts as a function for coordinating this balance.

Long-term Cycle

Long-term cycles are periods which span from several decades to whole centuries. These periods are characterized by overarching technological, economic and regional developments. Technological innovations are often the easiest to identify. These include great innovative breakthroughs throughout the world history such as the steam engine, railway and motor power. Still, it will be challenging to pinpoint the exact moment in time where these have affected the freight rates.

Short-term Cycle

Short-term cycles range from periods from three to twelve years. These cycles are usually the most prominent, being more distinct and relevant for analysis. This comes from a limited access to long-term historical data, as well as uncertainty for the nearby and distant future. Short-term cycles have the following four stages (See Figure 3.2):

Trough Recovery Peak Collapse

Trough is where the freight rate has reached a minimum level or marks the end of a downturn. In this stage, the market experiences a saturation of ships and/or low demand. Freight rates reach a level below operational costs, and ships are consequently laid up. Low freight rates and pressured credit loans leads to negative cash flows, and banks defer granting new loans. Shipowners face difficult decisions, and ships may be sold for far less than book value. Second-hand prices may fall even down to the level of scrap prices.

The *recovery* stage occurs when the market balance is adjusting towards a supply-demand equilibrium. Freight rates increase above operational costs, and ships are taken out of lay-up. The second-hand prices increase and the future prospects brighten. The *peak* is then reached when the market experiences surplus demand. Freight rates increase far beyond operational costs, so that every ship is taken out of lay-up and operates at high transit speeds. Optimism spreads as the peak stage lasts on. Modern ships are sold at second-hand prices reaching above newbuilding prices, orderbooks increase and banks are willing to grant loans against strong asset collateral. At the point of surplus supply, when newbuildings reach the market and/or demand drops, freight rates will exhibit a *collapse*. The liquidity remains high during the collapse phase, and shipowners hesitate to sell or scrap ships.

Supply and demand forces will eventually force the market back to the trough stage, and the cyclical nature continues.

Short-term shipping cycles do, however, not always follow these four stages precisely. The recovery stage may end up in a second collapse or a long-lasting flat growth, called abortive recovery, inducing a prolonged trough stage. Large order volumes of cheap newbuildings as a consequence of investors predicting the market cycle is an example of counter cyclical activity. This means that supply which is meant to decrease, increases. Particularly heavily ordering of newbuildings in peak periods may cause the following trough and recovery stages to endure far longer, hence delaying a cyclical upswing. Newbuildings are usually delivered one to four years after orders have been made, causing increased pressure on the market when they are finally delivered.

A moving average is suitable to represent the short-term cycle (Stopford, 2009). The seasonal element can be leveled out with such a method. The average is centered around a given month, with an equal number of months on either sides. The average is subsequently shifted from month to month. The purpose is to capture the freight rate shape that best represents a short-term cycle (Stopford, 2009). Figure 3.2 shows the short-term cycle for two Baltic Exchange indices obtained from this method, adjusted for seasonal variations.

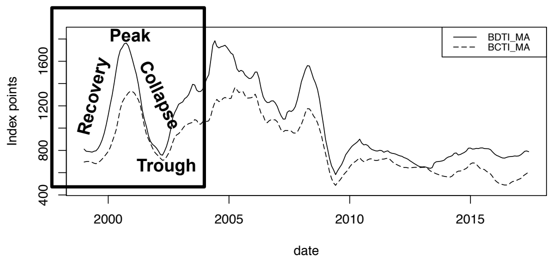


Figure 3.2: Short-term cycle. The four stages of a short-term cycle. Seasonal trends are leveled out with equally weighted moving averages. Baltic Exchange Dirty, and Clean indices are displayed.

Seasonal Cycle

Seasonal cycles are seasonal changes in the freight rate within a year, mainly determined by the seasonal demand. Regulation of supply is generally subject to a more long-lasting shift. In Figure 3.3, the seasonal cycle for the BDTI index is represented. Further, in Figure 3.4, a smoothed BDTI short-term cycle is represented along with the original BDTI index to illustrate the seasonal effects. Seasonality will be more thoroughly investigated in later chapters (see, further, Chapters 6 - 8).

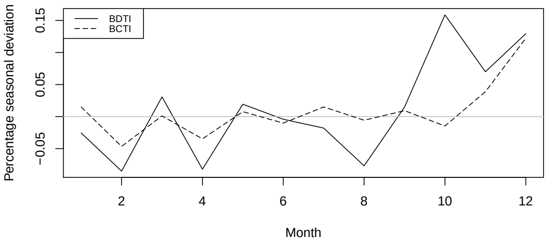


Figure 3.3: Seasonal cycle. Seasonal components for the BDTI index. Calculated as the mean seasonal change within the period August 2002 - June 2017.

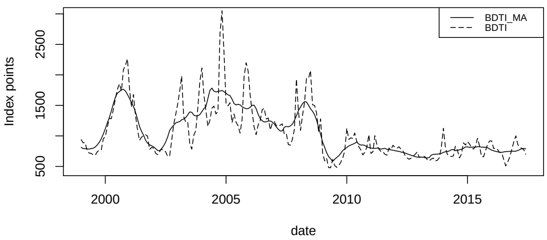


Figure 3.4: BDTI index. A smoothed short-term cycle overlaid the original BDTI index.

3.4 Risk and Cycles

Shipping cycles are defining the risk in the market. Shipping risk is defined as the «measurable liability for any financial loss arising from unforeseen imbalances between the supply and demand for sea transport» (Stopford, 2009). In shipping, the risk is divided between spot and term contracts. The spot market is where the shipping companies are at the greatest risk, but at the same time exposed to the greatest potential upside. In addition to potential upside with high freight rates, the spot market is more like a showcase than the long-term market (Fearnley, 2018). Ships that have traded in the spot market may as such be more likely to secure future contracts, as these are of higher awareness to several charterers. Charterers, on the other hand, would be at the greatest risk if they own their own fleet, as they then will be directly affected by their own activity. An «intermediate risk» for both parties is to enter into long-term contracts and/or speculative FFA trading.

Although a long-term market position is considered to be least risky for the charterers, they will in times of low capacity supply be at risk of paying a premium on the spot freight rate. The downside of lost revenues or increased operating costs from a lack of transportation options, is simply far too big compared to that extra potential freight cost. The shipping companies, however, will also take a certain risk of entering into long-term contracts, as they sometimes may be paid less at times when the spot rate is higher than their long-term rates. Time charter contracts for a certain period should reflect a weighted average of spot rates for the same particular period. The market is rarely quite efficient, and as a consequence, term and spot rates differ from each other during an observed period.

Revenue and Cost Risk

Risk in regards to earnings is depending both on income and cost. On the income side, owners are exposed to risk in the freight rate, utilisation (employment days and cargo capacity) and lifetime of the vessel regards to second-hand or demolition value. On the cost side, risks lie in capex costs (e.g., interest and amortisation subject to fluctuations in devaluation risk of currencies), opex costs (e.g., administration and operation, maintenance and docking, crew and insurance) and voyex costs (e.g., fuel, port and canal costs). Altogether, as the numerous factors that bear risks for the shipowners' earnings indicate, risk management measures are important to take. For instance, fuel costs accounts for about 75% of the total voyex costs, and shipowners actively manage this short-term risk by adjusting the ship speed in accordance to the activity level in the market. In the long-term, future expectations are vital in the decision-making process of when to order, sell or scrap a vessel. Forecasting techniques may be very useful to understand the underlying factors of the risks, and to make an opinion about the future.

Expectations Creates Challenges

Analysts have seen that wealthy shipping companies have a better ability of tackling a cyclical market, as their financial buffer against losses is higher (Fearnley, 2018). An alternating offer strategy that follows the cycle may turn out to be catastrophic for companies if the trough phase is lasting over a longer period (Stopford, 2009). In addition to normal supply and demand behaviour, cyclical uncertainty is also impacted by shipbrokers. Hampton (1991) has addressed the importance of human impact in both the long-term and short-term cycle: «In today‘s modern shipping market it is easy to forget that a drama of human emotions is played out in market movements». Shipbrokers have a desire to interpret the signals generated by the freight rate and, as an intermediary, will directly affect the freight rate cycle. Hampton points out the weakness of shipbrokers' rational for giving particular price signals as a reason for the «overresponsiveness» of the market.

However, it is not just about being right, but being right when others are not right. Hampton (1991) argues that some of the best opportunities arise this way, and that consensus is generally not a good indication. If both a shipping company and a chartering company have made proper analyses, and both have identified a peak in the market, naturally, the chartering company would prefer not to lease on long-term contracts, although this is desired by the shipping company. Consequently, it is important to evaluate at which stage of the cycle period you are at and at the same time assess the underlying psychology of the market, and not to be too concerned about the time horizon. If shipping companies find that they do not want to scrap ships as a result of expectations of an early upturn, the cycle will last longer. Shipowners, and other actors like brokers, who are guessing at which stage the market cycle currently is, as Hampton points out, makes it challenging to assess the duration of the cycle.

3.5 Oil Tankers

Seaborne trade of oil can be divided into two main groups; *crude oil trade* and *products oil trade*. In 2016, oil tankers[13] carrying these two cargo types accounted for a total of 22.9% of world seaborne trade (Clarksons Research Services Limited, 2017). Crude oil accounts for 33% of world's primary energy consumption, illustrating the scale of this energy source (BP, 2017).

In this section, we will at first give a more comprehensive description of the cargo itself, and put its corresponding trade in context to other trade. Next, we introduce the various tanker vessels and the current tanker fleet. Finally, we highlight some characteristics of the tanker market. We consider this knowledge to be important in order to understand the underlying drivers of the trade of oil.

Cargo

Crude oil and oil products are liquid bulk cargo.

Crude oil is a commodity. However, crude oil is not a «universal» commodity, many different gradings of crude oil are produced around the world. Crude oils have different quality characteristics, the two most important being sulphur content and density level. The sulphur content, expressed as percentage by mass, determines whether the oil is sweet or sour. Sweet oil has low sulphur content. The density level is measured in American Petroleum Institute gravity (API), and is a function of the oils specific gravity. API determines whether we are dealing with a light or heavy oil. Crude oil's quality grade is an important factor to consider for oil refineries, who are processing the crude oil. Refineries adjust their refinery facilities in order to process and make the most out of different crude oils. Processing of light, sweet grades of oil is far less sophisticated and energy-intensive to heavy, sour grades (EIA, 2018). Switching between gradings may be cost-intensive or infeasible. Hence, refinery capacity must be seen in the light of capacity by oil grading.

Oil products are the refined products (distillates) produced from a refinery process. These products are usually grouped into three categories; light distillates[14] (e.g., butane, LPG, gasoline, naphtha), middle distillates (e.g., kerosene, jet fuel, diesel fuel) and heavy distillates (e.g., heavy fuel oil, residues). Most refineries focus on products used for transportation. In fact, of world's oil consumption, about 45% is consumed by light and heavy duty transportation (ExxonMobil, 2017).

According to the U.S. Energy Information Administration (EIA) agency, pricing of crude oils is usually in the favour of the light and sweet oil gradings. Oil products such as gasoline and diesel fuel typically sell at a premium compared to more heavier products, and these products are more easily refined from light, sweet crude oil. The North Sea Brent and the U.S West Texas Intermediate (WTI) are examples of light and sweet oils (EIA, 2018).

To put crude oil and oil products trade into context, we have listed a breakdown of world seaborne trade of goods and raw materials by tonnes and tonne-miles, as well as their respective shares. As follows from Table 3.3, oil products' share of total tonne-miles is lower than its share of tonnes, indicating the fact that the average haul of oil products voyages is less than the average haul of world seaborne trade. If only considering crude and products, products' share of total trade drops from 35% of tonnes to 24% of tonne-miles. In line with theory (Talley, 2011), we observe that oil products voyages are significantly shorter than crude oil voyages.

Crude and Product Tankers

Crude tankers transport crude oils to refineries for processing. Seaborne transportation and distribution of distillates from the refineries to consumers is done by product tankers. Product tankers do also differ based on whether they carry *clean* (light and middle distillates) or *dirty* products (heavy and residue distillates, as well as, e.g., heavy crude oil)[15].

Liquid bulks require transportation in tanks and to

[13] Oil tankers are hereinafter simply referred to as tankers.

[14] Commonly referred to as "top of the barrel" products.

[15] Another common definition is to divide oil trade into dirty (crude oil and dirty oil products) and clean (clean oil products) trade. Dirty tankers may further carry, e.g., bitumen, molten sulphur, coal tar and pitch.

Table 3.3: Seaborne trade. Tonnes and tonne-miles of world cargoes (Clarksons Research Services Limited, 2017).

Cargo type	Million tonnes	Share of total (*oil*)	Million tonne-miles	Share of total (*oil*)
Crude oil	1,950	17.5% (*0.65%*)	9,578	17.4% (*0.76%*)
Oil products	1,067	9.6% (*0.35%*)	3,042	5.5% (*0.24%*)
Total Oil	3,017	21.1% (*100%*)	12,620	22.9% (*100%*)
Gas	355	3.2%	1,463	2.7%
Chemicals	283	2.5%	991	1.8%
Containers	1,727	15.5%	8,592	15.6%
Iron ore	1,411	12.7%	7,912	14.4%
Coal	1,140	10.2%	4,968	9.0%
Grain	480	4.3%	3,396	6.2%
Minor bulk	1,864	16.7%	10,934	19.8%
Other dry	852	7.7%	4,236	7.7%
Total cargo	11,130	100%	55,112	100%

be handled by pumping systems (Talley, 2011). In general, tankers can be identified by the flush freeboard deck with a series of pipes and vents covering the deck (Beard, 2011). Furthermore, due to oil spillage disasters over the history, regulations have been enforced so that tanker newbuildings now must be double hull vessels; seeing a final phase-out of single hull ships in 2020[16]. The degree of technicality of tankers is not particularly high compared to other, more specialised segments, but more advanced than dry bulk carriers.

In contrast to crude carriers which only carry crude oil, product tankers transport several batches of different distillates simultaneously. Coating of the tanks in clean product tankers is one of the main differences that allow for transportation of distillates (Jeffries, 2017). Clean product tankers also requires more sophisticated pumping and piping systems to support the number of cargo segregations. Dirty product tankers, or *high heat tankers*, carry dense, viscous cargo that require high heat to flow smoothly through piping systems. High heat tanks, or floating tanks, are allowed to expand and ensure a cargo temperature up to 250°C. Consequently, dirty product tankers demand higher sophistication than clean product tankers with regards to heating requirements, insulation, as well as valves and pumps (Wärtsilä, 2018), but have less complex segregation systems. Spread in freight rates between clean and dirty product tankers may cause some tankers to switch from clean to dirty or vice versa. Clean tanks require less cleaning than dirty tanks before loading new product, and would need a lower spread in freight rates to be sufficiently incentivised to switch. Cleanup of crude tankers is, nevertheless, very rare nowadays due to the high expense.

[16]The International Maritime Organization (IMO (2017)) enforced a 2015 phase-out date, but the Nigerian Maritime Administration and Safety Agency (NIMASA) extended the deadline to 2020 for Nigerian tankers.

Tanker Fleet

The tanker fleet comprises about 2,000 and 3,000 crude and product vessels, respectively. This corresponds to 19% of the merchant fleet, and 30% in capacity terms (>10 k.dwt) (Clarksons Research Services Limited, 2017). Another common measure of oil trade is *barrels of oil (bbl)*. In 2016, world seaborne exports of crude oil and oil products was 39.2 and 23.1 million barrels per day, respectively. A Very Large Crude Carrier (VLCC) tanker vessel of 270,000 dwt is capable of loading approximately 2 million barrels of oil[17]. Nonetheless, there is significant differentiation in dwt capacity across tanker vessels.

The highest volumes of crude oil are traded over intercontinental routes with the largest tanker ships. *VLCC* and *Suezmax* are the main ship types employed on long-haul voyages, while *Aframax* ships are usually employed on short to medium-haul voyages. Aframax ships are also used for carrying oil products. Aframax is considered to be the "workhorse" among the tankers, due to its ability to access most ports around the world. Smaller sizes, such as *Panamax* and *Handysize*, mainly carry oil products. Some ship classes are named based on the canal that will be transit-restrictive in terms of size[18] (e.g., Suezmax vessels in a laden condition are capable of transiting through canals that are of equal or larger size than the Suez Canal). VLCCs are too large to pass through any canal in laden, and must transit around the great capes of the world. Aframax, Panamax and Handysize ships used for oil products trade are often referred to as LR2, LR1 and MR, respectively[19]. An overview of the tanker fleet can be seen from Table 3.4.

As ships are being built larger, the ULCC class has

[17]1 barrel equals 159 litres. Assuming 0.88 specific gravity, a metric tonne then equals about 7 barrels of oil.

[18]Aframax vessel is named not because it travels around Africa, but after the Average Freight Rate Assessment (AFRA) tanker rating system.

[19]LR: Long Range. MR: Medium Range. Serves the purpose of separating crude and product tankers in the same dwt range.

been introduced in the upper range along with VLCCs. ULCCs are used almost exclusively on the longest long-haul routes, such as the Middle Eastern Gulf to the US Gulf and East Asia (Chapter 4). The largest vessels are also more commonly used as speculative *oil hubs* outside major import markets in certain market situations (Tradewinds (2017a); see discussion about contango oil pricing in Section 5.3). Table 3.5 shows the development of the productivity of the tanker fleet, measured as total million tonnes of oil traded by sea to total fleet tonnage. This indicates that tankers have become less productive following the period between 2005 and 2010. Demand has thus weakened relative to supply. Although supporting demand, some of the explanation may be that tonnage are now being transported over longer distances than before. Increased use of vessels as storage units, or oil hubs, may be a further support to demand. From Table 3.5, at least for crude, indications of a recovery can be seen. Furthermore, by dividing tonne-miles by tonnes an explicit measure of the average distance the cargo travels can be made. Results from this is also given in Table 3.5. We see a tendency of upward movement from 2015 to 2017 in average haul, which will strongly support a recovery from today's distressed rates if more oil flood into the market. Discussion on trade flows are further given in Chapter 4.

Market Characteristics

Oil companies are dependent upon independent shipowners' transportation capacities. There are several explanations of why oil charterers' interests ensure the need of independent shipowners (Talley, 2011). Inspiration taken from Talley (2011) lead us to identify five explanations. *Firstly*, there exists uncertainty of transportation volumes of oil. *Secondly*, the buying and the selling side of oil do not want to be dependent on the other for transportation. *Thirdly*, specialisation suggest more favourable operating economic conditions, as well as exemption from certain politics and governance. *Next to last*, cost of capital considerations of purchasing a vessel is avoided. *Fifth and lastly*, tonnage supply regulates in line with transportation demand. Accordingly, it is evident that a monopolistic structure of oil companies on tanker transportation is not in these companies' best interest.

Stopford (2009) identifies shipping markets as textbook examples of markets operating under conditions of nearly perfect competition, or efficient markets. Five characteristics are met under a perfect competition market structure (Investopedia, 2018b), which we translate to the tanker market. *Firstly*, shipowners offer almost identical products. Tanker technology is generally homogeneous within the various classes of tankers, and tankers are as such perfectly interchangeable. *Secondly*, all shipowners are price takers. The freight rate mechanism sets the price. *Thirdly*, the tanker market is known to have a high level of fragmentation. The top 20 ownership distribution of oil tankers portrays that these companies only own 34.5% of the tanker fleet capacity[20]. No single owner owns more than 2.5% of the total capacity; as of today, this is the privately owned Iranian company National Iranian Tanker (Clarksons Research Services Limited, 2017). In fact, economies of scale is not particularly evident in terms of the number of vessels a company owns. *Fourthly*, charterers have complete information about the worldscale rates charged on various routes, as well as tankers available for chartering. Freight rate manipulation is difficult with this transparency. *Lastly*, the market is characterised by ease of entry and exit. No complex administration structures are required for operations. With the necessary financing, everyone can order a newbuilding, and scrapping a vessel is, in most respects, straightforward. The ease of exit also translates to tankers' mobility of switching to routes in which freight rates are more profitable. Furthermore, referring back to the introduction of this chapter, we note that the tanker market comes under the category of tramp shipping, wherein vessels are hired on-demand. In other segments, such as those involved with liner shipping, obstacles of entry are greater.

In the context of freight rate prediction, these market characteristics possibly indicate that tanker market dynamics exhibit predictable behaviour over time, and hence serves as a theoretical support of translating past history into future predictions.

[20]Total capacity of owners with more than 10k dwt in their fleet.

Table 3.4: Oil Tanker Fleet above 10k.dwt (Clarksons Research Services Limited, 2017).

Vessel class	**Vessel size range** k.dwt	Avg. k.bbl	**Number of vessels**	**Fleet Capacity** k.dwt	Share of fleet	**Avg. age** (year)	**Typical haul**
Crude Tankers							
U/VLCC	200 - 442(d.d)	2,102	734	225,686	41%	9.3	Long
Suezmax	120 - 200	1,055	543	84,740	15%	9.5	Med-Long
Aframax	80 - 120	742	658	71,483	13%	11.1	Med-Short
Panamax	60 - 80	481	89	6,212	1%	12.9	Short-Med
Total Crude			**2,024**	**388,121**	**71%**	**10.1**	
Product Tankers							
Suezmax	120 - 200	1,072	16	2,539	0.5%	9.3	Long
LR2	80 - 120	804	342	37,354	7%	7.9	Med-Long
LR1	60 - 80	514	360	26,426	5%	9.3	Med-Long
MR	30 - 60	307	1,961	88,389	16%	10.2	Med-Short
SR	10 - 30	108	277	4,114	1%	16.4	Short
Total Product			**2,956**	**158,823**	**29%**	10.4	
Total Oil Tanker Fleet			**4,890**	**546,944**	**100%**	**10.3**	

Table 3.5: Tanker fleet productivity: Tonnes carried per dwt of tankers & Tonne-miles carried per dwt of tankers & Average haul in miles - dividing tonne-miles by tonnes. Includes tankers above 10 k.dwt (Clarksons Research Services Limited, 2017). *NA: data not available.*

A more detailed breakdown of each productivity constituent is given in Appendix C.

Year	**Ratio: Tonnes carried per dwt of**			**Ratio: 1k. tonne-miles carried per dwt of**			**Average haul [miles]**		
	Crude	Product	Total tankers	Crude	Product	Total tankers	Crude	Product	Total tankers
1990	5.6	NA	NA	26.6	NA	NA	4,719	3,040	4,268
1995	6.9	NA	NA	32.7	NA	NA	4,765	3,087	4,372
2000	7.6	10.50	8.2	36.5	29.4	35.1	4,803	2,797	4,031
2005	7.8	10.39	8.4	35.7	30,9	34.6	4,580	2,978	4,140
2010	6.3	7.56	6.7	29.4	22.3	27.4	4,641	2,950	4,099
2015	5.4	7.35	6.0	26.4	20.9	24.8	4,855	2,844	4,144
2017	5.2	7.04	5.8	26.4	20.1	24.6	5,042	2,861	4,270

4 Tanker Trade and Routes

Oceans, continents and ports are all important geographical terms to consider in the context of intercontinental tanker trade. Three economic centres dominate the world of shipping; North America, Europe and Asia. The most predominant oil trades are taking place in the oceans and seas between these continents. Shipowners face several logistical matters when employing the ships to carry out these trades, such as acceptable freight rates, size restrictions, voyage durations, voyage costs, labour, and congestions and disruptions at sea-chokepoints, canals and ports.

At first, we will briefly look at how the tanker route pattern has changed over the last decades. Then, we will look at how trade flows are today, and emphasise some of the major trading routes in terms of imports and exports of oil. Next, we will select and introduce more in-depth the routes of which freight rates we will analyse. Finally, we will look at the background of the freight rate data that we will use throughout this paper.

4.1 Tanker Trade

Historical Tanker Trade

Trade change, along with changing economies, is imperative to get right, as described by Stopford (2009): «One of the most fundamental principles of trade forecasting is to recognize [trade change and changing economies] and build it into the forecast». The choice of relevant determinants heavily rests on acknowledging this statement, moreover motivating somewhat why a model should be conditional on varying market conditions. In other words, economic forces and relationships important for trade in the past, may be of less importance today. Over the history, tanker trade has undergone a continuous development in trade dynamics.

Crude Oil. The first charter of crude oil by sea was conducted in 1861. Oil tanker trade is accordingly far younger than several other types of shipping trade. Tanker vessels carrying oil in bulk, using the outer ship-hull as tank compartment, were not seaborne before 1886. The Suez Canal opened for tanker transit in 1892, which considerably shortened voyage distances, and has since been shut down, reopened and enlarged several times. At this time, shorter distances supported demand tonnes more than it teared on demand miles.

Tonne-miles demand grew rapidly in the mid-1900s as the Middle East exported more and more oil, especially to Western Europe through the Suez Canal. The increase in average haul was profound during this period. Shipping costs and the price of oil was about the same in the 1950s (today, the price of oil is way higher than shipping costs). Oil majors, who also were tanker shipowners at the time, eventually faced high shipping costs by building supertankers (UL/VLCCs), thus exploiting economies of scale. The first VLCC sat sail in 1966. From then on, the tanker fleet grew rapidly.

Fleet growth combined with a decline in oil trade caused the tanker market to plunge in the late 1970s. Maturing transition from coal to oil as an energy source in Europe and Japan was part of the cause[21]. High oil prices, along with economic recessions were additional contributors. Furthermore, the average haul was struck by increased production of oil trading short-haul (e.g. North Sea), opening of Middle Eastern refinery capacities and pipelines, increased domestic production in importing regions (e.g. North America), as well as the reopening of the Suez Canal in 1975 following the closure in 1967. Smaller vessel classes, performed well under the growing short-haul trade, while VLCCs moved to medium-haul trades in the Atlantic. In 1986 the freight market improved again, when lower oil prices supported demand for Middle Eastern oil. From the beginning of the 1990s, and up until today, seaborne oil trade has grown greatly, but also been subject to cyclical downturns.

Oil Products. The history of crude oil trade and oil products trade is related. However, trade of products is different from crude trade. Not only in terms of cargo and vessel sizes and technicalities, but also from a geographical and route perspective. We mentioned oil refinery capacity in the context of crude trade above. Before 1960, the share of products tonne-miles to total oil tonne-miles was greater than it is today. In the 1960s, cost benefits of moving crude to refineries in proximity to products demand was brought to life in much larger scale, which was further supported by political matters. Western Europe, especially, expanded its oil refinery base substantially at the end of the 1950s. These refineries are still majors players in the oil trade today (see, e.g., port of Rotterdam in the description of «TC2» below). Naturally, the route patterns of crude and products trade were adjusted thereafter. Europe was now importing crude instead of products from the Middle East, and US built up their own refinery capacity instead of importing products from South America and the Caribbean. After that, constrained capability of building additional refining capacity and the need of balancing trade of various oil products, among other things, lead to increasing trade into US and Europe in the 1990s. From the 1990s, Fear East Asian products demand has had a solid growth until today (for detailed history, see Stopford (2009); Talley (2011)).

In summary, seaborne oil trade has developed from

[21] Maturing trade import of iron ore to China is today a potential concern for the Dry Bulk industry (Clarksons, 2017a).

a less speculative mode of trade, carefully planned by oil majors, to a volatile, market regulated industry. Natural resources and geographical locations play a vital role in the seaborne route pattern, and tonne-miles demand that follows – conflicted by a variety of stakeholders filling different roles.

Tanker Trade Today

The route pattern today consists of a comprehensive network of routes going across the entire world. Routes are commonly referred to with route codes as they are presented in the Baltic indices, whereof TD and TC are route codes for the BDTI index and the BCTI index, respectively. A complete list of the routes included in BDTI and BCTI can be seen in Appendix A.6.

Table 4.1 provides an overview of regional tonne-miles importers and exporters of crude oil and oil products. By analysing seaborne trade tables over time, one may be able to capture important changes in trade dynamics. The Middle East is by far the largest exporting region of oil in the world[22]. Due to the region's geographical location, several long-haul routes exist between the Middle East and large importing regions such as the Far East and North America. These long-haul routes contribute significant to tonne-miles demand. Understandably, an increase or decrease of exports from the Middle East will have large effects on tanker freight (as history above revealed). In terms of import, the Asian continent is the largest for both crude and products. Furthermore, Table C.4 in Appendix C provides a list of top 10 producers, consumers, importers and exporters of oil.

On the products side, analysts would track oil refinery capacity to see what distances the crude must travel. Surplus refinery capacity to oil products demand would be an indicator of potential exports of oil products. In terms of deficit trade, keeping track on domestic refinery capacity and output to production and demand may reveal a country's capability to export surplus products or crude; or the need of importing the deficit of either or both.

4.2 Selection of Routes

As previously mentioned, we will be looking in-depth at specific routes for the tanker market, and later attempt to model respective freight rates of these. But which routes should be considered? Our selection of routes is based on the following criteria:

1. **Route:**
   Major oil route, and part of the BDTI or the BCTI.
2. **Data time horizon and frequency:**
   Data available back to year 2002, and on a monthly frequency (see, also, Chapter 6).
3. **Ship type:**
   Both crude, and product vessels.
4. **Ship size:**
   Both similarity, and variation in vessel size.
5. **Freight rate (TCE) correlation:**
   As much variation as possible between the selected routes in terms of freight rate correlation.

After addressing the criteria above, the corresponding freight rates for the following routes were chosen: **TD1**, **TD3**, **TD7**, **TD12**, **TC1** and **TC2**. Table 4.2 presents an overview over the selected routes, and Figure 4.1 displays these routes on a world map. With these routes, we get a mix of routes that makes it possible to compare several interesting relationships; the same vessel size and oil trade, but different routes (TD1, TD3); crude oil and oil products trade on the same route (TD3, TC1); along with relationships regarding vessel sizes and routes across both types of trade. Next, we will give a description of some characteristics of each route.

Route Characteristics

Crude oil round voyages are often made up of one laden leg and one ballasting leg, i.e. vessels are only transporting cargo one way. Product tankers on the other hand, are not unusually transporting backhaul cargo on a round voyage, i.e. vessels are loading and carrying cargo both ways on a round voyage. These are important facts that have been considered when selecting determinants (Chapter 5). In regards to loading and discharging, terminals for UL/VLCCs normally have jetties[23] in deep waters outside a port. This is due to the massive beam and draft (size) of the largest tankers. These terminals are thus often referred to as offshore terminals. Ships discharge cargo using their own pumping systems, and cargo is pumped via offshore pipelines that are connected to storage tanks on-shore. When ships are loading oil, terminals' pumping capacity are used. Product tankers and smaller crude tankers are less size-restricted, so jetties are usually not needed. Product tanker terminals do, however, require a loading/discharge system that can manage a combination of products.

VLCC TD1: Ras Tanura - LOOP

Ras Tanura is the main export terminal of Saudi Arabia, and located in the Middle East Gulf (MEG). *Crude*

[22] About half of the seaborne oil exported originates from the Middle East (Talley, 2011).

[23] Jetty: in the same category as a pier, i.e. a structure that projects from the land out into the water.

Table 4.1: Seaborne trade by region. Import and export in 2017 of crude oil and oil products (Clarksons Research Services Limited, 2017).

Seaborne crude oil [b.tonne-miles]				Seaborne oil products [b.tonne-miles]			
Region	Export	Region	Import	Region	Export	Region	Import
AG	5,178	China	3,030	Far East	203	UK/Cont.	216
Caribs.	1,399	N.America	1,721	USG	190	L.America	136
WAF	1,320	Japan	1,000	UK/Cont.	149	China/Jap/Korea	94
U.S.*	242	India	732	AG	137	Med.	90
UK/Cont.	217	UK/Cont.	577	Baltic	101	N.America	89

*U.S. crude export is set to increase by 80% in 2018 according to CRSL forecasts.

Table 4.2: 6 selected routes, and corresponding route description. Route symbol in accordance to the Baltic Exchange indices.

Route symbol	**Route ports** (Countries)	**Vessel size**	**Cargo type**	**Oceans** (Seas)	**Baltic Exchange**
TD1	**Ras Tanura - LOOP** (Saudi Arabia - U.S.)	VLCC	Crude oil	Indian Ocean, Atlantic Ocean	BDTI
TD3	**Ras Tanura - Chiba** (Saudi Arabia - Japan)	VLCC	Crude oil	Indian Ocean, Pacific Ocean	BDTI
TD7	**Sullom Voe - Wilhelmshaven** (UK - Germany)	Aframax	Crude oil	(North Sea)	BDTI
TD12	**Antwerp - Houston** (Belgium - U.S.)	LR1	Dirty oil products	Atlantic Ocean	BDTI
TC1	**Ras Tanura - Chiba** (Saudi Arabia - Japan)	LR2	Clean oil products	Indian Ocean, Pacific Ocean	BCTI
TC2	**Rotterdam - New York** (Netherlands - U.S.)	MR	Clean oil products	Atlantic Ocean	BCTI

oil is imported to refineries in the US Gulf area after being discharged at the Louisiana Offshore Oil Port (LOOP). Once reaching the refineries, the oil is refined and distributed through a network of oil pipelines. VLCCs must transit the most important oil chokepoint in the world, namely Strait of Hormuz connecting MEG and the Arabian Sea. In 2016, 18.5 million barrels flowed through the strait (EIA, 2017c). At sea, VLCCs are sailing around the Cape of Good Hope, hence crossing over from the Indian Ocean to the Atlantic Ocean. Notable for the TD1 route, is that VLCCs may not perform a complete round voyage, but rather take on charter parties that lead them to ballast to, e.g., West Africa (WAF) for new cargo loads.

The port of Ras Tanura has several deep-water terminals that are linked to relatively short pipelines from large-reserve on-shore oil fields, securing this oil to be readily available for exports. The terminal has a series of storage tanks that are connected to jetties which VLCCs berth to. LOOP is located in the US Gulf off the coast of Louisiana. The terminal is a designated oil lightering terminal. In other words, oil is being transferred from VLCCs to smaller vessels, which again transports oil to ports along the Gulf that the VLCCs are size-restricted to enter. LOOP is the only deep-water oil terminal in the U.S. capable of offloading VLCCs. Pipelines are also stretching from LOOP to ports on-shore, connecting to 35% of the U.S. refining capability (Stopford, 2009). Going forward, LOOP will play an important role in the export of U.S. crude, as the U.S. will increase its exports substantially following the lifting of U.S. crude export ban in 2015.

EIA divides the U.S. mainland into to five so called Petroleum Administration for Defense Districts (PADD) (EIA, 2012). Of relevance to this thesis, are PADD1 and PADD3, which represents the Gulf Coast and the East Coast, respectively. EIA's database includes data on PADD-specific level, which our inclusion of determinants in Chapter 5 reveals.

VLCC TD3 *and* LR2 (Aframax) TC1: Ras Tanura - Chiba

Vessels on the TD3 (*crude*) and the TC1 (*clean*) route are sailing from Ras Tanura, Saudi Arabia, in the Middle East Gulf to Chiba, Japan, on the east coast of Japan. On the voyage, VLLCs and LR2s transit eastward through the Strait of Malacca. The Strait of Malacca, linking the Indian Ocean and the Pacific Ocean, is a key chokepoint for the trade between

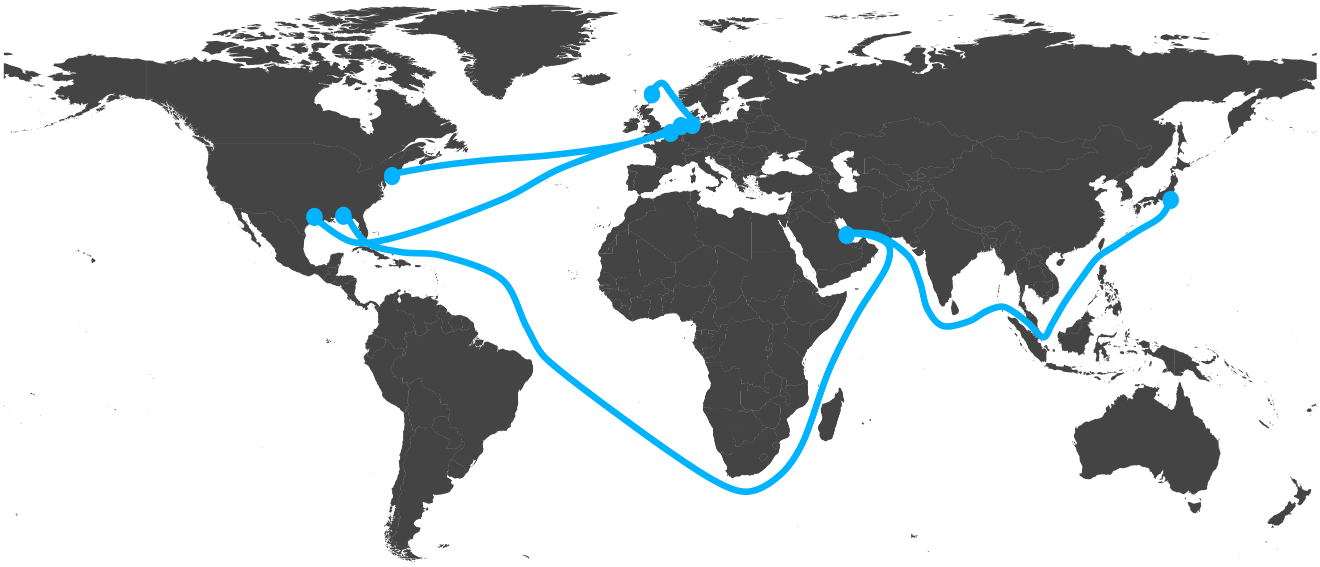


Figure 4.1: An overview of the specific tanker routes which are investigated in this thesis; spanning nine ports, seven countries and four regions. TD1: Ras Tanura [SA] - LOOP [US], VLCC; TD3: Ras Tanura [SA] - Chiba [JP] VLCC; TD7: Sullom Voe [GB] - Wilhelmshaven [DE] Aframax; TD12: Antwerp [BE] - Houston [US] LR1; TC1: Ras Tanura [SA] - Chiba [JP] LR2; TC2: Rotterdam [NL] - New York [US], MR.

the Middle East Gulf and Japan. According to EIA (2017b), nearly 16 million barrels of crude oil and oil products per day transited the strait in 2016. Clearly, the Strait of Hormuz is also a key chokepoint on this route (see TD1). The port of Chiba is Japan's second largest port in terms of cargo tonnage handled (Find a Port, 2017).

Japan is one of the major contributors to seaborne imports in the world, which is supported by their large-scale industrial capacity. Japan does possess very limited natural resources, and are highly dependent on foreign energy supply to cover energy demand. Most of Japan's imports of oil come exactly from the Middle East. Countries in the Middle East, relative to other world regions, are namely generally rich on natural resources. The economy of Saudi Arabia is heavily dependent on its exports of oil. The Middle East is one of the world's conflict areas, and oil trade has been disrupted several times over the history. Political instability, and the pressure it has on seaborne freight, is unfortunately difficult to safeguard against for shipping participants.

Aframax TD7: Sullom Voe - Wilhelmshaven

On the TD7 route, vessels transit from Sullom Voe, UK, in the North Sea to Wilhelmshaven, Germany, at the northern German coastline. Vessels on this route load crude oil from the storage terminal Sullom Voe on the northern Shetland Islands. Sullom Voe receives its oil from offshore oilfields in the Shetland Basin. According to Enquest, the terminal operator, the terminal has a current inflow rate of 0.13 million barrels per day (EnQuest, 2017). The port of Wilhelmshaven has Germany's only deep-water oil terminal and is the largest import port for crude oil (Niedersachsen, 2017). Similar to Ras Tanura, jetties handle the transhipments of oil.

The voyage to Wilhelmshaven is free of oil chokepoints, but bad weather causing colossal waves represents possibly the most significant bottleneck in up- (oil production) and midstream (tankers) operations in the North Sea. Histories of tankers running aground in the waters off the Shetland Islands do exist.

LR1 (Panamax) TD12: Antwerp - Houston

TD12, from Antwerp, Belgium to Houston, U.S., in the US Gulf is a dirty products route. The voyage is fairly straightforward, and no major oil chokepoints exist on the route, even though port congestions and unfavourable climate conditions occasionally occur. For instance, the Atlantic hurricane season in the U.S. Gulf may tighten available tonnage supply and cause tanker rates to spike, historically, around August (Teekay, 2017).

The port of Antwerp has a long merchant history, in which the 1500s marked Antwerp's golden age (Port of Antwerp, 2017). Today, Antwerp is a main competitor to, e.g., Rotterdam (see TC2 below). The port of Houston, on the other hand, is the largest port on the U.S. Gulf Coast (Port Houston, 2017). The aforementioned PADD3 district is thus relevant to this route also.

MR (Handysize) TC2: Rotterdam - New York

Vessels on the TC2 route cross the Atlantic Ocean with clean products cargo from Rotterdam, Netherlands to New York/New Jersey, U.S., at the US North-East Atlantic Coast[24]. Like TD12, the voyage is clear of major oil chokepoints.

Following the port of Antwerp's golden age, Dutch ports took over as Europe's maritime capital. Rotterdam[25] is today Europe's largest port, and a distribution centre for the European continent (Port of Rotterdam, 2017). The port of New York and New Jersey is the third-largest port in the U.S., and the largest on the East Coast (Port of NY/NJ, 2017). Above-mentioned PADD1 district is thus relevant to this route also.

Route Analysis

TD1, Ras Tanura – LOOP, is the route with the longest voyage duration among our selected routes, with approximately 78 days or 24,4450 miles if a complete VLCC round voyage is made. In comparison, Ras Tanura – Chiba has a duration of about 43 days or 13,308 miles for the same vessel (voyage details for all routes follow in Table 6.2). It is evident that the future development in route patterns is important to consider when analysing tonne-miles demand.

Oil moves from areas of surplus to areas of shortage, defined by Stopford (2009) as *deficit trade*. In terms of oil trade, deficit trade is undoubtedly the most relevant and make up the largest portion of trade. In addition to deficit trade, Stopford (2009) argues that we can look at two additional economic forces as underlying drivers of trade; *competitive trade* and *cyclical trade*. Competitive trade occurs when countries that are capable of producing oil, imports oil at a lower cost instead - this is what economists and market analysts commonly refers to as arbitrage trade. Cyclical, or temporary trade refers to trade (or the lack of it) due seasonal or temporary shortages. For instance, unusual cold winters will lead to unusual high demand for heating oils, failure on oil rigs in countries of shortage will suggest more import, and instabilities (war, politics etc.) in oil producing countries may lead to disruption of oil supply (and consequently oil trade). In Figure 4.2, we have visualised spreads in crude production and demand by countries/area of relevance for the selected routes. This serves to identify the countries that are in surplus or shortage of oil, and at what level they are in the need of oil trade; either crude oil for domestic refining or oil products for direct use.

[24]The TC2 is often treated as part of a triangulated Atlantic route including a ballasted leg between the US East Coast and US Gulf Coast (Tradewinds, 2017b).

[25]MEG-Rotterdam is one of the key benchmark indices used by owners, brokers and charterers to describe the health of the VLCC market (Kavussanos and Visvikis, 2016).

Viewed simplistically[26], the following can be read from the graphs in Figure 4.2 given a **negative** (*positive*) value in the spreads:

- **Spread 1)** Refinery Output − Oil Demand: **Import of products; or draw in inventory** (*export of products; or add to inventory*)
- **Spread 2)** Crude Oil Production − Oil Demand: **Import of crude and/or products; or draw in inventory** (*export of crude and/or products; or add to inventory*)
- **Spread 3)** Crude Oil Production − Refinery Output: **Import of crude; or draw in inventory** (*export of crude; or add to inventory*)

  *i.e. Spread 2 *(S2)* − Spread 3 *(S3)* = Spread 1 *(S1)*

EU4 (Germany, France, UK, Italy) is included here, since Rotterdam, Antwerp and Wilhelmshaven are all distribution centres for European oil trade. EU4 thus serves as a benchmark for European oil trade. Overall, Saudi Arabia is the only country not in the need of imports. For instance, U.S. oil demand exceeds U.S. oil production with approximately 10.7 million barrels per day (Sept. 2017). The U.S. is therefore dependent on imports to cover demand. Yet, balancing trade does occur. Countries in shortage of oil do therefore also export certain amounts of oil.

In the summer months of 2017, Saudi Arabia decided to cover high domestic demand, caused by the Middle East heat season, by cutting down on exports instead of increasing production. This was to support OPEC's cut-deal on production from November 2016. The main aim of OPEC with this cut deal is to constrain supply to increase oil prices. From Figure 4.2, we can see a clear dip in the Saudi Arabian spreads at the end of 2016, which have been translating negatively to the tanker market. OPEC compliance with output cuts is not impacting tanker trade favourably; at least in the short-term, until high oil prices potentially accelerate sanctioning of oil projects. Similar macroanalyses may be made to uncover relationships in other spreads. We will elaborate on these relationships in more detail in the determinants chapter (Chapter 5).

[26]This should, however, possibly be seen in the context of refinery capacities (and/or refinery utilisation) also, to get the full picture of deficit trade. Furthermore, changes in inventory (stock-buildings) may support analysis. Unfortunately, we do not have adequate data to perform such analysis.

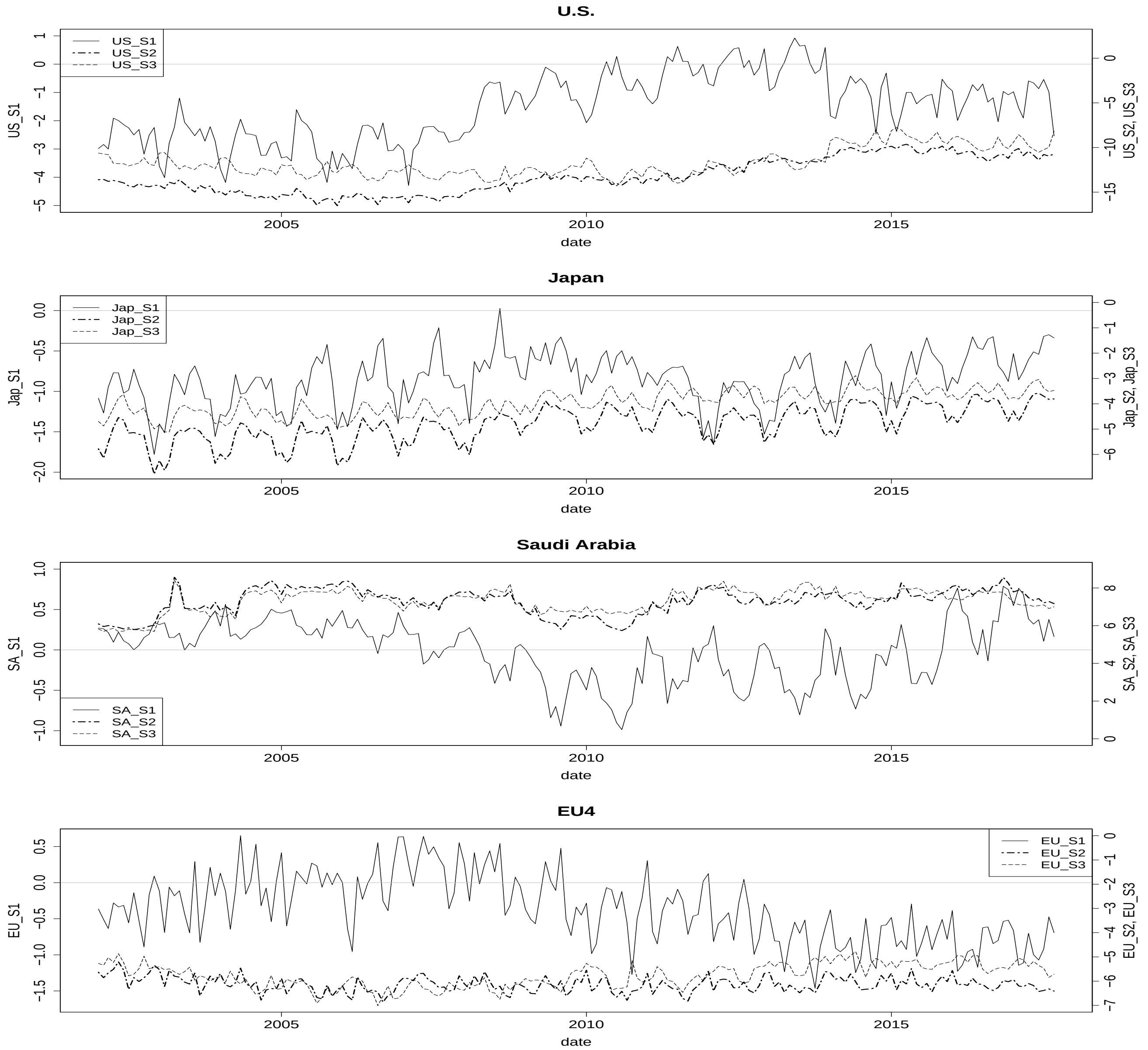


Figure 4.2: Figures from top to bottom: **USA**, **Japan**, **Saudi Arabia**, **EU4** (Germany, France, UK, Italy).

Spreads between combinations of Refinery Output, Oil Demand and Crude Oil Production for countries relating to our 6 selected routes (TD1, TD3, TD7, TD12, TC1, TC2). Serves as a visualisation of oil shortage or surplus.

Spreads are calculated as: S1) Refinery Output − Oil Demand ; S2) Curde Oil Production − Oil Demand ; S3) Crude Oil Production − Refinery Output. Numbers in million barrels per day *(Mbbl/d)*. S2 and S3 on secondary axis (RHS).

* Be aware that Saudi Arabian Refinery Output and Oil Demand are rated as "Not assessed" or "Use with caution" by the data source (Joint Organisations Data Initiative (JODI, jodidb.org)).

5 Factors Affecting Tanker Freight Rates

In the consecutive parts of this chapter, we will introduce hypothetical *route-specific* and *general* determinants for our 6 selected routes. We consider route-specific factors as factors solely focused on a single route. General-specific, or common factors, concerns either all routes or some subset of routes - A VLCC determinant might for instance affect more than one route. We group determinants[27] into three groups: Supply, Demand, and Economic and Non-fundamental (E&N). The classification of each group is, generally, based on the following:

i) **Supply:** ship and fleet related factors

ii) **Demand:** cargo trade, vessel fixtures, production and oil demand factors

iii) **E&N:** a variety of market price, financial and other factors possibly affecting supply and demand

In the rest of the chapter, a combination of market theory and empirical theory will be used to justify the relevance of the determinants we will be considering for the regime models.

Macroeconomic Landscape

UNCTAD (2016) claims that maritime trade flows continue to be largely determined by developments in the macroeconomic landscape. These are variables that hypothetically could explain the whole or subsets of the whole tanker market. On the contrary, microeconomic variables are individual ship specifications, charter party details and so on (see, for instance, Köhn (2008); Alizadeh and Talley (2011); Adland and Cullinane (2006); Abouarghoub et al. (2017)). Multiple empirical studies over the history find macroeconomic data to be of great influence on tanker freight rates. Findings include determinants such as global economic activity, oil prices, ordered newbuildings, growth in industrial production, trade in commodities at sea, deliveries of new vessels and the scrapping rate (Hawdon (1978); Beenstock and Vergottis (1989); Abouarghoub et al. (2012); Abouarghoub et al. (2017)). *Our models will as such be based on macroeconomic variables.*

Abouarghoub et al. (2012) identifies that both exogenous and endogenous factors affect shipping business cycles. Exogenous factors belong to *shipping demand*, while endogenous factors are concerned with *shipping supply*. The fact that demand for oil seaborne trade is derived demand, supports empirical findings of freight rate dynamics being influenced by macroeconomic events (Stopford (2009); Abouarghoub et al. (2012)). Global exogenous, macroeconomic events such as the most recent global boom caused by the Chinese activity in mid-2000s or the financial crisis in 2008, lead to significant changes in the demand for shipping services, which in turn lead to endogenous capacity adjustments of shipping supply. Supply adjustments must be considered in the light of both actual events and future expectations. The level of endogenous reactions to exogenous factors, i.e. lead time for cycle transition, will further depend on the current capacity and fleet utilization situation. Kavussanos and Visvikis (2016) argue that short-term cycles are largely supply driven, with a few exceptions from external shocks - «originates from timing effects and mass psychology on very fragmented markets with low entry barriers for vessel ownership». Supply regulating behaviour is supported by maritime economic theory of efficient shipping markets (see Section 3.5).

Since we in this study are dealing with 6 different routes and a manifold of variables, factors will be grouped into subgroups when presented here. For instance, all determinants on oil import and export are gathered into a subgroup called «Oil Import and Export», belonging to «Demand»[28]. We will also address the various determinants' hypothesised directional (positive) impact on freight rates. Table 5.1 provides a complete overview[29]. At the end, a brief discussion on "non-capturable» factors and factors that could have been useful is given.

5.1 Supply-Driving Factors

Shipping supply is broadly given by the available fleet's tonne-miles capacity. Tonne-miles is made up of two factors, namely deadweight tonnage and productivity (see, further, Section 3.2). The former is quite unambiguous. The latter is imponderable; non-easy access or non-capturable data. Productivity refers to speed, deadweight utilisation, port time, loaded days at sea and lay-ups[30] (Stopford, 2009). Other than productivity factors, supply relies on factors such as the size of the fleet, shipbuilding activities and the rate at which ships are scrapped (Alizadeh and Talley, 2011).

Aforementioned theory suggests that supply factors exhibit cyclical behaviour to freight rate movement, i.e.

[27] Interchangeably referred to as «factors», «determinants» and «external variables» etc.

[28] Instead of listing and mentioning all single individual variables that we have tested, we refer to the Appendix, although some specific variables are mentioned where found natural.

[29] Hypothesises are based on a "most cases" basis. Thinking of all possible cases where a directional impact would yield opposite effect will make our heads spin. Luckily, we have statistical measures to reveal the «truths».

[30] Factor added by authors.

they are endogenously related. Cyclicality causes the signs of correlation to change across the lag structure. When hypothesising the expected coefficient sign in the regressions models, we therefore consider real-time (0-lag) correlation.

Fleet Size

The size of the tanker fleet refers to the total deadweight capacity (dwt) in the market, i.e. it is a direct measure of shipping supply. In the short term, the capacity is fairly constant and not subject to large fluctuations. Medium to long-term, capacity will move up and down along with newbuildings (deliveries) and demolitions, respectively. When demand is low, typically only a portion of the fleet will be on charter, while surplus capacities are laid-up and/or marketed in the S&P market.

We have linked fleet determinants for VLCCs, Aframaxes, LR2s, LR1s (Panamaxes), and MRs (Handysizes) to respective routes, including dwt (current and yr/yr growth), contracting, orderbooks, deliveries, demolitions and removals.

We expect to see distinctions in the dynamics of models concerned with different vessel sizes. See for instance Jing et al. (2008), who find that asymmetric characteristics are distinct for different vessel sizes and market conditions in the dry bulk market. van Dellen et al. (2011) draw similar conclusions for freight rates in both the dry bulk and the tanker market, where they conclude that different vessel sizes regardless of trade (e.g. crude vs. products) are suited for different models (see also Abouarghoub and Mariscal (2011); Abouarghoub et al. (2012)). Moreover, maritime theory suggest that smaller vessels have greater ability to switch to different routes and cargoes (IAME, 2014). Accordingly, freight rates for larger vessels may prove to be more sensitive to, e.g., fleet capacity adjustments.

Supported by maritime theory in the introductory chapters, we hypothesise positive and negative relationship to freight rates for capacity down-adjustments (demolitions, removals) and up-adjustments (dwt increase, deliveries), respectively. We do not, directly, consider possible effects of interchangeable fleets, i.e. capacity may move to other routes/markets. In other words, we do not include fleet size determinants across routes (e.g. Aframax on VLCC routes, neither total tanker fleet). However, (as mentioned earlier), we might discover higher volatility among larger vessels, attributed to small vessels ability to interchange.

Newbuilding Contracts. Newbuildings require huge capital investments, and the lead time from contract signing to delivery could be lengthy (1-4 years). The history reveals that the profitability of a newbuild investment is heavily dependent on the ability of timing the cycle, and thus is very reliant on accurate forecasts (Fearnley, 2018). Timing decisions are complex. Several shipowners do not have the liquidity of taking on high risk in trough periods. Hence, they rather tend to contract newbuilds when the market is strongly recovering or about to peak. This could in turn cause a rapid surplus capacity if the market is in decline during delivery, which would lead to an unavoidable collapse. If a shipowner company, nevertheless, manage to time delivery against a peak, cash inflows could yield a considerably competitive advantage in the trough phase to come. Figures in Appendix illustrate the close positive relationship between newbuilding contracts and freight rates (Figure D.3). Such relationship was acknowledged by Zannetos (1964) as far back as in 1964.

Maritime regulations affect the rate of contracting. For instance, double hull regulations in the 1990s, sulphur emission limitations in emission control areas in 2015, and the upcoming ballast water treatment contribute(d) to economic aging of existing vessels. Less efficient vessels are forced to leave the market as more eco-friendly and efficient vessels enter the market (Clarksons, 2017a).

Orderbook. The total number of newbuilding contracts shipyards currently have, make up the orderbook. Quotations of newbuilding prices depend upon orderbooks, and both measures follow a cyclical pattern towards freight rates. An increase in freight rates will trigger the placing of newbuilding orders, thus constraining capacity at shipyards and increasing prices. The opposite yields in periods of recession. Furthermore, orderbook lag effects indicate that an orderbook buildup will affect freight rates negatively in a two-three years' time. In real time, however, we anticipate a positive relation to freight rates.

Down-payment of a ship is usually spread over several instalments. Therefore, we could expect that the orderbook does not necessarily reflect the correct number of ships that will enter market, since cancellations or delays do occur (Fearnley, 2018). We have included orderbook numbers both in deadweight tonnage and percentage of fleet although we consider the latter to be most relevant (see, for instance, Tsolakis (2005) and references within).

Deliveries. We have already suggested that ordering activity of ships is high when freight rates are high. Deliveries on the other hand, is anticipated to have the opposite effect, both in real time and across several years of lags. This is expected when considering lead times of delivery in the context of cycle duration. Delivery tops therefore coincide with trough periods, as can be seen from figures presented in the Appendix (Figure D.3). So why is this? We mentioned in Section 3.3 that shipowners could be labelled with a certain degree of short-mindedness. One main reason is that, during boom markets, a potential great profit can be made. If you do not catch the train, some other will - which in its own sense accelerates the cycle process. Some will unfortunately always catch a train heading for a cliff (Clarksons, 2017a). It is fascinating how the history repeats itself. Paradoxically, if everyone follows a recommendation of ordering vessels at the bottom (counter-cyclical activity), newbuilding prices will inflate, and once delivered, will push freight rates down and consequently prolong the cycle. Yet, over the history, shipping magnates have succeeded well in buying cheap vessels during recession periods.

Demolition and Removal. Ship demolitions, or scrapping, reduce the tonnage of the tanker fleet. A demolition decision has almost real-time effect on the market. Aspects regarding technical innovations, maritime regulations, market prospects and cost of classing influence scrapping decisions. Removals refer to vessels converted to another ship type, or vessels that for some other reason are taken out of the market.

Fleet Age

Average age determinants for aforementioned vessel sizes are included. Scrapping normally occurs when vessels are about 25 years (Clarksons, 2017a). Demolitions are more likely to increase when the age profile is becoming heavier right-tailed. An older fleet suggests shorter lead-time for capacity down-adjustments in terms of scrapping. In this sense, an increased age profile is hypothesised to be positive related to freight rates. Additionally, when the market improves owners tend to hold onto their ships. Furthermore, the fleet becomes younger when vessels are delivered. As we described earlier, deliveries are unfavourable to the market.

Empirical literature in regards to the fleet age in a macro context is far more limited than in a micro context. Köhn (2008) describes that the risk of hiring an aged vessel should require a discount in the freight rates. He indicates that a two-tier tanker market has emerged post-OPA90 (Oil Prevention Act of 1990). Although from the same paper, Kohn refers to Tamvakis (1995) who does not find any consistent difference in freight rates between younger, double-hull or single-hull vessels. Further, Tamvakis finds that weak market conditions are likely to disfavour freight rate discrimination. Strandenes (1999) also find that a two-tier tanker market only can be marginally seen in freight rates under some short-lasting period. We expect that these matters would not be of any particular significance in a macro environment.

Vessel Prices

Price dynamics have repeated themselves over the history. For instance, in 1995, Velonias (1995) exemplified with several cases that high ordering during boom periods leads to higher prices, both in the S&P and the newbuilding market. S&P prices and newbuilding prices have historically developed in parallel (Beenstock and Vergottis (1993); Tsolakis et al. (2003); Kavussanos and Visvikis (2016)). In contrary to boom periods, market collapse has the opposite effect. In the aftermath of the financial crisis in 2008, prices fell sharply across all tanker vessels (Clarksons, 2017a).

While newbuilding and S&P prices depend on both the activity in the freight market and steel commodity prices, scrap prices are more dependent on the latter. Scrap prices are namely quoted in dollars per lightweight tonnage (i.e. the amount of steel). Countries reliant on scrap metal import, such as Southeast Asian countries, imports only a small share as vessel scrap. Hence, demolition values are less influenced by the freight market (Kagkarakis et al., 2016). Further, safety and environmental concerns with beaching[31] vessels have triggered EU to impose new standards for world-wide dismantlement[32]. Several countries have ratified a convention saying that all EU flagged vessels are required to be dismantled in EU-approved facilities, thus increasing the number of qualified scrapyards (The Maritime Executive, 2017). Increased care of dismantlement will possibly cause demolition values to decrease, and consequently decisions of dismantlement will be less incentivised, which isolated is negative for freight rates.

Tsolakis et al. (2003) provides an empirical analysis of S&P prices in the dry bulk and tanker market. Their findings indicate that S&P prices of different segments and ship sizes react differently to changes in determining variables, and proposed that analyses should be carried out at a disaggregated rather than an aggregated level. This supports the inclusion of route-specific vessel prices (i.e. VLLC prices for TD1 and TD3). Furthermore, newbuilding prices and time charter rates are found to be leading predictors of S&P prices, both in the short and long run. They also provide results indicating that tanker orderbooks (as percentage of the fleet) has a negative effect on S&P prices in the long run. This is in line with the hypothetical orderbook lag effects presented earlier. Other econometric articles on S&P tanker prices found, e.g., the state of the freight market, deadweight tonnage, newbuilds, scrap prices, age, voyage rates, time charter rates and orderbook to have significance (Köhn (2008); Pruyn et al. (2011); Zhong and Shi (2007); Adland and

[31] Refers to vessels laid ashore for dismantlement; often under very poor and dangerous working conditions.

[32] EU Ship Recycling Regulation entered into force on Dec. 30 2013. Further to be followed by the IMO Hong Kong Convention of May 15th 2009, which will enter into force 24 months after countries representing 40% of world merchant shipping have acceded to the convention.

Koekebakker (2007)).

5.2 Demand-Driving Factors

Demand tonne-miles is also made up of two factors, namely cargo volume and the distance cargo travels (see Chapter 3). The demand for tanker services is heavily dependent upon international trade in crude oil and oil products. This is in turn derived from global and regional economic activity, as well as import and export of (other) energy commodities (Stopford, 2009).

In contrast to supply factors, demand factors are generally exogenously related to freight rates. Furthermore, as we saw in Section 3.2, demand is price inelastic to freight rates. Hence, correlograms of demand variables will not portray the same cyclical relation as supply variables. Interesting lag structures may, however, exist.

Oil Demand

Arguably the most important factor of derived shipping demand. Demand placed on a good or service is a result of changes in the demand or price of some other related good or service (Alderton and Rowlinson, 2013). Hence, demand for tankers is influenced by the demand for crude oil and oil products. If oil is not consumed, people would have few incentives to trade oil. Therefore, future trade dynamics are heavily influenced by this factor. We saw in Section 3.5 that 45% of the world's oil is consumed on the road. Topics regarding electrical vehicles' disruption on oil demand are popular these days. Oil demand itself is price inelastic to freight rates. We hypothesise demand to be positively related to freight rates.

Oil Import and Export

Oil import and export cover the tonnes part of the tonne-miles equation, and are direct measures of oil being traded in the world. Naturally, crude oil and dirty and clean products trade are expected to have a positive impact on freight rates of respective routes. *Seaborne* crude oil and products trade numbers, both import and export, are also added to respective routes where available. We note that these numbers are generally very similar[33] to total import and export, indicating that most of the oil trade is traded by sea – which support the relevance of "non-seaborne-specific" import and export numbers. The positive relationship between seaborne oil trade and tanker freight rates is widely covered in literature (see, for instance, Tsolakis (2005); Stopford (2009); and Anyanwu (2013)).

In Chapter 4, we outlined trade patterns. Here it became evident that the Middle Eastern Gulf (or Arabian Gulf (AG)) produces and exports a vast share of oil. This leads us to include regional import and exports variables as well. We figure it is positive that high regional activity levels, in proximity to our selected routes, cancel out available fleet capacity. EU4 serves as a benchmark for the need of European oil trade (TC1, TC2), and North Sea exports is deemed relevant (TD7), even though Sullom Voe gets its oil from the UK shelf. Furthermore, regarding TD1, TD12 and TC2, we are able to capture route-specifics almost on port level, with trades into PADD3 and PADD1[34]. While China is a major player in oil trade (largest crude importer), Chinese crude export and products import and export are not added because of poor data quality. Data are too likely to be ambiguous, according to footnotes from one of the data sources we use[35]. In general, Chinese industry numbers are not among the most credible numbers that market participants have available (Clarksons, 2017a). However, we find Chinese crude imports from SIN to be of good quality, as import numbers are more transparent. Further, India crude imports (3rd largest crude importer) and combined US, EU4 and Japan crude imports are added as general-specific variables to all routes.

The U.S. is not self-supplied with oil, which Figure 4.2 from Section 4.2 indicates. Imports of crude oil or oil products are necessary. Products are, however, not imported in high volumes, most seem to be refined by domestic refineries. Of products imported to the US, gasoline is the dominant product. The consumption of gasoline is higher than the consumption of diesel, partly since the sweat and light quality of the U.S. WTI oil (see section 3.5) does make refinement ideal for gasoline. The majority of the U.S. car fleet is gasoline-driven as a consequence. According to JODI (2017), they import about 0.5 million of barrels per day of gasoline. In contrast to Europe, where diesel became the main transport fuel after Europe backed for a switch from gasoline to diesel in the late 1990s, which doubled diesel imports between 2001 and 2014. In Japan, the consumption levels of distillates are fairly evenly distributed (Clarksons, 2017a).

Vessel Fixtures

Vessel fixture refers to the hiring of a vessel. We have included vessel fixtures for all relevant vessel sizes on both a total and a route-specific level. One of the fixture series we use has been constructed to "suit" the specific route - Aframax fixtures related to TD7 is composed of Aframax AG-Cont, Baltic-Cont., Baltic-Med., Med-Cont. and FarEast-Cont. fixtures. Fixtures for

[33] Caution must be taken regards to multicollinearity (addressed in the Methodology chapter).

[34] EIA provides excellent data on US oil trade

[35] Joint Organisations Data Initiative (JODI) (see Chapter 6).

other routes are available quite precisely. For instance, VLCC AG-West and VLCC AG-Japan fixtures are added to TD1 and TD3, respectively. Moreover, fixtures on routes hypothesised to affect our selection of routes are also added. For instance, MR Med-USA fixtures is added to TC2. Rising fixture numbers are hypothesised positive to freight rates.

We do also add a factor describing VLCCs that are *due in the MEG* in the month in question. These VLCCs arrives the Gulf, e.g., when a round voyage is performed, and hence increasing the available supply capacity. A negative relationship is therefore hypothesised. We find these factors only to be applicable to TD1 and TD3. Unfortunately, the empirical literature on vessel fixtures on a macroeconomic level seems to be limited, and we have not come across any literature that include these factors in forecasting. Tamvakis (1995) finds, however, that a two-tier market in terms of vessel age can be described on the basis of whether fixtures are US-bound or non-US-bound.

Crude Oil Production

Oil produced is an explicit indicator of tanker demand. Zacharioudakis and Lyridis (2011) explore tanker market elasticity with respect to oil production, and shows that OPEC oil production is the crucial non-supply external variable affecting the tanker market. He argues that this variable is able to embody political, economic and direct and indirect psychological matters to the shipping market, including wars, OPEC decisions on production and production productivity. Increasing oil production is hypothesised to be positively related to tanker rates. However, some caution must be taken regards to this hypothesis. In Figure 4.2, we saw how production related to demand. Hence, increased production may actually offset the need of imports in countries short of oil. We therefore hypothesise a negative relationship for U.S. oil production on the TD1 route only.

OPEC and global oil production are added as general-specific variables to all routes. Certain production numbers are also added as general-specific variables to subsets of routes. For instance, US oil production is relevant for TD1, TD12 and TC2.

Refinery Output

Refinery output refers to the output of finished products only (JODI, 2017). We have only included refinery outputs for products routes, and we have been precise by adding clean and dirty refinery outputs to respective routes (i.e. dirty on TD12 and clean on TC1 and TC2). Refinery output could be a better forecast indicator than import and export, since its lag structure is more likely to contain information that could describe future trade flows. After all, import and export describe cargoes that are already transported. However, refinery output is somewhat complicated to address. We use, among other things, Figure 4.2 in Chapter 4 to understand whether output is likely to cover domestic or foreign demand. TC1 is quite straightforward, as Saudi Arabia is not in the need of imports. TD12 and TC2 depend on the aforementioned; if output is covering domestic demand, less import is needed and/or less is exported. For TC1, increased Japanese output would yield less need of imports. Nevertheless, we hypothesise a positive relationship with freight rates (except Japanese output) - more oil is generally a positive sign.

For crude routes, we have not included refinery output, partly since we do not have the specifics of what refinery input is domestic crude production and what is crude imports. Isolated, increased U.S. production could yield higher refinery outputs, which will not be positive for TD1. For TD3, increased Japanese refinery output would be positive, since they do not produce any oil. Moreover, refinery of crude is taking place after it has been imported.

Refinery Utilisation

Refinery utilisation refers to refineries' output level as a percentage of its capacity level. EIA provides time series on district-specific level. We are therefore including PADD1 and PADD3 refinery utilisation levels to TC2 and TD12, respectively. Trade is not necessarily unidirectional, as described in Section 4.2. Our hypothesis is therefore neutral for these variables.

The impact of increasing capacity and utilisation on trade flows varies by region. On the one hand, crude trade may offset some products import demand. On the other, increased capacity may also support exports. For instance, Latin America is short on light products, and its sanctioned refining projects do not add up to this shortage, which in turn supports products export from the US Gulf (McKinsey, 2015). High capacity additions may further put pressure on utilisation if demand growth slows down.

Europe's sanctioning of refinery capacity is slowing down, while we see built ups in China, Middle East and the U.S. When refinery capacity change, trade patterns also change. Lower refining capacities in Europe will lead to the need of more long-haul products routes (LR2 vessels) across, e.g., the Atlantic Basin towards Europe. In 2017, the Chinese government tightened scrutiny over taxes and shifted quota policies, causing refinery margins to tighten, and also cut off export quotas of products, trapping products output for domestic demand. So called non-state-driven "teapot refineries" produce approximately 12% of Chinese crude demand. Seeing squeezed refinery margins, these refineries get incentives to shut down, and thus cap Chinese crude imports. In the unsureness of import quota adjustments, refineries built high inventories, which have translated negatively to tanker shipping (McKinsey, 2015). However, in 2018 China has relaxed its policies towards these independent refiners.

5.3 Economic and Non-fundamental Factors

Economic and imponderable factors such as external influences (e.g. market prices), financial markets, boundaries (e.g. politics and weather) and behaviour and psychology (see Section 3.4) may have significance in addition to fundamental supply and demand factors. We will try to capture the influence of such factors through 10 additional groups of Economic and Non-fundamental variables. As with supply and demand factors, we will rationalize the hypothetical significance and empirical studies that lies behind our selection.

Gross Domestic Product (GDP)

World GDP is a measure of the development of the world economy. Stopford (2009) describes GDP as the single most influential factor on shipping demand. He compares the world business cycle[36] with shipping business cycles, and find them to coincide. In short, he argues that investments and consumption trigger trade and vice a versa. Klovland (2002) examines business cycles, commodity prices and shipping freight rates for periods pre-WWI, and finds a close timing relationship between the *upper* turning points of these three factors. On the other hand, trough periods were generally less synchronized with business cycle troughs. Along with empirical literature and economic theory, Klovland's findings support our assumption of asymmetry between the peaks and troughs of shipping cycles. Abouarghoub et al. (2012) investigates the tanker market between 1994 and 2010, and finds the correlation between GDP growth and oil seaborne trade growth to be significant. During the most recent economic boom period, GDP increased 20%, while total seaborne trade grew by 21.5%. After the financial crisis, both GDP and oil seaborne trade fell. As expected, Abouarghoub et al. (2012) concludes that continuous changes in demand for oil seaborne trade have profound effect on tanker earnings.

In the relation to seaborne trade, GDP acts as a measure of the successfulness of global interaction - the way regions collaborate and generate global GDP. The world is becoming increasingly globalised, and trade and flow of cargo is thus essential in global value generation. Furthermore, the world is growing, both in complexity and population (energy demanding), and cities are becoming larger and need access to international trade by sea. The constant change and increased complexity of the pattern of world trade will continue, as developing countries will account for above 90% of population growth from today to 2030 (Kavussanos and Visvikis, 2016). Kavussanos and Visvikis (2016) do, however, argue that attempts from economists to correlate GDP growth with seaborne trade growth are not particularly successful.

In terms of GDP, UNCTAD (2016) do help us further motivate why a regime model is suitable. They find long-term trade-GDP elasticity to vary across different historical time periods, in which estimates suggests that cyclical factors to trade slowdown is more noticeable during crisis and recession periods.

World GDP is the only GDP selected. GDP on country-level is rarely accessible on higher frequency data, and is nevertheless highly correlated with industrial production which will be introduced later.

Time Charter Rate

1-year time-charter rates have been included for ship sizes relevant for each route. Shipowners' and charterers long-term expectations about the future is reflected in term as well as FFA rates (see Section 3.1). Shipowners closes on term time-charter contracts when they do not want to bear the risk of varying spot rates. The time-varying risk premium of spot contracts thereby suggests a higher spot rate today than future term rates. Tsolakis (2005) identify spot rates as the major determinant for period rates, both in the dry bulk and tanker market, thus confirms validity of the expectations theory[37] of the term structure relationship in shipping freight rates (see, also, Alizadeh and Nomikos (2011)). However, he finds exceptions in the Aframax and Panamax tankers models, whereby fleet changes were found to be statistically significant instead.

A positive change in time-charter rates means that the market is expecting spot rates to improve (Köhn, 2008). Stopford (2009) does somewhat challenge this hypothesis, by arguing that shipping demand may be affected by higher transportation costs, which is the case when freight rates increase. We find this unlikely, and consider demand very inelastic to freight rates (see section 3.2). For instance, a VLCC transporting 2 million barrels of oil for 25,000 \$/day on a 43 days TD3 round voyage (see Table 6.2) would require a hire of 0.009% of the cargo value, if assumed an oil price of 60 \$/barrel.

[36]Business cycle: periodic fluctuations in the rate of economic growth.

[37]Expectations theory: the hypothesis that long-term rates contain a prediction of future short-term rates (Investopedia, 2018a).

Exchange Rate

Exchange rate refers to the value of a nation's currency in terms of another currency. We include exchange rates for currencies against the US dollar for respective routes. Additionally, we have included a dollar index and a euro index to track shipowners' revenues and the European economy, respectively. A trade has two sides; money is either received or paid. Japanese as importers, for instance, would benefit from a strong Japanese economy in relation to Saudi Arabian or US economy. Cullinane et al. (2005) refer to Tvedt (2003) who claims that the Japanese economy is a major driver of dry bulk shipping. We expect this to hold for tanker shipping as well.

Referring to Section 6.2, we identified that certain voyage costs are converted from local currency to US dollars. Freight rate income is normally received in US dollars (see Section 3.1). A strong dollar would thus be positive in terms of income and voyage cost. Hence, both cost and revenues for shipping participants are concerned with exchange rates fluctuations. Furthermore, shipyards in, e.g., the Far-East normally receive payment based on dollar quotes. A strong dollar could thus lead to an increase in newbuilding orders. Overall, exchange rate fluctuations is hypothesised neutral.

Consumer Price Index and Money Supply

A consumer price index (CPI) is a measure of prices of consumer goods and services. The percentage change in CPI is often used as a measure of inflation in an economy. We motivate this inclusion by referring to Ringheim and Stenslet (2017), who find that US CPI is statistically significant when predicting the BDTI index, but has a negative impact as opposed to their initial hypothesis. A possible explanation is that a weak CPI reflects declining prices, which raise consumer power, and consequently demand for traded oil products. Furthermore, we have included M1 money supply. M1 refers to funds that are readily accessible for spending (FRED, 2017). We therefore anticipate this to have a positive influence on freight rates, as purchasing power to buy oil products increase with increasing spending funds. Japanese, US and European CPI are included for respective routes. Japanese and US money supply are included for TD3 and TC1, and TD1, TD12 and TC2, respectively.

Interest Rate

Our intuition is that interest rates have a negative relationship to freight rates. Our rational is that a rise in interest rate will increase the cost of capital and as a result lower people's purchasing power and the liquidity of most shipowners, and oppositely a decline in interest rates will result in higher investment willingness. Most shipping loans are financed based on a LIBOR base rate. We include three different 3-month LIBOR interest rates, based on Euro, Yen and US Dollar. The world of shipping is intercontinental. Hence, our motivation lies in the possibility of these variables to capture somewhat distinct dynamics in the freight rates.

Zhong and Shi (2007) found interest rates to have a negative effect on VLCC S&P prices. Tsolakis (2005) argued in his excellent PhD thesis, investigating the four shipping markets as well as ship finance of bulk shipping, that shipping demand is negatively affected by the LIBOR rate.

Industrial Production

Oil trade should be somewhat proportional to the industrial production of each country (Velonias, 1995). Above-mentioned factors, such as oil demand and interest rates, do also affect a nation's industrial production. According to EIA (2018), relationships between economic growth, oil prices and oil consumption are determined by structural conditions in each country's economy. Hence, we expect the impact of industrial production on freight rates to vary depending on the country in question. Japan, for instance, being heavily dependent on oil imports, are dependent on oil trade in all oil related industries. The industrial production numbers in Europe and the US are probably less sensitive to oil trade. The industrial sectors in Japan accounted for 30% of total oil consumption in 2013 (EIA, 2017a).

Industrial production numbers, on a year-on-year basis, are included for countries relevant to our selection of routes. Further, OECD industrial production is added as a proxy of oil consumption in large economies. We have also chosen to include industrial production numbers for India and China, being major players in the world economy as well as oil trade. Besides, India and China are not captured in the OECD factor. In fact, in the first decade of this century, non-OECD oil consumption rose 40%, whereby oil consumption in the OECD fell (EIA, 2017d). In the period from 2010 until today, OECD oil consumption growth has been rather flat overall, but increased somewhat between 2015 and 2017.

Crude Oil and Oil Products Price

The impact of oil prices on tanker freight and the economy as a whole has been extensively researched, but the relationship is not easily quantifiable. Hypothesis differ between a positively, negatively and twofold effect (see, for instance, Stopford (2009); Poulakidas and Joutz (2009); and UNCTAD (2016)). We support the latter. Oil prices' effect on rates must be seen in the context of the underlying drivers of the oil price, namely oil supply (production) and demand (consumption). The oil price decreases because supply increases and/or demand drops. In times when demand keeps up with supply growth, the freight market will benefit. When supply grows, and demand cannot keep up, oil prices decline, and the freight market benefits from higher oil flow. Low oil prices then initiate demand for oil, and oil prices rise. Increasing demand and oil prices, in turn, give incentives to increased supply. No wonder why people have different views on a unidirectional impact on freight rates; it is simply not unidirectional. Clarksons Platou summarise it nicely, "*Oil price dynamics have a mixture of positive and negative effects for shipping, but certainly remain crucial given the key role of oil both for shipping and for the wider economy*" (Clarksons Research Services Limited, 2017). This further motivates why a regime model may be appropriate.

Poulakidas and Joutz (2009) analyse, using cointegration techniques and Granger causality[38], the relationship between spot rates on the TD4 (WAF – US Gulf) route and the oil market from 1998-2006. They find a feedback between spot rates and WTI and crude stockbuilding. They conclude that rising oil prices put an upward pressure on spot rates. Hence, the oil price might serve as a proxy for oil demand.

We have in this study decided to focus on three different crude oils, namely the North Sea *Brent*, the United Arab Emirates *Dubai* and the U.S. *West Texas Intermediate* (WTI). The Brent[39] is a natural choice, since it is the most common benchmark crude oil. The reason for this is twofold; it is easily refined (light and sweet, see Section 3.5) and transportable. The North Sea oil is waterborne cargo, which make it ideal for seaborne trade. Brent could, for instance, be either loaded onto FPSOs[40], directly onto oil tankers or pumped to oil terminals like Sullom Voe (see TD7, Section 4.2). According to ICE futures, Brent is the source of pricing of 60% of the world's traded oil (ICE, 2013)[41]. The WTI is ranking next below Brent as a common benchmark. In contrast to Brent, WTI transportation is onerous in terms of shipping, since it is extracted onshore and transported by pipelines and rail. However, WTI is both lighter and sweater than the Brent blend. The quality of the Dubai falls in between Saudi Arabia - Light and -Heavy in both sulphur content and density (which are both loaded at Ras Tanura (TD1, TD3, TC1), and is also a more common benchmark oil than Saudi oil, and thus seem as an appropriate choice. It becomes evident that dynamics in regards to oil production and transportation of different crude oils further complicates hypothesising the directional impact of oil price movements on tanker rates. Moreover, differences in oil prices may generate arbitrage trade opportunities. Furthermore, these three benchmark oils also directly relate to the trading routes we are focusing on. However, we do not expect to see big differences when testing variables, considering the high correlation between the price of these oils.

In the *Oil import and Export* category above, we suggested that gasoline is the main fuel source in US road transportation. Gasoline prices in the US has a high elasticity to oil prices[42]. In July 2017, Reuters reported that low domestic gasoline prices at the start of the US summer driving season had encouraged diversion of tanker shipments from Europe, due to smaller profit margins, and incentivised domestic refineries to export (Reuters, 2017). This leads us to additionally include the US gasoline price (US Gulf Spot) on the TC2 route.

Crude Oil Forward Price. We have constructed a crude oil forward curve from Brent 6-month and 1-month delivery prices (6m – 1m). The rationale for this inclusion is to capture the contango[43] structure in the oil market. During contango periods, oil suppliers and oil traders have incentives to store, primarily on VLCCs, rather than trade crude at once. This exemplifies that it is not only the supply side that is psychological and speculative driven in the tanker market. The effect of increased storage due to a contango market is stronger when oil prices are relatively low compared to historical levels or when the price has dropped fast (Fearnley, 2018). Furthermore, contango in oil pricing favours long haul routes where the largest ships are employed, such as TD1 and TD3 – i.e. to arrive at offloading latest possible.

Bunker Price

Fuel, or bunkering cost, accounts for about 75% of total voyex costs[44]. Tsolakis (2005), in his investigation of the freight market, finds that bunker prices are significant in the long run for large tankers only (VLCC and Suezmax). He attributes this to the fact that larger vessels trade on long-haul routes, and thus earnings

[38] The Granger causality: statistical hypothesis test for determining whether a variable is useful in forecasting another.

[39] Brent is extracted from the BFOE oil fields in the North Sea (Brent, Forties, Oseberg and Ekofisk).

[40] Floating Production Storage & Offloading units

[41] Great article that covers the essentials of crude oil and oil products trade.

[42] In comparison to countries in Europe, like, e.g., Norway, where government regulations make gasoline and diesel prices far more inelastic to crude oil prices.

[43] Contango refers to future oil prices being higher than current prices, thus indicating a positive sentiment in the market. Backwardation refers to the opposite effect.

[44] Of total costs (capex, opex and voyex), bunkering accounts for 40-50%. Capex and opex accounts for about 25% and 20%, respectively (Kavussanos and Visvikis, 2016).

of these vessels are more exposed to fluctuations in voyage costs. We have chosen to include two bunker variables, namely 380 cSt Japan and 380 cSt Philadelphia. These were found to vary the most in relation to crude prices although they do move in highly in parallel with each other. It is further highly likely that vessels on some of our selected routes bunker at ports in Japan and Philadelphia. Furthermore, bunker costs used in Clarksons calculation of TCE rates are based on 380 cSt fuel (see Section 6.2). Our hypothesis is, however, neutral, in line with our oil price hypothesis above.

Shipping Index

The ClarkSea shipping index is included to capture general freight market sentiment and possible cointegration across shipping segments (see, for instance, (Talley, 2011)). The index is an indicator of earnings for all the main commercial vessel types, and is weighted according to the number of vessels in each segment (Clarksons Research Services Limited, 2017). We further add three tanker indices; BDTI, BCTI and Clarksons Average Tanker Earnings Index. We are including these instead of lagged observations of the dependent variables. Motivated by theory from introductory chapters, our hypothesis is that freight rates are driven by cyclical momentum. Hence, we expect relationships to be positive.

Stock Index

Lastly, our selection of variables concludes with a set of stock indices. These are S&P500 (US), MSCI World, MSCI Emerging, Tadawul (Saudi Arabia) and Nikkei225 (Japan). S&P500 measures the value of the 500 largest companies listed on the New York Stock Exchange. MSCI World and MSCI Emerging are designed to measure equity market performance for large and mid-cap companies across 23 developed markets and 24 emerging markets, respectively. Tadawul is the Saudi stock exchange, and we link it to TD1, TD3 and TC1. Nikkei is Japan's leading index of Japanese stocks, and we link it to TD3 and TC1. S&P500 and MSCI indices are tested across all routes. A positive relationship is hypothesised, as improving market conditions lead to higher trade activity.

Shipping markets exhibit, as we know from Chapter 3 periods of extreme volatility. In the hope of capturing psychological effects, we have included the VIX index (motivated by Section 3.4). This index tracks the market's volatility expectations of the S&P500[45]. Recently, the VIX index has been very calm (see Figure D.3 in the Appendix). From Section 3.2, we identified that freight rates are more volatile during peak than trough periods. Hence, we expect the VIX to have a positive impact on freight rates.

[45]Implied volatilities of S&P500 options.

5.4 Additional and Non-capturable Factors

Supply. In general, shipowners are concerned with four core factors when employing a ship on a voyage. These are: distance, ship size, ship type and ship speed (Stopford, 2009). These factors are all determining factors of freight rates. We are able to compare freight rates in terms of *ship size* (VLCC, Aframax etc.) and *ship type* (crude or products), and also *distance* (route), which are constant factors, but we do not have a long enough time frame on time series for *ship speeds* to include these in the models. In shipping, slow steaming is used as a way of managing excess capacity. Besides speed, other fleet productivity data are not included. Furthermore, referring to Clarksons' TCE calculation in Section 6.2, we do not have time series' for port costs. However, these costs are minor in comparison to, e.g., bunkering cost, as described earlier. Missing port costs could, nevertheless, be captured somewhat by exchange rate variables. Moreover, congestions and disruptions in relation to ports, canals and chokepoints are not captured by factors presented above. Additionally, weather factors would have been useful to collect, but we argue that these effects may be captured by incorporating seasonality in our models (see Chapter 7).

Demand. Initially, we thought of including export variables for non-route-specific countries that a vessel may ballast to and load new cargo at before eventually (or potentially) returning to its port of origin. This could indirectly capture capacity that cancels out. For instance, exports from West Africa (WAF) could be especially relevant to TD1. However, for several reasons, we concluded not to add them after all. Furthermore, market participants do also follow refinery margins. In general, we hypothesise that higher refinery margins will translate into higher attractiveness of crude import (refinery demand), thus being positive for crude tanker shipping. For products trade, the relationship is more complex. As for vessel speed, we were only available to gather such data from 2008. Lastly, oil inventory and strategic petroleum reserve (SPR) factors could help in the insight of whether oil has been exported/imported or been added/drawn from inventory.

E&N. Ringheim and Stenslet (2017), forecasting the BDTI and BDI indices, include a high yield bond spread, which is motivated to reflect investor's willingness to invest money. However, the variable was not implemented in the best models, as other variables where found to have better predictive power. Lyridis et al. (2017) and Velonias (1995) identify, among other factors, political decisions, war and climate conditions as influencing factors on tanker demand. Politics is, however, not easily captured by time series. Knowledge of politics does make it possible to identify factors that may be of high importance in the future, such as the rescinding of U.S. export bans in 2015, which leads us to identify US exports as an important factor going forward (see Chapter 9). Further, costs of sulphur emission regulations are included in Worldscale rates, thus it could be argued that this political enforcement is somewhat captured. Lastly, geopolitical conflicts can potentially disrupt and limit tanker trade significantly. Conflicts in oil producing countries can cause cut-offs in oil production and distribution channels.

Table 5.1: Hypotheses overview of the directional impact of determinants on freight rates. Hypotheses are based on the impact a *positive* change in the determinant has on the freight rate on real-time lag (0-lag). Exceptions made to each determinant group are given at the end of the table. See Chapter 5 for a description of each of the determinant groups. A complete overview of specific determinants included for each group is given in Appendix A.1.

Determinant Group	**Hypothetical impact**
Dependent Variable	
TD1	TCE Ras Tanura − LOOP; VLCC route
TD3	TCE Ras Tanura − Chiba; VLCC route
TD7	TCE Sullom Voe − Wilhelmshaven; Aframax route
TD12	TCE Antwerp − Houston; LR1 route
TC1	TCE Ras Tanura − Chiba; LR2 route
TC2	TCE Rotterdam − New York; MR route
Panel A: Supply	
Fleet Size, Mdwt & Yr/Yr growth	(−)
Fleet Size, Newbuilding contracting	(+)
Fleet Size, Orderbook Mdwt & Perc. of fleet	(+)
Fleet Size, Deliveries	(−)
Fleet Size, Demolition & Removals	(+)
Fleet Age	(+)
Vessel Prices	(+)
Panel B: Demand	
Oil Demand	(+)
Oil Import and Export	(+)
Vessel Fixtures	(+)
Crude Oil Production	(+)
Refinery Output	(+)
Refinery Utilisation	(+)
Panel C: Economic & Non-fundamental	
Gross Domestic Product	(+)
Time Charter Rate	(+)
Exchange Rate	(*neutral*)
Consumer Price Index & Money Supply	(−) & (+)
Interest Rate	(−)
Industrial Production	(+)
Crude Oil & Oil Products Price	(*neutral*)
Bunker Price	(*neutral*)
Shipping Index	(+)
Stock Index	(+)

Comment on hypothetical exceptions:
Panel A: Supply:
*non exceptions made
Panel B: Demand:
* Vessel Fixtures, VLCC Due (TD1, TD3): (−)
* Crude Oil Production, US Production (TD1): (−)
Panel C: Economic and Non-fundamental:
* Oil Products Price, US Gasoline Spot Price (TC2): (−)

6 Data Analysis

In this study, we model the logarithmic return distribution of 6 different oil tanker freight rates. The rest of this chapter is structured as follows. First, we examine stylized facts about shipping freight rates from a theoretical point of view, which are of relevance as supporting theory and motivation to the modelling methodology and results in the forthcoming chapters. We briefly examine three characteristic properties of freight rates; *mean reversion*, *seasonality*, and *distribution and jumps*. Second, the background story of how Time Charter Equivalent (TCE) freight rates are calculated and quoted by Clarksons Research Services Limited is outlined. Third, we present some brief comments on data gathering and sources. Next, pre-processing of some of the data gathered is discussed along with a presentation of the dependent times series. Finally, descriptive statistics of the selected freight rates are given.

6.1 Stylized Facts of Freight Rates

Mean reversion

Stationarity, and the mean-reverting property that follows for shipping freight rates, is a significant property of short-term cycle behaviour according to maritime literature (Koekebakker et al., 2006). It is generally stated that cyclical behaviour of a time series is a strong indication of mean-reverting properties in the data. Mean-reversion theory suggests that supply-demand dynamics in the market eventually will force the freight rate back to its mean. As described in Section 3.2, there is a continuous regulation of supply in the shipping markets, which is causing freight rate fluctuations around the long-term mean.

Stopford (2009) identifies that shipping markets operate under conditions of nearly perfect competition, and are as such usually held as textbook examples (see Section 3.5). In these markets, mean-reversion is usually apparent. Koekebakker et al. (2006) identifies that freight rates are unsustainable at extreme highs and lows, due to the potential for supply adjustments. At high freight rates, scrapping of vessels slows down, and delivery and order book placements continues unabated. As the tonne-miles capacity in the market increases, the supply curve will gradually shift to the right, and freight rates will revert down to levels more close to or below the mean. In the times of low freight rates, downward regulation of supply will yield opposite effect. Moreover, in a perfectly competitive market, maritime economists argue that freight rates cannot exhibit asymptotically explosive behaviour, which implies rejection of non-stationarity.

The rate at which the freight rate returns to its mean could be distinctive for which state the supply-demand curve is at, either the elastic or inelastic part of the supply. Adland and Cullinane (2006) shows that freight rates in the tanker market only reverts to its mean in the extremes of the distribution (trough and peak), and else is exhibiting a non-stationary process over most of its empirical range. This kind of non-linear mean-reverting behaviour backs our motivation of implementing regimes in modelling. Regime switching methodology is set out in detail in Chapter 7.

Several empirical research articles in the maritime literature have explored the stationary property of the freight rates. Koekebakker et al. (2006) argues that maritime economic theory suggests non-linear stationary dynamics, and are using various methods to test this themselves for the dry bulk and tanker markets. Their conclusion, after performing a non-linear version of the augmented Dickey-Fuller (ADF) test, is that freight rates, in line with maritime theory, exhibit a non-linear stationary behaviour. Koekebakker et al. (2006) identifies that a substantial body of empirical research suggest non-stationary behaviour of freight rates, and further identifies that this conclusion is mostly drawn when traditional linear unit root tests are used. Using such tests may be a pitfall, since they are known to have low power against relevant non-linear alternatives. Non-linear alternatives may be better when dealing with a highly persistent price process, which is the case for freight rates in the short-run. The persistent behaviour is partly a consequence of the continuous slow supply regulation against the demand due to lead times between ordering and delivering of vessels (Section 3.2), which indicates that it usually takes time for the cycle to move into a new cycle stage. Oppositely, demolition of vessels may take time as shipowners hesitate to make these decisions. Non-stationarity may therefore be too easily captured when using linear unit root tests if the time-horizon for the data set is short.

Further discussion on stationarity and its implications on modelling is given in Chapter 7.

Seasonality

From Section 3.3 it became evident that tanker rates exhibit seasonal variation, both from a theoretical perspective and from a visual representation of the seasonal cycle. This kind of behaviour is broadly documented in maritime literature across all market segments (Alizadeh and Nomikos (2009); Kavussanos and Alizadeh (2002)). Seasonal cycles and events in the underlying market, such as seasonal variation in oil demand, oil production, refinery outputs and heating/cooling demand causes demand for tanker services to vary. Furthermore, we identified weather conditions as a significant bottle-neck on certain routes. In periods, weather can limit the capability and availability of a portion of the fleet, leaving the rest of the fleet with potential lucrative contracts. The influence of such seasonal effects vary across different market segments and routes. In Chapter 7, we will address how we address and incorporate seasonal effects in our prediction models.

As for the mean-reversion property, seasonal changes in freight rates exhibit asymmetric properties between different supply-demand states. In recession periods (collapse and trough), seasonal changes are less evident in comparison to expansion periods (recovery and peak), due to the supply curve being elastic and inelastic during recession and expansion stages, respectively. This is more evident in time series with monthly or higher frequency (Alizadeh and Nomikos, 2009). As such, distinct seasonal patterns and effects across different regimes may occur.

Distribution and Jumps

Demand could be subject to short-term rapid change, thus causing sudden jumps as the supply-demand equilibrium changes. Jumps tend to be of greater magnitude and of higher frequency when the equilibrium is closer to the inelastic part of the supply curve, i.e. in expansion periods. The combination of excess demand and short-term constrained supply will cause strong freight rate movements. Oil companies have high alternative costs, so they would usually rather pay a premium if there is shortage of ships. In more troublesome times, with excess supply, a rapid change in demand would not have the same strong effect. Higher volatility, and consequently more frequent jumps, at higher freight rate levels has historically been the case (Fearnley, 2018). Hence, the distribution of the freight rate is expected to feature non-zero and higher levels of skewness and kurtosis, respectively (Abouarghoub et al., 2014).

Maritime literature provides strong evidence of clusters in rates' return distribution, whereof large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes (Abouarghoub et al., 2014). This means that a jump tend to be followed by a new jump or be persistent at the current level. Climate conditions affecting supply is for instance a reason for sudden jumps in the freight rates. Weather was also mentioned in the context of seasonality, indicating that jumps may occur more frequently in different seasons.

Shipowners' awareness of the significant lead time from placing an order to delivery of a ship can occasionally cause shipowners to make early investments. Shipowners trying to predict freight rates could therefore trigger a decrease of the likeliness for jumps, since delivery of early investment ships contribute in meeting potential excess demand.

In summary, considering shipping theory and empirical literature, we set out to prove or disprove the following regards to each *fact*:

i) *Mean reversion*: Freight rates to revert to its mean faster in high volatility regimes, and hence these periods to have a shorter duration than low volatility periods.

ii) *Seasonality*: Freight rates to have more evident seasonal effects in high volatility regimes.

iii) *Distribution and Jumps*: Freight rate changes to be of higher magnitude and extremes to occur at higher frequency in high volatility regimes, also referred to as level effects in the conditional variance.

6.2 TCE Earnings Calculation

We continue this chapter by giving a more comprehensive explanation of the time series that are to be forecasted, the dependent variables. Such knowledge is deemed important for sound judgement in the selection and interpretation of hypothetically significant determinants, which was covered in the previous chapter (Chapter 5).

The route-specific TCE rates, or voyage earnings, are gathered from the Shipping Intelligence Network (SIN) database of Clarksons Research Services Limited (2017). In the document *"Sources & Methods for the Shipping Intelligence Weekly"*, CRSL gives detailed descriptions of the freight rate data in SIN. We will now outline the CRSL's assumptions made for the TCE calculations. The complete structure of the calculations are presented in more detail in Appendix A.7.

CRSL states that the purpose of the TCE time series is to provide an *estimate*[46] of the daily earnings of ships which is implied by the current level of Worldscale rates. These estimates are calculated on the basis of characteristics of a set of reference vessels trading on various routes, which CRSL have consulted Clarksons Platou brokers in order to determine. Furthermore, CRSL provides TCE data for both a modern and an older reference vessel for each route. We note that the reference vessels used for calculations have changed over the time series period. At the most recent date of our chosen time horizon, time series are based on *c.2000 built* vessels for crude routes and *c.2010 built* vessels for products routes. Either way, rates based on

[46] In fact, Baltic Exchange indices are also estimates, and are calculated in a similar manner as TCE rates modelled in this study. In general, awareness of potential biasness from using data made up by assumptions is important. CRSL: "The use of different assumptions can potentially make a significant difference to the results of earnings calculations.".

different reference vessels are almost fully correlated on the same route despite vessel age.

Calculations of earnings are in line with TCE theory presented in section 3.1, i.e. net revenue earned on the voyage, after deducting bunker costs, port and canal costs as well as commissions, and finally dividing by the days for the voyage. Certain cost factors are not accounted for, such as: i) waiting time at port, ii) off-hire time, iii) other voyex.

Freight rates are converted from Worldscale using flat rates for the respective routes. *Cargo volume* loaded is predetermined, and intended to represent the most common volume currently prevalent on the route in question; considering restrictions concerning vessel capacity, ports (draft) and canals (draft, beam). *Bunker cost* is a function of the days for the voyage, with consumption per day at sea and in port[47] and price of bunker[48] as parameters. A representative bunker port(s), and thus a corresponding regional bunker price, have been selected for each route. *Port cost* is the sum of loading and discharge costs based on cargo size and ship type. *Canal cost* is only added when a voyage includes passage through a canal. Important to notice, is that costs are converted from local currency to US $ at the current exchange rate. The *duration* of the voyage in days is determined on the basis of *sea time* and *port time* (and potentially canal transit). Sea time is calculated from the voyage distance and predetermined speeds for laden and ballast, plus an additionally sea margin of 5%. As indicated earlier in this chapter, for some tanker routes, a shorter ballast voyage may occur (e.g. for TD1), and is thus accounted for in calculations when considered realistic. Regarding port time, an average of two days both at loading and discharging is used for tanker vessels. Further, earnings are not attempted optimised by any means, e.g., by adjusting speed or fuel type.

In table 6.1 and 6.2, we have listed the reference vessels CRSL use in calculations for our selection of routes, as well as assumptions made in terms of route-specific factors mentioned above. As can be seen, the same reference vessel can be used for several routes (here TD1 and TD3), while voyage details differ. TD1 is the only route that is assumed not ballasting back to its port of origin, as the shorter ballast distance portrays. In the tables, we have only included vessel and voyage details used in the most recent part of the time series', i.e. January 2009 onwards. Prior to January 2009, some deviation occur.

The fact that most TCE rates are calculated based on vessels trading on a round voyage, excludes the favourability that shipowners may have by achieving higher utilisation for their vessels, i.e. ballasting less than 50% of the time at sea. Triangulation is a well-known strategy for maximizing utilisation, whereof especially smaller vessels load and discharge cargo at several ports before returning to the port of origin. TCE rates could thus be downwards biased compared to true spot earnings (see, for instance, Adland (2003)). Furthermore, CRSL provides TCE data on a specific route that is constructed on the basis of the arithmetic mean of the flat rates (see Section 3.1). Accordingly, these TCE rates do not directly reveal that vessel-specific earnings may deviate significantly from the reference vessel used in the calculations, due to distinctive fuel consumption and cargo capacity properties of each individual vessel. *Throughout this thesis, we therefore consider TCE rates that portray the average state of the market on each route for a certain reference vessel.*

6.3 Data Gathering and Sources

We mentioned in the previous chapter that we have chosen variables quite specifically to capture characteristics of each route. For instance, Aframax and LR2 vessel data are used on crude oil and oil products routes, respectively. When, e.g., LR2 data were not available, Aframax data were used as a substitute (*and so on*). Furthermore, data concerning oil types (e.g., import & export, demand, and refinery output) have been collected on crude oil, clean products and dirty products level according to the relevance for respective routes. Additionally, we mentioned general variables. These are data that are not unique to any specific route, but can be used on several routes or all routes.

In total, we have gathered 169 time series', whereof 6 dependent variables and 163 independent variables. Demand, Supply and E&N groups consist of 68, 51 and 46 variables, respectively. On route level, TD1, TD3, TD7, TD12, TC1 and TC2 consist of 31, 29, 23, 34, 28, 35 variables, respectively. There are also 29 additional variables that are common to all routes (see, also, Appendix A.2). We refer to Table 5.1 from the previous chapter for an overview of the hypothetical impact of each variable group. For a complete list of all variables included in the data selection, and details regarding data source, data description, unit and more, see Appendix A.1. Moreover, a complete list and description of data sources used is given in Appendix A.3.

Data are gathered on a *monthly frequency.* However, note that data variables vary in their format; whether they are quoted based on start of month (SoM), end of month (EoM), average, year on year, or sum. Above-mentioned appendices give a complete overview. Data that were only available on a higher frequency than monthly, were collected based on EoM quotes. Ad-

[47] Dirty products voyage calculations include allowance for heating consumption at sea and in port. In our study, this matter is relevant for the TD12 route.

[48] Worldscale.co.uk: "Annual bunker prices and consumption rates for standard 380 cSt fuel allowances are set out in the Preamble for each edition of Worldscale."

Table 6.1: Tanker reference vessels for earnings (TCE) calculations on selected routes (Clarksons Research Services Limited, 2017).

Route Reference vessel	Dwt	**Cons. at sea** [tonnes/day] Laden	Ballast	**Cons. in port** [tonnes (gross)]
TD1: Ras Tanura - LOOP and **TD3: Ras Tanura - Chiba** VLCC (c.2000 built, dh[49])	300,000	93	80	225
TD7: Sullom Voe - Wilhelmshaven Aframax (c.2000 built, dh)	105,000	50	50	100
TD12: Antwerp - Houston Panamax (c.2010 built, heating dirty cargo)	74,000	42	27	76.5
TC1: Ras Tanura - Chiba (clean) LR2 (c.2010 built, products)	115,000	42	30	100
TC2: Rotterdam - New York (clean) MR (c.2010 built, products)	50,000	30	25	38

Table 6.2: Voyage assumptions for earnings (TCE) calculations on selected routes (Clarksons Research Services Limited, 2017)

Route	**Cargo size** [tonnes]	**Voyage dist.** [miles] Laden	Ballast	**Voyage time** [days] Sea time	Sea margin	Port time	**Total voyage time** [days]	**Speed** [knots] Laden	Ballast
TD1: Ras Tanura - LOOP	280,000	12,225	1,970	38.4	2.0	4.0	45.4	15.0	15.0
TD3: Ras Tanura - Chiba	265,000	6,654	6,654	37.0	1.8	4.0	42.8	15.0	15.0
TD7: Sullom Voe - Wilhelmshaven	80,000	600	600	3.4	0.2	4.0	7.6	14.5	14.5
TD12: Antwerp - Houston	55,000	5,100	5,100	33.4	1.7	4.0	39.1	13.5	12.0
TC1: Ras Tanura - Chiba (clean)	75,000	6,654	6,654	43.6	2.2	4.0	49.8	13.5	12.0
TC2: Rotterdam - New York (clean)	37,000	3,383	3,383	22.2	1.1	4.0	27.3	13.5	12.0

ditionally, note that we have constructed some of the variables by summing time series together. This matter variables for clean oil products, like demand, import and export, and refinery output. Time series' were generally only available on single clean distillate level[50], so we summed these together.

As for the *time horizon*, we have collected data from February 2002 to June 2017. However, considering a maximum lag of 6 months, the "effective period" we consider spans from August 2002 to June 2017. We discuss, further, how we deal with lag structures in the following chapter (Chapter 7).

[50]See Appendix A.1 for the breakdown of clean constituents summed together.

6.4 Data Pre-Processing and Presentation

Pre-processing of some of the variables were required. We found four out of six dependent variables to have the occurrence of negative values, i.e. negative TCE earnings. We handled this by qualitatively smoothing the value of these data points in accordance with observations before and after the data point, while ensuring that the direction of change remained unchanged[51]. TD1, TD3, TD7 and TC2 had the occurrence of 1, 4, 4 and 1 negative values, respectively. Further pre-processing required, such as the concept of stationarity, is given next in the Methodology chapter (Chapter 7).

In Figure 6.1, we have presented the $/day freight

[51]As an alternative, we could have added a constant term to the entire time series'. Considering the low count of conflicting values, we found this to be of less importance (as well as avoiding unnecessary "human-error bias").

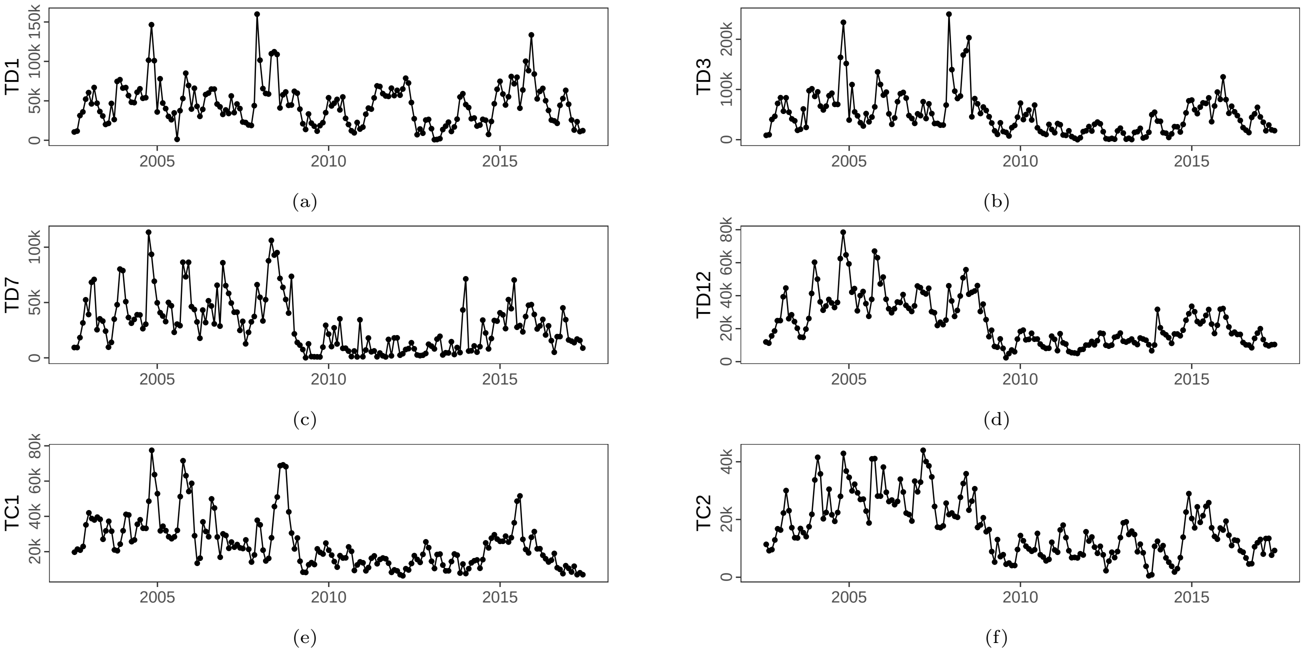


Figure 6.1: Displaying the freight rate time series' for each selected route.

Table 6.3: Descriptive statistics for the relevant variables 1/3.

Variable	Mean	Median	Min	Max	SD	Kurtosis	Skew	JB	ADF
TD1	45.7k	44.3k	.8k	159.9k	27.1k	2.134	1.076	71***	-3,9**
TD3	50.2k	42.2k	.0k	250.1k	41.8k	4.990	1.851	296,4***	-3,4*
TD7	30.2k	26.5k	.1k	113.5k	24.7k	0.572	1.009	33,7***	-3,7**
TD12	24.2k	19.7k	2.4k	78.5k	14.8k	0.563	0.962	30,8***	-3,4*
TC1	24.6k	21.3k	6.3k	77.5k	14.5k	1.769	1.350	80,1***	-4**
TC2	17.3k	15.7k	.5k	44.0k	10.0k	-0.336	0.647	13,4***	-3,6**
US_dem	19.7	19.6	17.8	21.7	0.8	-0.867	0.190	6,4**	-0.900
Jap_dem	4.8	4.8	3.6	6.8	0.7	-0.226	0.506	8,1**	-7,1***
Eur_dem	7.6	7.5	6.6	8.7	0.5	-0.872	0.149	6**	-3,9**
US_fdem	0.5	0.5	0.2	1.1	0.2	-0.978	0.300	9,5***	-2.900
Eur_fdem	0.5	0.5	0.2	0.9	0.2	-1.293	0.107	12,4***	-5,2***
Jap_cdem	3.0	2.9	2.0	3.7	0.4	-0.667	0.170	3.9	-6***
Eur_cdem	5.7	5.7	5.1	6.4	0.3	-0.226	-0.183	1.3	-3,9**
US_cdem	14.7	14.7	13.6	15.8	0.4	-0.526	-0.099	2.1	-1.900
AG_exp	17.3	17.5	13.0	21.4	1.5	0.464	-0.113	2.3	-2.400
NS_exp	2.6	2.4	1.4	4.9	0.9	-0.380	0.811	20,9***	-1.500
Bel_fexp	0.1	0.1	0.0	0.2	0.0	1.717	0.589	33,9***	-3,3*
US_fexp	0.3	0.3	0.1	0.6	0.1	-0.346	-0.034	0.8	-2.600

rate development for the 6 selected routes, adjusted for aforementioned negative values.

6.5 Descriptive Statistics

Tables 6.3, 6.4 and 6.5 display a selection of descriptive statistics for the monthly freight rates in their original form, as well as the explanatory variables. We present the four statistical moments, mean, standard deviation, skewness, and kurtosis, of the respective dependent variables. Furthermore, statistical tests describing normality and autocorrelation are presented[52][53].

By looking at the column for the ADF test statistic, it is apparent that we will have to perform transformations on some of the variables.

[52] See Appendix B for a mathematical formulation of the tests. Normality: JB test. Autocorrelation: (A)DF test.

[53] Engle's ARCH test could have been a method to verify ARCH effects, as a high ARCH order is necessary to catch the dynamics of conditional variance in freight returns. However, this has been widely approved by researches conducted earlier (see, for instance, IAME (2014); Abouarghoub and Mariscal (2011)).

Table 6.4: Descriptive statistics for the relevant variables 2/3.

Variable	Mean	Median	Min	Max	SD	Kurtosis	Skew	JB	ADF
USBel_fexp	0.0	0.0	0.0	0.1	0.0	2.652	1.149	95,1***	-5,4***
SA_cexp	0.4	0.3	0.0	1.2	0.3	0.394	1.272	50,5***	-1.100
Ne_cexp	1.2	1.3	0.7	1.7	0.3	-1.489	-0.169	17***	-2.200
US_cexp	1.1	1.0	0.2	2.6	0.7	-1.449	0.240	17***	-3,2*
USNe_exp	0.2	0.2	0.0	0.5	0.1	-0.509	0.388	6,3**	-4,4***
SA_exp	7.0	7.2	5.3	8.3	0.7	-0.164	-0.751	17,2***	-3,3*
VLCC_fix_west	19.5	19.0	6.0	41.0	6.1	0.448	0.529	10,3***	-3,4*
VLCC_fix	158.3	156.0	89.0	230.0	26.0	0.056	0.251	2.0	-3.000
VLCC_due	10.6k	10.6k	.6k	26.8k	4.4k	0.233	0.177	1.5	-4,5***
VLCC_fix_east	74.6	73.0	40.0	121.0	17.1	-0.500	0.112	2.0	-2.900
VLCC_fix_jap	7.4	7.0	0.0	22.0	4.4	0.453	0.798	21,1***	-3,3*
Afra_fix_sum	60.0	61.0	11.0	121.0	19.0	0.251	-0.259	2.7	-3.100
Afra_fix	445.7	415.0	285.0	736.0	107.7	-0.984	0.523	15,2***	-2.000
Pana_fix_US	5.5	5.0	0.0	14.0	2.9	-0.087	0.595	10,8***	-4,1***
Pana_fix	47.0	46.0	21.0	90.0	12.7	0.355	0.625	13***	-2.700
Afra_fix_east	18.7	18.0	4.0	47.0	8.5	-0.040	0.553	9,3***	-3,7**
Afra_fix_us	0.3	0.0	0.0	5.0	0.7	10.451	2.779	1072,7***	-4,9***
MR_fix_US	34.6	35.0	9.0	63.0	12.4	-0.675	0.181	4.1	-3,2*
MR_fix	569.5	527.0	272.0	1109.0	173.5	0.496	0.946	29,3***	-2.100
US_sea_imp	6.7	7.1	3.9	9.1	1.6	-1.344	-0.349	16,8***	-3,9**
US_imp	9.0	9.1	7.1	10.8	1.1	-1.300	-0.192	13,4***	-3,8**
SAUS_imp	1.3	1.3	0.7	2.2	0.3	0.124	0.254	2.1	-3,5**
SAPadd3_imp	0.8	0.8	0.4	1.5	0.2	0.637	0.332	6,8**	-3,3*
PADD3_imp	4.9	5.3	2.5	7.0	1.2	-1.294	-0.304	14,9***	-4,1***
Jap_imp	3.2	3.2	2.5	4.3	0.4	-0.615	0.321	5,7*	-7***
Eur_imp	5.6	5.6	3.9	7.1	0.8	-1.147	-0.144	10,1***	-3.100
Ge_imp	2.0	2.0	1.6	2.4	0.2	-1.285	0.197	13,1***	-3.000
Bel_fimp	0.1	0.1	0.0	0.2	0.0	-0.130	0.470	6,8**	-3,5**
US_fimp	0.3	0.3	0.1	0.7	0.1	0.615	0.797	22,5***	-4**
Padd3_fimp	0.1	0.1	0.0	0.2	0.0	0.254	0.797	19,9***	-3,3*
Jap_cimp	0.3	0.4	0.0	0.7	0.2	-1.587	-0.236	20,1***	-2.300
Ne_cimp	0.8	0.9	0.3	1.4	0.3	-1.499	-0.185	17,4***	-2.100
US_cimp	1.3	1.2	0.7	2.4	0.3	0.211	0.545	9,5***	-3.100
Major_imp	15.6	15.9	11.0	19.3	2.6	-1.416	-0.253	16,5***	-3,4*
Chi_imp	4.1	4.0	1.2	8.7	1.8	-0.656	0.436	8,7**	-2.400
Ind_imp	2.9	2.7	1.4	4.6	0.9	-1.339	0.201	14,2***	-3,9**
BelUS_fimp	0.0	0.0	0.0	0.1	0.0	1.392	0.714	30,9***	-4,1***
BelPadd3_fimp	0.0	0.0	0.0	0.1	0.0	0.070	0.179	1.1	-4,9***
NeUS_cimp	0.1	0.1	0.0	0.3	0.1	0.287	0.692	15,3***	-3,4*
NePadd3_cimp	0.1	0.1	0.0	0.2	0.0	0.230	0.676	14,4***	-3,7**
ME_prod	23.0	22.9	19.3	26.9	1.4	0.496	0.583	12,5***	-2.000
NA_prod	15.8	14.6	12.6	20.1	2.2	-0.756	0.915	29,4***	-1.400
US_prod	6.4	5.6	4.0	9.6	1.6	-0.872	0.856	27,6***	-1.800
SA_prod	9.2	9.4	7.1	10.7	0.8	-0.479	-0.424	7**	-3,2*
NS_prod	3.8	3.7	2.3	6.0	1.0	-0.906	0.527	14,2***	-2.000
OPEC_prod	30.0	30.0	24.7	33.9	1.8	0.238	-0.380	5*	-3.000
W_prod	88.1	86.8	76.2	98.8	5.4	-0.714	0.140	4.1	-2.600
Padd3_refuti	89.6	90.1	59.8	99.1	5.3	5.902	-1.582	344,6***	-3,7**
Padd1_refuti	83.7	85.5	57.8	97.7	8.7	0.202	-0.829	21,3***	-2.900
US_fout	0.6	0.6	0.4	0.7	0.1	-1.277	-0.231	13,4***	-3,4*
Bel_fout	0.1	0.1	0.0	0.2	0.0	-0.148	0.226	1.6	-3,8**
SA_cout	1.1	1.0	0.7	2.0	0.3	0.250	1.167	42***	-1.100
Ne_cout	0.9	0.9	0.6	1.1	0.1	0.825	-1.128	44,2***	-3.000
US_cout	14.5	14.5	12.8	16.4	0.8	-0.629	0.027	2.7	-3,8**
Yen_USD	104.1	107.0	76.6	123.9	13.8	-0.926	-0.510	14***	-1.400
USD_Pound	1.7	1.6	1.2	2.1	0.2	-0.359	0.129	1.3	-3.000
USD_Eur	1.3	1.3	1.0	1.6	0.1	-0.502	-0.056	1.8	-2.600
SDR_USD	1.5	1.5	1.3	1.6	0.1	-0.716	-0.330	6,9**	-2.500
Euro_index	125.1	126.6	96.7	155.3	12.7	-0.509	-0.061	1.8	-2.600
USD_index	81.9	81.0	69.1	104.1	8.4	-0.550	0.548	11,1***	-2.100
GDP_w	100.7	99.6	75.1	124.5	14.1	-1.084	-0.130	8,9**	-3,2*
US_CPI	216.1	218.0	180.7	245.0	19.1	-1.160	-0.327	12,9***	-2.000
Jap_CPI	101.2	100.7	99.2	104.0	1.4	-0.776	0.783	22,8***	-1.700
Eur_CPI	92.4	92.2	79.6	102.0	6.9	-1.302	-0.293	14,9***	-0.400
Ind_US	1.6	2.3	-13.6	8.2	4.0	4.629	-1.991	286***	-4,2***
Ind_Jap	0.7	1.8	-38.4	31.3	9.4	4.639	-1.051	200***	-4,1***
Ind_Eur	0.7	1.4	-19.1	9.3	4.9	4.578	-1.887	270,2***	-4,1***
Ind_OECD	1.0	1.8	-18.9	10.5	5.0	5.966	-2.148	414,5***	-4***
Ind_China	12.6	13.3	5.4	23.2	4.6	-1.136	0.014	9,3***	-3,6**
Ind_India	5.0	5.3	-5.1	17.6	4.6	-0.328	0.225	2.2	-3.100
LIBOR	1.6	0.9	0.2	5.6	1.7	-0.113	1.163	41,1***	-1.800
LIBOR_Yen	0.2	0.2	-0.1	1.0	0.3	1.077	1.499	77,6***	-1.700
LIBOR_Eur	1.6	1.3	-0.1	5.3	1.5	-0.614	0.711	17,9***	-2.400
Jap_money	.5k	.5k	.4k	.7k	.1k	-0.137	0.868	22,9***	1.500
US_money	2.0k	1.7k	1.2k	3.5k	.7k	-1.077	0.647	21***	-1.000
VLCC_1tc	41.4k	38.3k	18.0k	90.0k	16.9k	-0.332	0.609	11,9***	-2.300
Afra_1tc	23.5k	20.1k	13.0k	43.5k	8.3k	-1.076	0.465	14,8***	-2.500
Pana_1tc	21.1k	18.6k	12.5k	37.1k	6.9k	-1.081	0.577	18,5***	-2.300

Table 6.5: Descriptive statistics for the relevant variables 3/3.

Variable	Mean	Median	Min	Max	SD	Kurtosis	Skew	JB	ADF
LR2_1tc	23.8k	21.0k	13.5k	42.5k	7.9k	-1.033	0.512	15,6***	-2.500
MR_1tc	17.8k	15.0k	12.0k	30.5k	5.3k	-0.822	0.844	26,3***	-2.100
MR_3tc	15.6k	14.0k	10.5k	21.8k	3.3k	-0.996	0.653	20***	-2.400
US_gaso	1.9	1.9	0.7	3.3	0.7	-1.125	0.136	9,6***	-2.100
Bunker_Jap	437.0	386.3	174.5	781.0	184.7	-1.308	0.262	14,5***	-1.600
Bunker_Phil	396.1	346.1	141.6	745.9	176.8	-1.294	0.293	14,7***	-1.500
Brent	71.2	64.8	24.7	137.2	29.7	-1.190	0.265	12,3***	-1.700
WTI	68.7	66.3	26.1	140.0	25.6	-0.940	0.199	7,5**	-1.900
Dubai	68.3	63.8	23.3	131.2	29.4	-1.235	0.229	12,6***	-1.600
Oil_price_index	130.2	121.6	46.5	249.7	53.1	-1.194	0.212	11,6***	-1.800
ClarkSea	18.3k	14.3k	7.4k	48.5k	9.7k	0.341	1.086	36,9***	-3.000
ClarkAve	25.1k	21.9k	6.3k	79.7k	14.1k	0.575	0.935	29,4***	-3,4**
BDTI	1.0k	.9k	.5k	3.1k	.4k	2.970	1.559	142,6***	-4,5***
BCTI	.8k	.7k	.4k	1.9k	.3k	0.410	0.946	28,6***	-4***
VIX	19.5	16.7	10.4	59.9	8.4	4.633	1.941	280,4***	-3.000
SP500	1.4k	1.3k	.7k	2.4k	.4k	-0.631	0.658	15,9***	-1.500
VLCC_age	8.5	8.5	7.5	9.9	0.6	-1.007	0.005	7,2**	-1.500
Afra_age	9.1	9.0	8.0	11.6	0.9	0.338	0.912	26,3***	-2.900
Pana_age	10.3	9.5	7.5	16.7	2.6	-0.062	1.050	33,5***	-0.500
MR_age	10.8	10.1	8.5	15.1	1.9	-0.738	0.712	19,2***	0.000
VLCC_down	435.7k	272.6k	.0k	3780.8k	609.6k	8.144	2.398	684,5***	-3,6**
VLCC_deliveries	988.4k	920.8k	.0k	3643.4k	647.9k	0.910	0.790	25,7***	-2.400
Afra_down	185.0k	176.0k	.0k	674.2k	154.1k	-0.209	0.700	15,1***	-3,7**
Afra_deliveries	498.0k	444.0k	.0k	1632.1k	300.9k	0.534	0.681	16,5***	-2.700
Pana_down	51.9k	.0k	.0k	357.6k	73.1k	2.944	1.694	154,7***	-4,8***
Pana_deliveries	135.7k	138.6k	.0k	513.1k	122.0k	-0.233	0.738	16,8***	-3,4*
LR2_down	63.9k	.0k	.0k	496.7k	107.9k	4.064	2.030	253,1***	-3,4*
LR2_deliveries	179.8k	112.5k	.0k	753.7k	167.2k	1.225	1.150	52,2***	-2.900
MR_down	138.1k	121.7k	.0k	527.4k	99.8k	1.760	1.225	70,1***	-3,7**
MR_deliveries	424.9k	375.9k	75.5k	970.0k	191.5k	-0.051	0.663	13,3***	-2.000
VLCC_mdwt	163.4	160.5	122.3	222.3	28.6	-1.242	0.214	12,5***	-2.200
VLCC_yy	0.0	0.0	-0.1	0.1	0.0	0.160	-0.410	5,4*	-3,9**
Afra_mdwt	82.6	87.3	51.2	107.7	16.9	-1.345	-0.374	17,4***	-1.700
Afra_yy	3.3	3.7	-3.1	11.3	3.3	-0.600	-0.078	2.6	-4,4***
Pana_mdwt	18.4	21.1	7.1	25.9	6.5	-1.376	-0.498	21,3***	-0.800
Pana_yy	8.9	7.5	-4.3	24.2	8.5	-1.407	0.210	15,7***	-4,1***
LR2_mdwt	19.8	21.1	7.3	36.1	8.7	-1.384	0.043	14***	-2.500
LR2_yy	11.5	10.8	0.0	38.3	7.4	3.172	1.539	150,2***	-3,2*
MR_mdwt	65.3	69.4	40.0	91.6	15.7	-1.276	-0.127	12,3***	-2.300
MR_yy	5.6	5.4	0.3	11.9	2.8	-0.787	0.355	8,2**	-3.000
VLCC_new	1171.8k	634.5k	.0k	10385.6k	1536.0k	10.125	2.670	1003,5***	-5,2***
Afra_new	550.7k	440.0k	.0k	5060.4k	641.3k	13.795	2.859	1706,7***	-3.100
Pana_new	151.3k	.0k	.0k	2089.1k	250.6k	19.990	3.509	3431,5***	-4,7***
LR2_new	211.7k	104.5k	.0k	1887.0k	335.6k	7.775	2.569	665,3***	-2.900
MR_new	477.0k	376.0k	.0k	3784.1k	474.1k	12.818	2.613	1467***	-3,9**
VLCC_order	39401.2k	34430.3k	17337.4k	79755.1k	16721.2k	-0.951	0.569	16,2***	-1.700
VLCC_order_fleet	24.9	19.9	9.1	54.4	11.5	-0.815	0.667	18,1***	-1.800
Afra_order	11503.0k	11750.1k	3798.7k	19834.6k	4653.4k	-0.852	-0.037	5,2*	-1.900
Afra_order_fleet	15.2	15.7	3.9	27.5	7.5	-1.430	-0.090	15,1***	-1.600
Pana_order	5479.1k	5297.9k	939.8k	10387.2k	2818.0k	-1.236	0.131	11,6***	-2.500
Pana_order_fleet	38.5	30.5	3.9	82.4	27.3	-1.538	0.268	19,5***	-2.600
LR2_order	6606.0k	5705.7k	1553.6k	14161.9k	3394.0k	-0.289	0.773	18,6***	-2.400
LR2_order_fleet	41.5	31.8	5.9	107.2	28.5	-0.125	0.991	29,9***	-2.600
MR_order	15573.3k	13859.3k	6853.3k	27780.9k	5945.6k	-0.582	0.787	21,1***	-2.100
MR_order_fleet	25.9	23.0	7.5	50.8	12.4	-0.958	0.507	14,3***	-2.300
VLCC_price	104.8	99.0	62.5	162.0	23.4	-0.305	0.489	7,8**	-2.400
VLCC_SP	64.3	58.0	34.0	135.0	25.6	0.071	0.962	28,1***	-2.300
VLCC_demo_price	339.4	342.5	143.0	582.5	93.3	-0.395	0.222	2.5	-3.000
Afra_price	54.8	53.5	33.5	82.5	10.5	0.013	0.372	4.2	-2.600
Afra_SP	33.6	29.0	16.0	64.0	14.0	-0.863	0.728	21,3***	-2.100
Afra_demo_price	6.7	6.8	2.5	12.4	1.9	0.481	0.271	4.2	-2.900
Pana_price	46.9	45.5	29.5	68.0	8.5	0.012	0.567	9,8***	-2.200
Pana_SP	37.9	36.0	20.0	62.0	11.0	-0.503	0.704	16,7***	-2.400
Pana_demo_price	5.4	5.4	2.1	9.5	1.5	-0.373	0.114	1.3	-2.800
MR_price	38.2	36.0	26.3	53.5	6.7	-0.325	0.727	16,7***	-2.400
MR_SP	19.0	16.0	10.5	37.0	7.2	0.855	1.394	65***	-2.300
MR_demo_price	3.4	3.4	1.2	6.7	1.1	0.463	0.383	6,3**	-2.700
Brent_forw	0.6	0.9	-6.0	10.2	2.8	-0.270	0.251	2.3	-2.600
Tadawul	7.6k	7.0k	2.4k	19.5k	2.9k	2.915	1.327	119,7***	-2.900
Nikkei	13.1k	11.9k	7.6k	20.6k	3.7k	-1.268	0.312	14,6***	-1.700
USD_SAR	3.8	3.8	3.7	3.8	0.0	100.252	-8.807	79062,2***	-4,8***
MSCI_w	1.3k	1.3k	.7k	1.9k	.3k	-0.973	-0.084	6,9**	-2.600
MSCI_e	.8k	.9k	.3k	1.3k	.3k	-0.551	-0.685	16,3***	-2.600

7 Modelling Methodology

We now present our approach for creating a forecasting model. In this chapter we will look at the underlying theory and reasoning behind the approach, before presenting and interpreting the results in the following chapter. We will create separate models for all 6 routes. A novel variable selection approach based on regularization will be used to rank and statistically evaluate the potential candidate variables from the last chapter, Chapter 5. This ensures we have a solid foundation in the decision making of which variables to include in the final model(s). To evaluate the forecasting model for each route, we will be considering three main objectives:

Objectives

Assess the forecasting capabilities of our regime-model: It is imperative to know how well the model(s) actually forecasts the freight rate. Regime-switching forecasting models will be created for each route and compared with benchmark models on out-of-sample data. Common performance metrics will be applied, and the impact of the different determinants on the freight rate will be reviewed.

Evaluate the route impact: We will investigate the impact of creating route-specific models. Will a route-specific model have a significantly different performance, compared to a generic tanker model? We will assess this by comparing each route model against a generic tanker model which uses non-route-specific variables.

Consider a parsimonious model: The issue of overfitting, whereby the model becomes too tailored to the random noise in the sample data, is a risk often incurred in models which include too many variables. We therefore investigate whether a parsimonious model will fare better in out-of-sample forecasting than a more comprehensive, sweeping model with more variables.

Two models will be created for each route. One based on domain knowledge and a theoretical foundation, containing a broad selection of features. The other, a feature-limited model, based on the best performing variables from our regularization approach for each route.

Note, this means that we will be working with a total of 12 models (excluding benchmarks), one variable rich and one parsimonious model, for each route.

We will also be considering the significance of the selected predictors, the regimes, vessel size, oil trade types (crude or products) amongst others.

7.1 Stationarity

To be able to forecast properly and get meaningful results, it is imperative that the time series are stationary. Here we will present the concept and definition of stationarity and why it is necessary. We will also present our method of choice for testing this. Lastly, we explain how we will transform the data.

Definition of stationarity

A time series is said to be stationary if its probability distribution does not change over time (Pelagatti, 2013). A more formal definition is given by Alexander (2008a) whereby a discrete time stochastic process $\{X_t\}_{t=1}^{T}$ is stationary if

1. $E(X_t)$ is a finite constant
2. $V(X_t)$ is a finite constant
3. The joint distribution of (X_t, X_s) depends only on $|t-s|$

The first condition simply implies that the expected value should be the same for every observation, i.e. it does not trend at all. The second condition means that most observations will be in the proximity of the expected value, and not drift too far away, with the distance depending on the size of the variance. The third condition implies that the joint distributions between the variables are the same at any point in time, whether they are two steps apart, three steps apart and so forth. Since the definition of strict stationarity is generally too strict for everyday life, a weaker definition is usually employed. This only requires that the covariance is independent of the time it is measured, instead of the whole joint distribution (Alexander, 2008b). The latter definition is used in this article, and is simply referred to as «stationary».

Common issues

Many time series are non-stationarity, whereby the expected value and variance changes over time. These time series cannot reliably be used to forecast as they are unpredictable. The results from a model based on non-stationary data can be spurious, signifying relationships between variables where none exists. In such cases it is common that values like the adjusted R^2, the t-statistic and the Durbin-Watson statistic exhibit inflated or deflated values (Baumöhl and Lyócsa, 2009).

A stationary time series should not have a trend, nor seasonality. These will influence the time series at various points. However, a time series could consist of cycles. Cycles are very much present within shipping. Although cycles can sometimes be misinterpreted as seasonality, they are different. Cycles are not of fixed length. This in turn means we cannot accurately foresee where the various peaks and troughs of the cycles will be (Hyndman and Athanasopoulos, 2014). Cycles within the time series is therefore not expected to be an issue for stationarity.

How we test for stationarity, and transform the time series

A method of identifying non-stationary time series' is by examining autocorrelation function (ACF) plots. For a non-stationary time series the autocorrelations will slowly reduce, while for a stationary time series the autocorrelations is expected to quickly drop to zero.

Furthermore, one can test for stationarity using methods like an augmented Dickey-Fuller (ADF) test, see Appendix B.10, where one tests for the presence of a unit root.

Looking at Figures D.3 - D.9, most of the time series appear to be non-stationary, displaying pronounced trends. Non-stationary variables can in several cases be made stationary through transformations. This includes calculating the difference between two consecutive observations and stabilizing an increasing variation in the observations with logarithmic transformation. Furthermore, the removal of any deterministic trend and seasonal adjustments by decomposition can accomplish similar results. Initially, we attempt to perform logarithmic difference transformations, before pursuing other methods like absolute or percentage differences. Working with logarithmic differences is preferred due to the nature of prices often being approximately log normally distributed, as well as for the sake of computational brevity (Hudson and Gregoriou, 2014).

The augmented Dickey-Fuller test will be performed on all time series to evaluate stationarity. For non-stationary series, the necessary tranformations will be applied until the null hypothesis of a unit root can be rejected with a significance level of 5%. Time series that remain non-stationary will be excluded.

7.2 Standard Multiple Regression

Before attempting any advanced methods of modelling the freight rate, we run a regular multiple regression technique using ordinary least squares (OLS) to corroborate any relationships in the data (Alexander, 2008a). See a full explanation of multiple regression in Appendix B.9. Finding indications of relationships using OLS would provide the foundation for more complex methods.

7.3 Determining Lags

As we are creating a model to forecast ahead of time, it is imperative that we utilise lagged independent variables. The problem lies in selecting the best fitting lagged time series and/or combinations of lagged time series. For example, the fleet size three months ago might provide better predictions than the fleet size one month ago. We are therefore looking for lagged time series that can provide signs and patterns of the freight rate movements ahead of time.

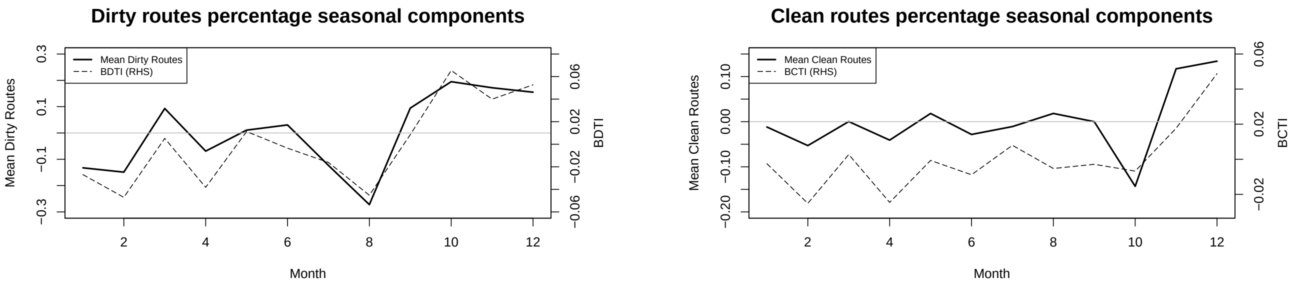
How many past lags are within our scope? If we are to brute force the selection of lags, and attempt every possible combination, we could quickly run into a computational problem in regards to the vast amount of possible combinations - often referred to as the *curse of dimensionality*. With approximately 170 variables, and looking at 6 lags for each variable, the total number of combinations equals 170^6, which is already a number in the trillions. Evaluating 12 lags for each variable would result in numbers the septillions. To limit the scope, we therefore decide to consider 6 months of lags.

The approach to determining lags will be to qualitatively assess cross-correlograms between each dependent variable and all relevant independent variables. This method ensures that we can include market domain knowledge in the lag selection, and make choices that pertain to a theoretical foundation. Variables that do not have any lags with a significant cross-correlation, will not be evaluated, and will default to a lag of 1.

7.4 Seasonality

Seasonality involves patterns that repeat every year (Diebold, 2017). As was evident in the literature review, the tanker freight rate appears to exhibit signs of deterministic seasonality. We rely on the findings of others, notably Kavussanos and Alizadeh (2002); Ringheim and Stenslet (2017); and Alizadeh and Nomikos (2009), and assume the presence of seasonality.

We therefore want to account for the seasonality that theoretically should be present in the tanker market. Our approach will be to create monthly dummy variables. A simple dummy variable takes the value 0 except during specific periods, where it takes the value 1. Such a dummy allows the model to shift during this period, resulting in a better fit (Grotenhuis and Thijs, 2015). As we are using monthly dummies, we need to include 11 variables in our regime-switching model. Note, we only include 11 variables, as the last month will be captured by the intercept. The associated coefficients can be interpreted as a measure of the effect



(a) Dirty routes

(b) Clean routes

Figure 7.1: Deterministic seasonality based on the mean seasonal change of our selection of dirty tanker (TD1, TD3, TD7, TD12) and clean tanker (TC1, TC2) routes, which is represented alongside their respective Baltic Exchange index. Serves as a documentation of the presence of seasonality in the time series'. The incorporation of seasonal dummies is discussed in Section 7.4. Calculation is based on the entire time series period from August 2002 to June 2017.

of each month on the dependent variable, relative to the omitted month (the intercept). The deterministic seasonal component can be seen in Equation 1.

$$S_t = \sum_{i=1}^{s} \gamma_i D_{it} \tag{1}$$

Where γ_i is considered a seasonal factor, and D_{it} is a dummy variable with e.g. $D_{1t} = (1, 0, 0, \dots)$ as described by Diebold (2017). These dummy variables will later be included in the regime-switching model.

Seasonality is generally accounted for as a fixed and known frequency in modelling (Hyndman, 2014). However, as we are modelling two distinct regimes that are hypothesised to have distinct qualities, we expect seasonal effects to be different between these regimes.

In Figure 7.1, we present the mean of the seasonal changes of our selection of dirty and clean routes. We anticipate these seasonal components to be broadly reflected in the supposed low volatility regime of the regime models, as this regime should have the highest weighting of total number of observations, and thus be considered the «normal state» regime (see, for instance, Abouarghoub et al. (2014), (2011) and Abouarghoub et al. (2012)). In Appendix A.4, a complete representation of each route's seasonal components are represented alongside their respective Baltic Exchange index.

Our mix of dirty tanker routes seems to exhibit higher irregularity and magnitude in their seasonal derivatives than our mix of clean routes. As occasionally discussed up to his point, the monthly increase in freight rates throughout the late fall and early winter months, November and December, is induced by a mix of the derived energy shipping demand and seasonal climate effects in this period of the year. This seems to be the case for both the dirty and the clean routes. As for the rest of the year, we observe that the seasonal pattern is quite similar, although components of the clean routes are fluctuating much tighter around the 0 percentage line; indicating less volatile earnings for these vessels. A decrease in January is followed by an increase in February, and then the sign of the double-derivative shifts each month until June. The biggest difference is seen from June to October, where the dirty routes continue the negative trend until they increase from September throughout December, while the clean routes continue fairly stable before having a dip in October, and then accelerating through November and December. Similar deterministic seasonal patterns are found in the study by Kavussanos and Alizadeh (2002), although they are investing a 20-year period pre-2000. In the following chapter, Chapter 8, we will address whether we are able to match these effects in our models.

7.5 Variable Selection

A preliminary candidate predictor analysis will be performed. We are interested in examining if the initial variables we have selected are appropriate, and rank them based on how well they can presumably explain the freight rates. This will provide an indication, as well as necessary support in the decision making of which variables to include in the final model (alongside the theoretical foundation gathered in the earlier chapters). Our method of choice for ranking candidate variables, will be a combination of stability selection with the penalization of LASSO accompanied with a random tuning parameter, hereby referred to as simply *stability selection*. As mentioned by Hasinur et al. (2016), by properly balancing goodness of fit and model complexity, penalization approaches can lead to parsimonious models.

As we are managing approximately 170 different time series, it is becoming necessary that we utilise such methods to gain statistical insight into this data. We were limited to only approximately 180 observations, by the amount of available past observations for many of the time series'. We are facing a situation where the amount of available variables are very close the amount of observations. Regularization techniques would allow us to estimate models with fewer observations than parameters (Hyndman, 2014).

We have deemed certain approaches inadequate or unsuited for these purposes. Among these are *least absolute shrinkage and selection operator* (LASSO), *principal component analysis* (PCA), *stepwise regression* and *ridge regression*. LASSO does select the best performing candidates, while at the same time forcing other candidates close to zero. As such, it fulfils the need for feature selection, but lacks in being able to interpret the data. The same is true for PCA as it creates new linear relationships, which in turn reduces the ability to interpret and explain the factors. Stepwise regression is avoided due to the potential pitfalls such as yielding high, biased R^2 values, the standard errors of the parameters being too small and it providing p-values without proper meaning (Harrell, 2001). Lastly, ridge regression could potentially force the coefficients to spread out similarly between correlated variables, and is thus avoided (Doreswamy and Vastrad, 2013).

Instead, as mentioned in the introduction, we will utilise the method of *stability selection* (Appendix B.5), a novel approach to feature selection, which combines subsampling with selection algorithms.

How does it work?

Stability selection is a technique on its own. This is simply the technique of running a model on multiple subsets of the data in order to select the best variables. A common interpretation of the practical implementations available is to use subsamples of both the observations (i.e. various time periods) and subsamples of the actual variables. This means that a model is run on a smaller amount of observations, with fewer variables, many times. Stability selection does however require a selection algorithm.

A typical selection algorithm is LASSO, which utilizes ℓ_1-penalization. Here the coefficients of most variables are forced to zero. The tuning parameter λ determines how strict it should be in leaving out variables. A simple enhancement of LASSO which has great practical benefits is randomized LASSO (Meinshausen and Bühlmann, 2009). The difference here is that regular LASSO is dependent upon specifying a good initial set of available tuning parameters Λ, while randomized LASSO can function very well without such rigid initial values. Randomized LASSO uses randomly selected tuning parameters λ for each run. For one single model run, this would provide inadequate results, but when implemented along with stability selection which is based on multiple runs, the results are very favourable.

How does this work for us?

The underlying idea is to run variable selection routines on subsets of the data, with various subsets of variables. This approach enhances the performance of existing methods. After numerous runs, the results can be aggregated by for example examining how frequently a variable ended up being selected as relevant. The strongest candidates will have scores close to one, while the weaker relevant candidates will have non-zero scores. These latter candidates would be selected during runs where the strongest candidates might not be present. Irrelevant candidates should have scores close to zero. This method is useful for both variable selection to reduce overfitting, as well being able to interpret data (Hasinur et al., 2016).

Approach

In order to look for and select interesting candidate variables, we will be taking multiple approaches with the help of stability selection.

Firstly, in the way variables are included in the multiple runs. Stability selection will be run on a set of all variables, as well as route-specific subsets of variables. These route-specific subsets contain the variables that we have already deemed most likely to impact the various freight rates, based on our domain knowledge. This approach allows us to review the robustness of the variables that we have picked, as well as look for any new ones, which might play a significant role.

Secondly, two different fractions of samples will be utilised for each randomized design, one smaller sample fraction and one larger sample fraction. In other words, stability selection's regression models will be run on data samples spanning either a few months, or several years. This will allow us to identify candidate variables that better fit a *multi-regime model*. We expect to find a discrepancy in the variables that are frequently selected for short time periods versus long time periods, by the stability selection algorithm. Some variables might perform far better on shorter time horizons, which better coincides with the abnormal high volatility state of the market. Longer time horizons are potentially a better representation for the normal state, and different variables could have a higher degree of impact on the freight rate.

The model will be tuned with a sample fraction of 0.05 and 0.3. This represents approximately 9 and 54 observations respectively. Each stability selection run will be performed with 10,000 different models, and a maximum number of iterations equal to 100,000.

After having run the regularizations, we *hand-pick* a selection of the most promising candidate variables for each route. We hand-pick in order to apply further domain knowledge during the selection, and avoid just performing an automatic selection of variables.

Additionally, we will use the top candidates for each route to create parsimonious regime-switching models, to compare their performance to our domain knowledge regime models. This is discussed further in Section 7.8.

7.6 Model Formulation

Chow test

To look for initial empirical signs of multiple regimes in the dependent data, we will perform a *Chow test*. This is a simple and intuitive way of uncovering structural breaks. As seen in Alexander (2008a), a regression model consisting of structural breaks in the parameters at time t^* can be defined as

$$\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\epsilon}_1, \quad \text{for } t = 1, \ldots, t^* \\ \boldsymbol{y}_2 &= \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon}_2, \quad \text{for } t = t^*+1, \ldots, T \end{aligned} \tag{2}$$

where

$$\boldsymbol{y}_1 = (Y_1, \ldots, Y_{t^*})', \quad \boldsymbol{y}_2 = (Y_{t^*+1}, \ldots, Y_T)',$$
$$\boldsymbol{\epsilon}_1 = (\epsilon_1, \ldots, \epsilon_t)', \quad \boldsymbol{\epsilon}_2 = (\epsilon_{t^*+1}, \ldots, \epsilon_T)',$$

$$\boldsymbol{X}_1 = \begin{pmatrix} X_{11} & X_{21} & \cdots & X_{k1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1t^*} & X_{12t^*} & \cdots & X_{kt^*} \end{pmatrix}$$

$$\boldsymbol{X}_2 = \begin{pmatrix} X_{1,t^*+1} & X_{2,t^*+1} & \cdots & X_{k,t^*+1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1T} & X_{2T} & \cdots & X_{kT} \end{pmatrix}$$

We will perform the Chow test as presented by Alexander (2008a), where one compares the combined residual sum of squares of multiple estimates of linear models for sub periods of the data, with one linear estimate for the whole period. We will do the following steps:

1. First we run a regression of the model up to time t^*. We then proceed to run a regression on the rest of the data. We must calculate the residual sum of squared errors on each of these sub-models.

2. We then add the two residual sums of squares to get the unrestricted residual sum of squares, RSS_U.

3. We proceed to run a regression on the whole data set. This represents the scenario where there are no regimes present in the data. We estimate the restricted residual sum of squares, RSS_R.

4. We complete the test by using an F test to see if there is a significant difference between these two models - the one containing structural breaks and the other ordinary linear regression model.

Markov regime-switching model

As we have seen up until now, Markov regime-switching seems to be an appropriate approach to modelling tanker freight rates. These models allow for greater flexibility than a standard multiple linear regression, and can account for different dynamics in each regime. They appear fitting for the tanker freight rate which exhibits a more volatile behaviour in states where the effects of supply and demand constrain the market.

Markov regime-switching has its foundation in Hamilton's paper released in 1989, where he statistically formalized that different economic states can affect the behavior of the economic variables.

As we walk through the modelling process, we use an ordinary linear regression model as a placeholder, for the sake of brevity and simplicity. The Markov switching regression model can then be expressed as follows

$$Y_t = \alpha_{s_t} + \beta_{s_t} X_t + \epsilon_{s_t t}. \quad \epsilon_{s_t t} \sim N(0, \sigma^2_{s_t}) \tag{3}$$

The model is assumed to have normally distributed homoscedastic errors (Alexander, 2008a). The regime is given by the latent variable s_t, which can be realized into two different values:

$$s_t = \begin{cases} 1, \text{ if state 1 governs at time t,} \\ 2, \text{ if state 2 governs at time t.} \end{cases} \tag{4}$$

It is assumed that this state variable follows a first-order Markov chain, whereby the probability of being in any state is only dependent upon the previous state. In such a Markov chain, it is assumed that the transition probabilities are constant. The transition probabilities determine the probability of being in a certain state at time t, given a specific state in time $t-1$. The matrix of transition probabilities can be written as

$$\boldsymbol{\Pi} = \begin{pmatrix} \pi_{11} & \pi_{21} \\ \pi_{12} & \pi_{22} \end{pmatrix} = \begin{pmatrix} \pi_{11} & 1-\pi_{22} \\ 1-\pi_{11} & \pi_{22} \end{pmatrix} = (\pi_{ij}).$$

In this equation, π_{ij} represents the probability of going from state i to state j. We are also able to find the unconditional probability of, for instance, regime 1 by

$$\pi = \frac{\pi_{21}}{\pi 12 + \pi 21}. \tag{5}$$

The full set of parameters for the model can be given by the following vector

$$\boldsymbol{\theta} = (\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma_1, \sigma_2, \pi_{11}, \pi_{22})'. \tag{6}$$

To represent the states of the Markov chain, we use a random state indicator vector ξ_t, whereby the ith element is equal to 1 if $s_t = i$, and 0 otherwise. For a two-state regime model this vector can be written as

$$\boldsymbol{\xi}_t = \begin{pmatrix} \xi_t^1 \\ \xi_t^2 \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{if in state 1 at time t,} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{if in state 2 at time t,} \end{cases} \tag{7}$$

We consider these states unobservable, and we cannot be certain of which state the series is in at any time. However, we can estimate the conditional probability of being in a state. We are interested in the expected state vector at time t, at time $t-1$. We therefore introduce a new variable $\boldsymbol{\xi_{t|t-1}}$ for the conditional expectation of the state indicator $\boldsymbol{\xi_t}$ at time t, given all the information up to time $t-1$. From the definition of the transition matrix, this can be expressed as the transition matrix multiplied with the state indicator at time $t-1$

$$\boldsymbol{\xi}_{t|t-1} = E_{t-1}(\boldsymbol{\xi}_t) = \boldsymbol{\Pi}\boldsymbol{\xi}_{t-1}. \tag{8}$$

Our parameter estimation will be based on maximum likelihood estimation. Simply put, this is a technique to find the most likely function that explains observed data. In other words, if the data were to have been generated by the model, what parameters were most likely to have been used? (Halls-Moore, 2016). It is therefore essential to build the likelihood function using the sample data and the model. The estimation process is complicated by the necessity to estimate the conditional probabilities as well. We are therefore required to perform further iterations at each algorithmic step. In other words, for each step of the algorithm where log likelihood function is maximized, it is also necessary to estimate the conditional probabilities.

Now, to have a look at the sub-iterative process. The probability density of the normal distribution with expectation μ and standard deviation σ can be expressed as follows

$$\phi(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right].$$

We also assign starting values for the conditional expectation of the state indicator

$$\hat{\boldsymbol{\xi}}_{1|0} = \begin{pmatrix} \hat{\xi}_{1|0}^1 \\ \hat{\xi}_{1|0}^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

as well as the initial model parameters. We will utilise the values from a regular multiple regression as a starting point. This means that the parameters for both states will be equal from the start. The iteration proceeds as follows from $t = 1$:

1. $f_t(Y_t|X_t; \hat{\boldsymbol{\theta}}) = \hat{\xi}_{t|t-1}^1 \varphi(Y_t; \hat{\alpha}_1 + \hat{\beta}_1 X_t, \hat{\sigma}_1) + \hat{\xi}_{t|t-1}^2 \varphi(Y_t; \hat{\alpha}_2 + \hat{\beta}_2 X_t, \hat{\sigma}_2)$

2. Set
$$\hat{\boldsymbol{\xi}}_{t|t} = \begin{pmatrix} \hat{\xi}_{t|t}^1 \\ \hat{\xi}_{t|t}^2 \end{pmatrix} = \begin{pmatrix} \frac{\hat{\xi}_{t|t-1}^1 \varphi(Y_t; \hat{\alpha}_1 + \hat{\beta}_1 X_t, \hat{\sigma}_1)}{f_t(Y_t|X_t; \hat{\boldsymbol{\theta}})} \\ \frac{\hat{\xi}_{t|t-1}^2 \varphi(Y_t; \hat{\alpha}_2 + \hat{\beta}_2 X_t, \hat{\sigma}_2)}{f_t(Y_t|X_t; \hat{\boldsymbol{\theta}})} \end{pmatrix}$$

3. Set $\hat{\boldsymbol{\xi}}_{t+1|t} = \hat{\boldsymbol{\Pi}}\hat{\boldsymbol{\xi}}_{t|t}$.

4. Set $t = t + 1$ before returning to step 1, while repeating the iteration until t = T.

This iterative process provides us with two necessities:

- a set of conditional state probabilities $\{\hat{\boldsymbol{\xi}}_{t|t}\}_{t=1}^T$
- a set of conditional densities $\{f_t(Y_t|X_t; \hat{\boldsymbol{\theta}})\}_{t=1}^T$

The elements of conditional state probabilities serves us with the (conditional) probability of being in state 1 or 2.

We continue by estimating the model parameters $\boldsymbol{\theta}$ by maximizing the log likelihood function.

$$\ln L(\boldsymbol{\theta}) = \sum_{t=1}^T \ln f_t(Y_t|X_t; \boldsymbol{\theta}). \tag{9}$$

For each iterative step of this maximization, we perform the sub-iteration to get the conditional densities and the set of conditional state probabilities, using the currently estimated model parameters.

Specific regime-switching model

The regime-switching regression model we will be implementing can then be formulated as follows

$$\begin{aligned} Y_t^r &= \alpha_{s_t}^r + \beta_{1s_t}^r X_{1t} + \beta_{2s_t}^r X_{2t} + \ldots \\ &+ \beta_{ns_t}^r X_{nt} + \sum_{i=1}^k \gamma_{is_t}^r D_{it} + \epsilon_{s_t t}^r. \\ \epsilon_{s_t t}^r &\sim N(0, \sigma_{s_t}^2) \end{aligned} \tag{10}$$

where X_{ht} represents the explanatory variables, $\beta_{hs_t}^r$ is the associated route specific regime determined coefficient, with $h \in \{1, \ldots, n\}$. Furthermore, $\alpha_{s_t}^r$ denotes the state dependent intercept, $\sum_{i=1}^k \gamma_{is_t}^r D_{it}$ represents the seasonal dummy component as seen in Section 7.4 with $k = 11$, and $\epsilon_{s_t t}^r$ the residual of each observation. The regression model will be estimated for each of the six tanker routes $r \in \{TD1, TD3, TD7, TD12, TC1, TC2\}$, using different explanatory variables for each model.

This leaves the following model parameters to be estimated

$$\begin{aligned} \boldsymbol{\theta}^r = (&\alpha_1^r, \alpha_2^r, \beta_{11}^r, \ldots, \beta_{n1}^r, \beta_{12}^r, \ldots, \beta_{n2}^r, \\ &\gamma_{11}^r, \ldots, \gamma_{g1}^r, \gamma_{12}^r, \ldots, \gamma_{g2}^r, \\ &\sigma_1^r, \sigma_2^r, \pi_{11}^r, \pi_{22}^r)'. \end{aligned} \tag{11}$$

Forecasting with the regime-switching model

Having defined how we will establish the model itself, we also need to determine how the model will be utilised to generate forecast results at time $t + 1$.

An inital problem that arises is how we are going to select which regime's forecasted value to use as our predicted value - the regime-switching model consists of two regression models which both produces a prediction. A solution would be to obtain forecasts as the weighted average of the outputs from both models, weighted by their associated probability.

To determine this probability for time $t + 1$, one can obtain the filtered probability at time t. Multiplying this filtered probability with the transition matrix yields predicted probabilities of each regime at time $t + 1$.

In this way, we establish a forecast which incorporates both regimes at all times. The model takes lagged variables as input, produces a prediction for each regime, which is weighted by the estimated probability of being in each regime.

The forecast horizon is 1-month-ahead. The regime-switching model parameters are not re-estimated for each step. We vary the origin from which forecasts are made, but maintain a consistent forecast horizon, as described by Hyndman (2006). The residuals from the forecast will later be evaluated.

7.7 Assumptions

Underlying the regular OLS method are a set of assumptions. We will investigate and ensure that these OLS assumptions are met for the regime regression models. If they are not, the model may be inefficient, or highly biased. There are multiple ways of assessing this, and we will consider two general approaches. We will be performing analyses both qualitatively (plot-based) and quantitatively (test-based), to evaluate most of the assumptions. This ensures a better understanding of the underlying characteristics, and also validates the test approaches for each assumption.

Overview of the OLS assumptions

1. Linear relationship
2. No autocorrelation in the errors
3. Homoscedastic (constant) errors
4. Normally distributed errors with zero mean

Approach to testing the assumptions

1 There should exist a *linear relationship* between the dependent and independent variables. If no such relationship exists, the model will be incorrect and unreliable. Erroneous results from the extrapolation is thus likely. To investigate this qualitatively, we study the plot of fitted values versus residuals. For the assumption to be met, we expect to see a fairly flat horizontal line, with no significant patterns. The individual points should be distributed evenly around this horizontal line, and have a roughly constant variance.

2 A common issue in time series regression models is *autocorrelation* in the errors, whereby the error terms of different observations are correlated with each other. This is often a sign of room for improvement in the model, and in the extreme cases, a sign of a highly misspecified model. To assess this qualitatively, an inspection of residual autocorrelation plots will be performed. This plot shows the correlation of the residuals with its own lags. Neighbouring residuals with significant autocorrelation can indicate a need to reevaluate the model.

To analyse this assumption quantitatively, several options are available (Gujarati, 2011). The Durbin-Watson statistic (Appendix B.10) is a test for detecting significant residual autocorrelation at lag 1. We are looking for statistic values deviating significantly from 2, which pertain to autocorrelation. Small values of DW indicate positive autocorrelation and large values indicate negative autocorrelation (Alexander, 2008b). A drawback of the basic Durbin-Watson approach is the lacklustre amount of lags. We therefore also include the Breusch-Godfrey test. This test is more general than a standard Durbin-Watson statistic, and statistically more powerful. We are working with monthly frequency data, and have limited the amount of lags to six. As the test allows for a specified order of serial correlations to be tested, we will be investigating the results from orders between one and six.

3 *Heteroscedasticity* is a display of differently dispersed errors. The residual terms should all have the same variance. If the residuals on the other hand are heteroscedastic, then the standard error of the regression estimates cannot be trusted due to unreliable confidence intervals. This can cause our confidence intervals for out-of-sample predictions to be unrealistically narrow (Nau, 2017).

For the qualitative approach, we will be examining the plot of residuals against fitted values. In the case of heteroscedasticity, we are looking for disparities in the variance of the errors.

As for the quantitative method, we will be employing the Breusch-Pagan test (Appendix B.13). This test for heteroscedasticity is performed using an auxiliary regression of the squared residuals on the explanatory variables (Gujarati, 2011). The resulting statistic $n\hat{R}^2$

follows the chi-squared distribution.

4 The residuals are expected to be *normally distributed*. Violations of this assumption can cause issues for the estimation of confidence intervals for the forecast, as well as determining whether the model coefficients are significantly different from zero (Nau, 2017). Regression analyses are quite robust to deviations from the normal distribution, which means the residuals only have to be approximately normally distributed (Gujarati, 2011).

When reviewing the normality assumption qualitatively, an effective approach is to examine the Q-Q plot, as well as consider the histogram. For the Q-Q plot, the standardized residuals are plotted against ordered expected residuals, whereby the points should roughly fall on a diagonal line. If there is no misspecification in the model, these two plots can be interpreted sufficiently.

There exists several quantitative tests for normality, and the most popular approach is the Jarque-Bera test (Appendix B.12). The method takes the skewness and kurtosis as input, which provides a statistic. With the null hypothesis being that the residuals are normally distributed, the statistic is evaluated against the chi-square distribution.

7.8 Model Assessment

How will we evaluate the route-specific regime models? To simplify comprehension, we have divided most of the assessment into three objectives, as mentioned before. We will consider the benefit of implementing models with few variables. We will assess the forecasting capabilities of the model. Lastly, we will evaluate how relevant it is to specify models distinctly for a route. This is to ensure handling 2x6 regime models, as well as a plethora of benchmark models is manageable. The data will be split into a training set, and a never-before-seen test set. Various performance metrics and tests will be applied to the forecasts of these models, and compared with benchmark models.

The problem of overfitting

Overfitting can be thought of as fitting a model to the randomness in data. Such a model would perform vastly better when creating the model, than when actually using it on real world data. Researchers sometimes add variables to their model in the hopes of increasing the R^2 value, mistakenly believing that this complexity will make a better model (Gujarati, 2011). However, the inclusion of unnecessary variables might lead to a loss of efficiency of the estimators, reduced degrees of freedom, as well as the problem of multicollinearity. In other words, an overfit model corresponds too closely to a particular set of data. As we are interested in modelling the future, it is imperative that the model generalises well to never-before-seen data.

We will therefore be *partitioning the data into two sets*, a training set and a test set. The model will initially be trained on the training set. The forecasting model will then be evaluated on new validation data, other than the data that was used to train the model. As a common consequence of overfitting is poor performance on the validation data set (Domingos, 2012), we can evaluate how specialised the forecasting model(s) are. The data is split in a 80/20 ratio, with the most recent 20 % of the data earmarked for validation.

Besides partitioning the data set, another method to mitigate overfitting is by adding a regularization term to the evaluation function (Domingos, 2012). We implement this in the best candidate variable selection, randomized LASSO. LASSO ensures only a few best candidate variables are selected, and forces most features to zero.

We also ensure to consider parsimonious models to evaluate how well such models generalise, see 8.5.

Other approaches also exists, such as stepwise multiple regression where the selection of variables is automated. One builds a model by successively adding or removing variables based on statistical tests (Nau, 2017). We will not be employing such approaches, as we are interested in adding our domain knowledge to the modelling.

Performance measures

To assess the performance of the forecasting models, a range of various measures will be used. Several metrics are used both due to the commonality of their inclusion, as well as the compromises that exists for them individually. The compromises include how easy it is to understand and compute, if it can be compared across series, and how well it functions with zero or close-to-zero values. The selection of forecast metrics include scale-dependent metrics, percentage error metrics and scale-free error metrics. We will specifically be looking at the root-mean-squared error ($RMSE$), the mean absolute error (MAE), the mean absolute percentage error ($MAPE$), the median absolute percentage error ($MdAPE$), and the mean absolute scaled error ($MASE$).

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T} (y_t - \hat{y}_t)^2}{T}} \tag{12}$$

$$MAE = \frac{\sum_{t=1}^{T} |y_t - \hat{y}_t|}{T} \tag{13}$$

$$MAPE = \frac{1}{T} \sum_{t=1}^{T} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \tag{14}$$

$$MdAPE = \underset{t \in T}{median}\left(\left|\frac{y_t - \hat{y}_t}{y_t}\right|\right) \quad (15)$$

$$MASE = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{|y_t - \hat{y}_t|}{\frac{1}{T-m}\sum_{t=m+1}^{T}|y_t - \hat{n}_{t-m}|}\right) \quad (16)$$

Here, $\hat{y}_t$ represents the forecasted value, y_t is the actual value, T is the number of samples, $\hat{n}_t$ is the forecasted value of a one-step «naive forecasting method», and m is the seasonalperiod (or 1 if non-seasonal).

The MAE works well on a single series. It is simple to understand and calculate. It does not however, fare well on comparison between several series, as it is scale dependent. MAPE and Mdape are based on percentage errors, and can thus be used to compare between different data series. Despite this, these measures do not perform well with errors close to or equal to zero. The MASE is scaled based on the in-sample MAE from a naive forecast method. This means it can be used to compare across series, and does not have trouble with infinite values (Hyndman, 2006).

Benchmark models

To get an indication of the relative performance of our regime-switching models, we need to compare them with the performance metrics of benchmark models. We will use the following four benchmarks.

Random walk The next value is composed of the current value in addition to a white noise error term. This should be considered a minimum requirement to outperform (See Appendix B.7).

Mean The forecasts of all future values are simply equal to the mean of the historical data (Hyndman and Athanasopoulos, 2014) (See Appendix B.8).

Multiple linear regression The dependent variable is modelled as a linear combination of the explanatory variables (See Appendix B.9). These benchmark models will use the same variables as the regime-switching models for the different routes. From this we can assess the usefulness of including regimes.

ARIMA The forecasts of an autoregressive integrated moving average model consists of lags of a stationarized series (autoregressive) and lags of the forecast errors (moving average), as well as a constant (See Appendix B.6). A best fit $ARIMA(p, d, q)$ will be estimated for each route. The model is commonly applied to time series and should provide a potent benchmark.

Diebold-Mariano test

To better evaluate the forecasting capabilities of our regime-switching model, we are going to utilise the Diebold-Mariano test (See Appendix B.14). This test allows us to compare two forecasts of interest and gauge if one is significantly different than the other. The quality of each forecast is evaluated on some loss function of the forecast errors. For this test, the null hypothesis is that the methods are equally accurate on average. It must be noted that the DM test is not intended for model comparison, but merely to compare forecasts, as explained by the original researcher Diebold (2013). Hence, we will ensure the DM test is only used to assess predictive accuracy.

I Considering a parsimonious model

We attempt to increase the accuracy of the regime model by decreasing the complexity. In other words, we will also evaluate the performance of a regime model with far fewer explanatory variables than the initial regime model based on domain knowledge. Such a parsimonious model will be created for each route.

The parsimonious regime model will be generated by only including the most highly ranked variables from our candidate variable selection approach, based on *randomized LASSO*.

The parsimonious regime models will be evaluated against the domain knowledge regime models, based on common performance metrics, as well as the Diebold-Mariano test for forecasting accuracy.

II Assessing the forecasting capabilities of our regime-model

As it is crucial to evaluate the predictions performance of the regime models, we will apply them to our out-of-sample data set. The regime models will be compared against the previously mentioned benchmark models. Only the best performing regime models will be evaluated (either domain knowledge models or parsimonious models). This means we will have 6 regime models, 6 mean benchmark models, 6 linear regression benchmark models and 6 Arima benchmark models, one for each route.

The models will be evaluated based on common performance metrics, as well as the Diebold-Mariano test for forecasting accuracy.

III Evaluating the route impact

We will consider how impactful a route-specific model is. How can this be evaluated? To measure this, we will make a generic forecasting regime model based on explanatory variables with no specific relation to

Table 7.1: Global benchmark. Variables to be used solely for the generic benchmark regime model based on global variables. This model will be utilised to assess the impact of specifying models specifically for routes. If such a generic model greatly outperforms the route specific models, claims can be made regarding the need for such models.

Variable	Brief reasoning
Major_imp	World oil trade
Chi_imp	The largest importer (went past US in 18')
W_prod	World production measure
OPEC_prod	Oil trade and oil price is largely dependent on OPEC
USD_index	USD is the most important currency in shipping
Euro_index	The EU consumes a lot of oil
GDP_w	World economic measure
Ind_China	A globally impacting economy
Ind_US	A globally impacting economy
LIBOR	The most common interest rate benchmark
Brent_forw	The state of the oil market, and future expectations
Brent	The current state of the oil market
ClarkSea	A global shipping measure
ClarkAve	An overall tanker index
MSCI_w	Global state and economic measure
SP500	Global state and economic measure
VIX	World psychology and market expectations
Afra_yy	Aframax is the workhorse in the tanker market*

*Should ideally be a general tanker variable, but we were unable to collect this.

individual shipping routes. In other words, a benchmark regime model will be formulated which consists solely of global variables. This generic model will be compared to the route-specific regime models. Its ability to forecast will be evaluated based on common performance metrics, as well as the Diebold-Mariano test. If such a generic model greatly outperforms the route-specific models, assertions should be made for the usefulness of including very specific variables for each route.

The variables in Table 7.1 are going to make up this generic, global benchmark model.

8 Results

8.1 Initial preparation

Stationarity

The augmented Dickey-Fuller test was performed on all variables, as seen in Table D.1. At a 5% significance level, approximately one quarter of the 169 variables were initially found to be non-stationary. Different measures were taken to get all the time series stationary. Most series were transformed with logarithmic differencing, but some required different measures like double differencing. After the transformations, *all* variables were found to be stationary with a 5% significance.

Determining lags

To determine the appropriate number of lags for each explanatory variable, a visual inspection of *cross-correlograms* on a route by route basis was performed. This allowed us to qualitatively assess the cross-correlograms (as seen in Figure 8.1) while utilising domain knowledge based on shipping theory and empirical research to manually select the lags. Lag correlation had to overcome a significance level threshold in order to get approved. A lag of 1 month was chosen if the variable did not pass this threshold, as all variables had to be lagged to be used in forecasting. Explanatory variables were considered with a maximum lag of 6 months. This method of lag selection could have allowed for a greater limit than 6 months, which will be discussed in Chapter 9. The selected lags for each explanatory variable for each route can be seen in Table D.2.

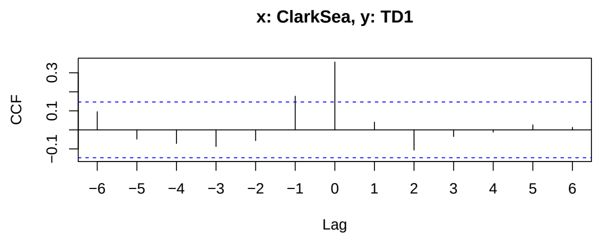


Figure 8.1: An example of a plot used to assist in the lag investigations.

Dealing with singular solutions

After performing an initial selection of variables based on our domain knowledge, and attempted to use these variables for modelling, we ran into an issue. We found that the selected variables did not provide a solution as a consequence of the design matrix not being invertible - the predictors were not independent, and their effects could not be uniquely identified. In other words, we needed to look for strongly correlated variables, and remove superfluous variables.

Checking the matrix rank: As an initial measure, we first constructed and ran a procedure based on the rankrank[54] of the matrix, in order to look for fully linearly dependent factors. The procedure excluded one variable for each iteration, and calculated the rank of the remaining matrix. This would allow us to identify any variable(s) that were fully linearly dependent. No such variables were found.

Issues can often occur when the matrix one is working with is «almost» not full. The problem is commonly referred to as multicollinearity. In our case, our likelihood model ran into trouble with singularities.

Review correlation between independent variables: The next step was to look into the correlation between the different variables. A calculation of the correlation matrix was done, and an accompanying visual representation added. The correlations were examined and potential variables were considered for removal. This includes the variable VLCC_{age}, among others, which provided irregular model results.

8.2 Candidate Predictor Analysis

Part one (i) of our research contribution (see Chapter 1) is finalized by the variable selection approach of stability selection with randomized LASSO. We will be looking for candidate variables with a high rank value[55] to validate variables that were initially included based on domain knowledge. Variables with a low score may be excluded, but we reserve the right to still include them if it seems viable from a theoretical stance. The low volatility state is expected to be the regime with the largest number of observations. Runs, hereby referred to as searches, were therefore made on differently sized subsamples - more precisely 30% and 5% of total training observations. We hypothesised that running based on these different fractions *might provide us with variables that are better at explaining each regime.*

In order to determine which variables to include in the domain knowledge and the parsimonious models, two search processes were performed:

i) **Domain knowledge model search:** These were searches which used the entire variable data set that we have gathered.

[54]This refers to the rank term from linear algebra.

[55]To avoid further confusion regarding rank and score of variables; *rank* value refers to the place a variable is ranking relative to other variables based on variables' individual *score* value from the stability selection.

The domain knowledge variable selection was based on a combination of stability selection and a theoretical foundation[56]. Stability selection was run for each route with all variables from the complete set available. The results were then aggregated, and the variables were ordered based on their (potential) performance ranking. In Table 8.1, we list the ranked results from the domain search on a route by route basis[57].

ii) **Parsimonious model search:** These were searches based on the subset of variables associated with each route, from the selection of variables in (i).

The best performing variables for each route were selected for the parsimonious regime models. We refer to the Appendix for a complete list of parsimonious ranking results (Table D.3).

Domain Search: Picking Domain Knowledge Variables

Based on the 30% stability search, Saudi Arabian crude export, followed by Chinese crude import provided the best fit for our two VLCC routes (TD1, TD3). Thus, a route-specific variable performed better than all the general variables. Among high ranked variables, we also find several VLCC supply factors, such as fixtures and newbuilding price. Furthermore, among more global variables, ClarkSea and VIX indices seem to be good indicators of VLCC freight rates. Medium good rankings belong to more general variables, such as world oil production and industrial production factors. We also find it interesting that more traditional variables, such as GDP, oil prices, certain stock indices, interest rates, Baltic indices, as well as capacity regulating supply factors do not perform particularly well.

On the last crude route, TD7, rankings of the best variables score closer to 1 than the two other crude routes. We do not observe the same top-one ranking of route-specific variables on TD7, although several import and export variables obtain close to 0.5 score, which would otherwise rank among the top on other routes. More variables are in general found to have a higher score on TD7 compared to other routes. Furthermore, it is encouraging to see that European consumer price index and European industrial production rank high, and significantly higher than, e.g., industrial production for some of the larger economies. TD7 is after all only concerned with European countries. In regards to supply, deadweight capacity and demolition price of Aframax all rank high. ClarkSea, world oil production, and VIX are the global factors ranking the highest. However, and again, several general factors are falling short.

When it comes to TD12, we observe a mix between the goodness of route-specific and general variables. The list of good-scoring candidates is shorter than crude routes. Chinese crude import is again the top-ranking variable, followed by four other general variables; ClarkSea, Panamax fixtures and deadweight capacity, and OECD industrial production. As with TD7, deadweight capacity and demolition price of the relevant fleet is ranking adequately high. Moreover, it is promising that some of the route-specific variables on a country and distillate level are scoring high. These include fuel oil export from Belgium and the US, as well as US refinery output of fuel oil. Additionally, factors such as US PADD3 fuel oil import and US fuel oil demand are performing medium well. A whole range of general variables are in fact beaten by these variables.

In regards to the clean routes, TC1 and TC2, the story with Chinese crude import and ClarkSea repeats itself. The two first high-ranked exchange rates are found on the TC1 route, namely US Dollar against the Saudi Arabian Riyal, and the Japanese Yen. Considering a route perspective, it is auspicious that these were performing compared to, e.g., pure Euro or Dollar indices. General-specific variables are more apparent on the TC1 route than the TC2 route, including, for instance, world oil production and a couple of supply variables. Furthermore, TC1 is the route with the fewest variables scoring sufficiently high. Scoring results are more interesting in a route perspective on TC2, as the second and third ranking variables are route-specific. These are also some of the more precise route-specific variables, as they involve both countries and type of oil trade; clean products export from Netherlands, and US clean products import from Netherlands. Similar to TD12, OECD industrial production ranks high on TC1 - hence it was included for both routes.

Based on the 5% search, route-specific variables are maybe even more noticeable among high ranked variables. Hypothetically, we find this to be in line with shipping theory and our own intuition, as high volatility states should be more irregular to underlying fundamentals. We observe that especially fixtures are ranking better in these searches. Fixture variables are intuitively more likely to exhibit inhomogeneous distributions as compared to other variables. This is due to charterers and shipowners potential for speedy responsiveness as opposed to, e.g., change in industrial

[56] This was done in conjunction with a rudimentary multicollinearity assessment by the evaluation of VIF, to limit issues stemming from this.

[57] For the sake of clarity, variables that were found to be associated with some sort of linear dependence issue are not excluded from Table 8.1, but are nevertheless ignored and not addressed further.

production. We believe some of the same reasoning is applicable to import and export factors as well, which also rank strong within these searches. Furthermore, the Brent oil forward curve ranks significantly better for most routes, including VLCC where its hypothetical significance probably is strongest (see Section5.3). We find this plausible, as psychology and future expectations may be more prominently reflected in a high volatility regime.

The fact that Chinese crude import is performing consistently well on all routes, is not surprising taking into consideration that China has been ranking among the top 5 crude importers during the relevant time horizon, ranking from 5th in 2002 to 1st today. We also note that the most consistent global indices are ClarkSea, world oil production and VIX. With these three variables, three key states of the freight market can be captured, namely the state of shipping markets in general, the flow of oil in the world, and market psychology, respectively. Oil prices, being a much-debated factor in all parts of the oil value chain, are consistently not scoring particularly strong on any of the routes. This is also true for several other economic variables, especially various stock indices, interest rates and exchange rates, as well as GDP. Surprisingly, 1 year-time charter rates also perform rather poor, except for MR vessels on TC2. Among the supply variables, we find deadweight capacities, in mdwt, to perform better than year-on-year capacity changes. Other supply variables, such as orderbook, demolitions and deliveries are not found to perform consistently well across our selection of routes. Vessel prices also occasionally score high.

As a sidenote, considering the historically high correlations between freight rates for tanker vessels, we were not surprised to see some of the same variables ranking consistently high. Fewer variables, however, seemed to be able to perform strong on products routes contrary to crude routes. Finally, we find it very encouraging that route-specific variables were able to rank very well across all routes despite the inclusion of many general variables.

Parsimonious Search: Picking Parsimonious Variables

In the parsimonious models, the best ranking variables from the second search are chosen. We have included a reasonable amount of high scoring variables. Results tend to be quite consistent with results from the domain search, while some deviation naturally exists as the selection process is based on simulation. For instance, rankings on TD1 are almost the same, while high-scoring variables on TD3 are now significantly shortened. Familiar variables from the domain search are though seen. At least, it is encouraging that two Japanese variables are included on TD3. On TD7 and TD12, top-ranking variables are also familiar from the domain search. On TC1, Aframax MEG - East fixtures is now at the top of the 30% search, as opposed to being at the top of the 5% search in the domain search. It is also noticeable that Japanese industrial production now ranks at the top in the 5% search. On TC2, a route-specific variable is ranking highest on the 30% search and second on the 5% search, namely products export from US to Netherlands. In fact, top 5 ranking variables on TC2 can be directly connected to this route. Moreover, the Brent forward curve seems to be a strong indicator also in the 5% search, ranking top 4 on TD12, TC1 and TC2. In general, we find route-specific variables to perform well in the parsimonious searches as well.

We believe that results from the parsimonious search are likely to have better validity and predictive power than results from the domain search, as the search is run without variables causing any misspecification. In subsequent parts of this chapter, statistical tests will reveal the predictive power of the various models.

8.3 Modelling

In the remainder of this chapter, we move on to part two (ii) of our research contribution (see Chapter 1), namely one-month prediction models of our 6 selected tanker routes. Results from the LASSO stability selection described above is incorporated.

Chow test

Before the model was developed, the Chow test for structural breaks was performed. The time series were evaluated and noteworthy periods were identified. These are seen in Table 8.2. The presence of structural breaks was assessed with the Chow test, by the estimation of regression models. We refer to plots of the time series' in Figure 6.1 when we comment on the individual routes below.

TD1 From surveying the plot of monthly freight rate movements, it appears to behave mostly the same. A couple of time periods appeared noteworthy. They were both tested for structural breaks, and the period from mid 2005 to late 2005 allowed us to reject the null hypothesis of no structural breaks.

TD3 The plot shows ordinary movements in most of the time span. The time series did, however appear to have noteworthy movements in the range from 2011 to 2014. A significant structural break was identified in the period from early 2012 to mid 2014.

TD7 For TD7, a longer period of a significant structural break was discovered - starting from mid 2009 and lasting to mid 2014. We rejected the null hypothesis with a 5% significance. The period can possibly be subdivided into several structural breaks with intermittent normal periods.

TD12 By visual inspection of the plot, one can see many periods spanning several months with higher changes than normal. Several structural breaks were identified, with a 0.01 significance. Most notable are the periods from mid 2009 to mid 2011 and late 2013 to early 2014.

TC1 The initial impression of this series was that it mostly behaves the same throughout the whole period. This impression is further strengthened by no rejections of the null hypothesis by the Chow test. There were no clear structural breaks to be easily identified.

TC2 The time series appears to follow the same patterns across most of the time period. A few unstable periods were identified around late 2013 and late 2014. We were able to reject the null hypothesis of no structural breaks for observations in this time span.

Abouarghoub et al. (2012) found noteworthy breaks for the tanker market. Particularly, they found the pre boom-cycle from 1990 to 2000 to be better captured by a two state regime model, and the period after 2000 to be suited to fit an even more volatile structure. As our data range was limited, beginning in mid 2002, due to the inclusion of explanatory variables, we were unable to capture similar results. However, Abouarghoub et al. (2012) did not explore the post-boom structure explicitly, and recommended this for future research. From our findings, some noteworthy changes in movements were determined for all series, except TC1. The results can be seen in Table 8.2. TD3 and TD7 had clearly visible structural changes. The others had certain periods which were identifiable.

Model Development

The purpose of the regime-switching models is to forecast tanker freight rates for different routes and vessel sizes. By combining the driving determinants of supply and demand that affect the spot freight rates, we established a prediction model for each route. The determinants were initially picked based on a theoretical foundation, and then validated as candidate variables with the use of a regularization (randomized LASSO) ranking approach (see Table 8.1). The models were trained based on the time period spanning from August 2002 to July 2014, while leaving the remaining period until June 2017 for validation of the out-of-sample, one-step-ahead forecast.

In total, 12 models were created. One domain knowledge model and one parsimonious model, for each of the 6 tanker routes. The fitted models were Markov regime-switching models consisting of different multiple regression models for each regime. Every model was fitted to two regimes, $k = 2$, and all variables were assumed to switch between these regimes. Domain knowledge models and parsimonious models are hereinafter referred to as TD1 and TD1_{pars} etc., respectively.

Resulting models are presented in Tables 8.5-8.10.

Regimes

In Table 8.3, we have provided an overview of regime properties for both domain and parsimonious models. In Figure 8.2, two distinctive regimes can be visually seen for the in-sample period for each route.

Remembering back to Section 6.1, we described three *stylized facts* of shipping freight rates, namely mean reversion, seasonality, and distribution and jumps. We will comment on the former and the latter here, while seasonality is being discussed later in Section 8.3.

At first, we note that the transition probability of switching from the low volatile regime to the high volatility regime is consistently lower than the probability of switching the opposite direction. This is true across all models, and in line with theory. From Table 8.3, we observe that the duration of high volatility regimes is lower than the duration of low volatility regimes. Furthermore, results indicate that low-volatility regimes have a larger weight of total observations than high-volatility regimes. This may suggest two things. First, freight rate levels are less sustainable at high levels, and thus reverts to its mean faster. Second, short-term momentum is less persistent in high volatility regimes. Ideally, the mean reversion property should portray a negative and a positive conditional mean for low levels of the freight rate and high levels of the freight rate, respectively. Adland (2003), In his PHD thesis, finds that non-linearity in the conditional mean is only statistical significant in the extremes of the freight rate distribution.

Shipping theory in Chapter 3 revealed the characteristics of the shipping supply curve, and suggested that freight rates exhibit higher volatility in the inelastic part of the curve. Adland (2003) argue and empirically proves that the volatility of TCE spot freight rates is an increasing function of the freight rate level. He further points out that freight rate volatility is non-zero in its empirical range except in its boundaries. Though in reality, freight rates do not reach these boundaries, hence making the true non-linear conditional variance curve impossible to fully replicate empirically. Yet, and as Adland (2003) argue, level effects in the conditional variance is an apparent maritime economic fact. Referring to Table 8.3, we clearly observe that this holds across all our models as well, by comparing volatilities as well as returns of low and high volatility sates against average freight rate levels in these states. Consequently, we confirm that freight rate changes are of higher magnitude and extremes to occur at higher frequency in high volatility regimes.

Additionally, Adland (2003) argues that a comprehensive model should incorporate residual lag effects in the conditional variance as well. We argue that we capture this somewhat indirectly by allowing volatility regimes to switch based on a given probability distribution. Furthermore, Adland empirically proves that the conditional variance, and the resulting magnitude exhibits a rich variation across bulk shipping sectors and vessel sizes (both dry bulk and tankers). He finds level effects in the conditional variance to be more evident across larger vessels sizes. This conforms somewhat with our findings in Table 8.3, whereof larger vessel sizes rank high on volatility in both states. At least, there is a clear indication that dirty routes are more volatile than clean routes. This is observable in both volatility regimes; however, more apparent for higher volatilities.

All in all, considering regime properties from Table 8.3, we are able to acknowledge empirical results from existing Markov regime-switching tanker literature, which we presented in the Literature Review (Chapter 2). In general, these studies were consistent in finding regime properties equal to those we have presented here (see, for instance, Abouarghoub et al. (2014); IAME (2014); Abouarghoub et al. (2012); Alizadeh and Nomikos (2011); Abouarghoub and Mariscal (2011)).

Parameter Evaluation

In the following, we address and interpret the economic impact of the selected variables from the LASSO process based on our hypothetical discussion from Chapter 6. Moreover, our interpretation will be mainly focused towards the regime with the largest share of total observations, i.e. what presumably is the low-volatility («normal») regime (see Table 8.3). All of the resulting coefficients from our domain and parsimonious models can be seen from Tables 8.5-8.10, and Tables 8.15-8.20, respectively.

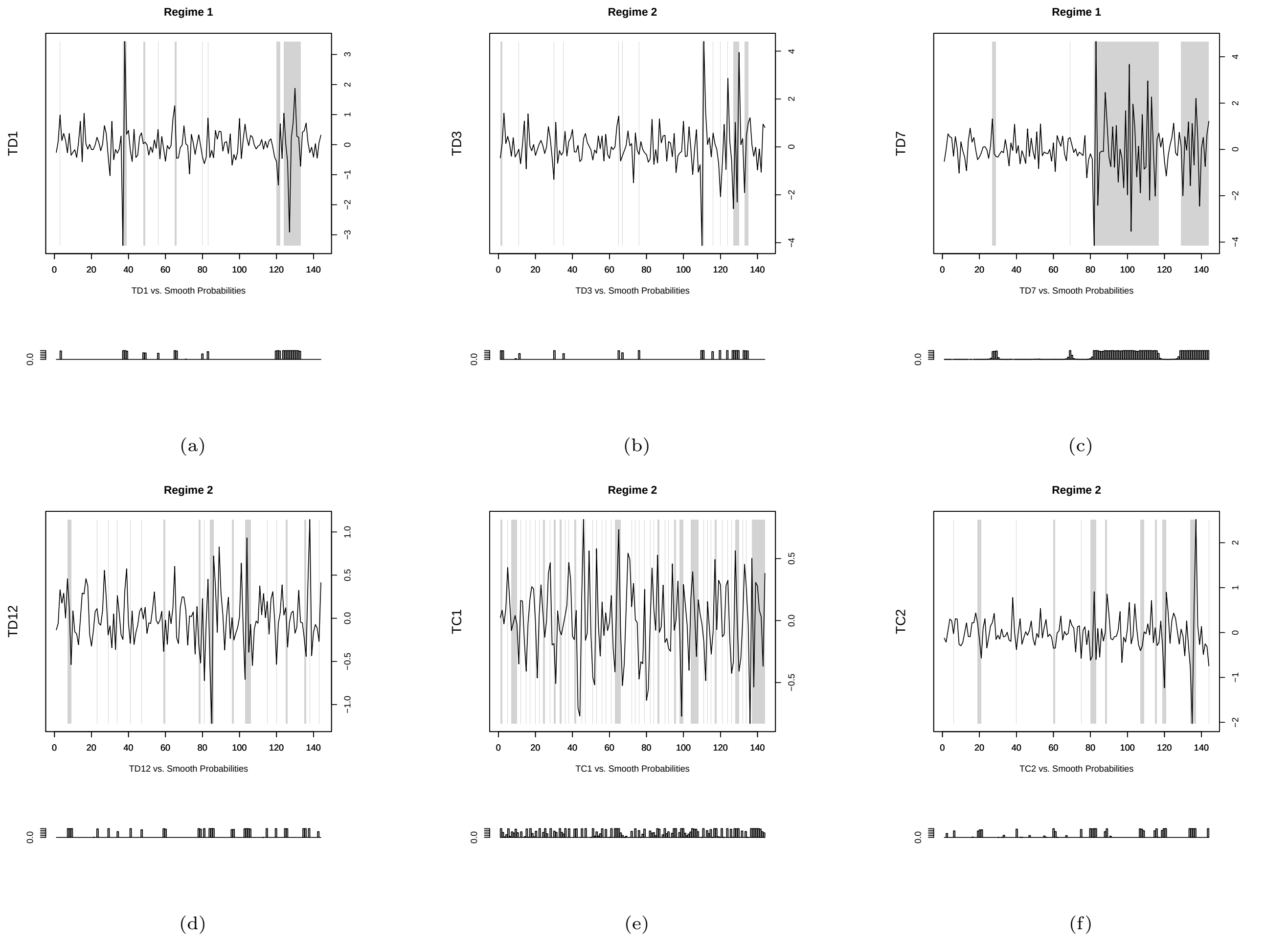
General observations of the domain models reveal that coefficient signs are very mixed from what we hypothesised in Chapter 5. This is in line with findings of Ringheim and Stenslet (2017), who also obtain a very mixed directional impact based on their initial hypotheses. In Table 8.4, a complete overview of our directional correctness regards to our hypotheses is given for each route model. Comparing seasonal dummies against each route's theoretical seasonal components from Appendix A.4 revels that seasonal dummies have a higher hypothetical coefficient correctness than determinant coefficients. Somewhat similar conclusions are drawn by Ringheim and Stenslet (2017), who find pure seasonal models to outperform factor-driven models for the BDTI index.

Remembering that our initial hypothesis was based on real-time lag effects (0-lag), we acknowledge that some of the hypothetical impact could differ based on variables' lag structure[58] (see Table D.2). We identified supply variables to possibly have the most interesting lag structures among our variable groups, due to their endogenous effect on shipping cycles. We do, however, have the same directional hypotheses for most of the demand and economic variables even when we are considering the lags chosen, although several arguments may exist to support changing these hypotheses after all. For instance, several import and export variables included in the models have surprisingly a negative coefficient based on a 1-month lag. "What goes up, comes down", i.e. if a lot of oil was transported last month, less oil could be transported in the current month. However, oil trade is likely to exhibit a non-stationary process, so we argue that a trend in increases/decreases should exist. Equally, several oil demand variables are found to yield a negative impact on freight rates in the domain models.

Based on Table 8.4, we observe that the accuracy of our hypotheses for the parsimonious models are generally quite mixed as well. Only *one* regime in *one* model conforms 100% with our initial hypotheses; the high volatility regime in $TD7_{pars}$. Seasonal dummies although seem to be equally accurate regards to our hypotheses as the domain models. Regarding some of the variables that we hypothesised natural, such as the WTI oil price and various exchange rates, we observe their directional impact to be mixed across the models they are included in.

Chinese crude import yields a negative impact on freight rates on both $TD1_{pars}$ and $TD12\ _{pars}$, while having a positive impact on $TD7_{pars}$. Nonetheless, we find a negative impact not to be completely off what we find rational, despite initially hypothesising a positive impact. The development in VLCC fixtures from MEG to China by Chinese shipowners deviates from freight rate changes on MEG – Far East routes, as opposed to fixtures from MEG to Far East by non-Chinese shipowners, which move almost in parallel (Fearnley,

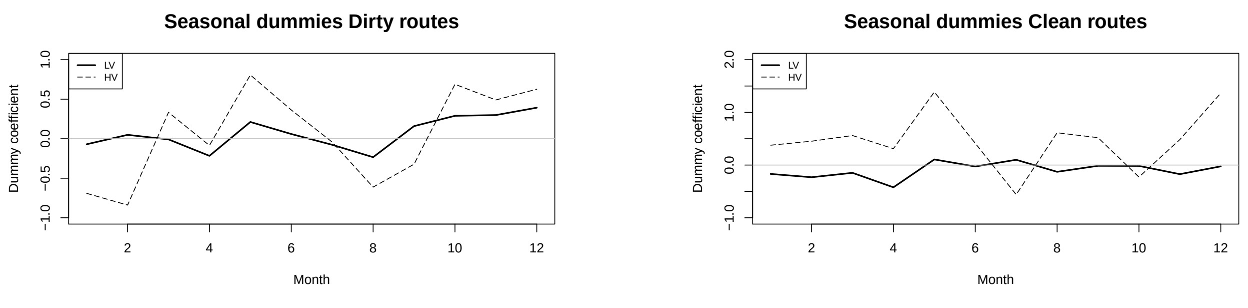
[58]Several data formats exist for the data collected; variables that are updated, e.g., in the start of month are as such in reality lagged by about 2 months to the freight rate if a 1-month lag is used. A complete overview of data formats can be seen in Appendix A.1.



(a) (b) (c)

(d) (e) (f)

Figure 8.2: Indication of which observations are associated with a possibly high volatile regime (gray scaled), for the ***parsimonious regime models*** The x-axis represents the in-sample time horizon. The small bars below each plot indicate the probability of being in the high volatility state at each time step.



(a) Dirty routes (b) Clean routes

Figure 8.3: Representation of seasonal dummy coefficients based on the mean of seasonal dummy coefficients from the ***parsimonious regime models***. Dirty tanker (TD1, TD3, TD7, TD12) and Clean tanker (TC1, TC2) routes.

2018). Further, and different from our initial hypothesis, 1-month lagged VLCCs that are due in the MEG have a positive impact on $TD1_{pars}$. This may be likely, as VLLCs that were due previous month may have taken on new charters, and thus fewer VLLCs may be due in the 0-lagged month. ClarkSea and Saudi Arabian exports are as expected positively related to $TD1_{pars}$. On $TD3_{pars}$, only 6-month lagged Japanese money supply is hypothesised correct according to the model. However, this variable is also the only one that is significant in the low volatility regime. On $TD12_{pars}$, $TC1_{pars}$, and $TC2_{pars}$, we find the Brent forward curve to be negatively related to freight rates. As none of the vessels on these routes are particularly used for storage, this impact may reflect that an increase in oil price have a negative impact on these routes, which the negative coefficient for WTI on $TD12_{pars}$ further may portray.[59]

More surprisingly is that there is lack of consistency in whether the low volatility regime has a better hypothetical coefficient impact than the high volatility state. We expected the low volatility regime to portray a closer to "normal state" condition. However, when the directional correctness is as limited as it is for some of the routes, this expectation is difficult prove righteously. Finally, we remark that some of our hypotheses certainly could be a wrong-pointers in terms of "real impact". However, that is exactly some of the "beauty" of statistics; we can reveal relationships we otherwise would not be able to prove/disprove with general knowledge.

Seasonal Dummies

Moving on to a closer assessment of the seasonal relationships; we saw from Figure 7.1 that seasonality was apparent in tanker freight rates. We restrict our interpretation of seasonal dummies to the parsimonious models. In order to get a better visual tool for inspection, we calculate the mean of the dummy coefficients for dirty and clean routes. Hence, we are able to compare these against the «theoretical» seasonal components from Figure 7.1. Resulting figures from this are given in Figure 8.3 (see Appendix A.5 for individual routes).

We are now ready to evaluate the last stylized fact from Section 6.1, namely *seasonality*. We observe from the resulting seasonal dummies that the seasonal magnitude are clearly different between the two regimes in our models (Figure 8.3). Furthermore, seasonal dummies are more stable for clean routes than for dirty routes, which is in line with the theoretical representation in Figure 7.1. It is also encouraging that the

[59] We leave further interpretation to the reader, as there exist endless of possible interpretations considering the large scale of variables; arguments from Chapter 5 may serve as a useful guide.

Table 8.1: Domain search. The ranked variables which are the basis for creating our domain knowledge regime models. This variable ordering was utilised to evaluate the benefit of including certain factors. They have been ranked by the method of stability selection with randomized LASSO. The fraction of included observations was tuned to 0.3 and 0.05 in order to look for variables that perform better in certain short time periods, i.e. we were attempting to narrow down variables for the high volatility regime. *Notes: Certain variables were not considered for use due to modelling conflicts, like* $VLCC_{age}$. *TD3 with a fraction of* 0.05 *did not provide any sufficient high-scoring results.*

TD1		TD3		TD7	
Regularization run: 0.3		Regularization run: 0.3		Regularization run: 0.3	
	score		score		score
SA_exp	0.581	SA_exp	0.578	Chi_imp	0.761
VLCC_age	0.555	VLCC_age	0.574	Eur_CPI	0.692
Chi_imp	0.507	Chi_imp	0.511	Ind_Eur	0.648
ClarkSea	0.446	ClarkSea	0.439	W_prod	0.612
VLCC_fix	0.438	VLCC_due	0.393	LIBOR_Yen	0.536
VLCC_due	0.429	VLCC_new	0.346	Afra_demo_price	0.503
VLCC_new	0.344	VLCC_price	0.346	Afra_mdwt	0.498
VIX	0.323	VIX	0.345	ClarkSea	0.490
VLCC_price	0.319	VLCC_fix	0.342	Eur_imp	0.458
W_prod	0.310	VLCC_fix_east	0.329	Afra_new	0.454
US_money	0.277	Yen_USD	0.326	VIX	0.445
USD_SAR	0.273	W_prod	0.324	Afra_down	0.443
ClarkAve	0.261	Ind_Jap	0.285	Ind_India	0.431
SAPadd3_imp	0.240	ClarkAve	0.280	NS_exp	0.430
Ind_India	0.234	Jap_CPI	0.265	Ge_imp	0.416
US_imp	0.233	USD_SAR	0.264	NS_prod	0.410
VLCC_deliveries	0.231	Ind_India	0.253	Ind_China	0.409
VLCC_down	0.222	VLCC_deliveries	0.247	Afra_SP	0.389
VLCC_fix_west	0.216	Jap_imp	0.243	Afra_price	0.387
USD_index	0.206	Jap_dem	0.235	Afra_deliveries	0.381
LIBOR_Yen	0.198	VLCC_down	0.233	Ind_imp	0.374
NA_prod	0.191	Jap_money	0.231	Brent_forw	0.366
Regularization run: 0.05		Regularization run: 0.05		Regularization run: 0.05	
	score		score		score
Chi_imp	0.195			Chi_imp	0.210
ClarkSea	0.180			NS_prod	0.200
VLCC_age	0.165			NS_exp	0.191
Major_imp	0.160			Ge_imp	0.188
VLCC_down	0.155			Afra_down	0.188
Brent_forw	0.155			Eur_imp	0.186
US_money	0.145			Ind_Eur	0.185
SAPadd3_imp	0.145			Ind_India	0.182

TD12		TC1		TC2	
Regularization run: 0.3		Regularization run: 0.3		Regularization run: 0.3	
	score		score		score
Chi_imp	0.445	Chi_imp	0.421	Chi_imp	0.475
ClarkSea	0.344	ClarkSea	0.320	Ne_cexp	0.429
Pana_fix	0.320	LR2_new	0.262	NeUS_cimp	0.425
Ind_OECD	0.311	USD_SAR	0.255	ClarkSea	0.365
Pana_mdwt	0.301	Afra_age	0.237	MR_deliveries	0.343
Bel_fexp	0.282	W_prod	0.233	Ind_OECD	0.336
US_fexp	0.268	Yen_USD	0.210	MR_3tc	0.326
LIBOR_Yen	0.254	Afra_fix_us	0.206	MR_1tc	0.269
US_fout	0.254	Afra_demo_price	0.203	USNe_exp	0.255
W_prod	0.238	LR2_order	0.203	ClarkAve	0.249
Pana_price	0.232	ClarkAve	0.185	W_prod	0.231
Pana_demo_price	0.224	LIBOR_Yen	0.178	Ne_cout	0.229
Padd3_fimp	0.221	VIX	0.173	MR_demo_price	0.220
US_fdem	0.213	Afra_down	0.167	MR_yy	0.216
ClarkAve	0.202	LR2_order_fleet	0.156	MR_fix_US	0.209
Pana_fix_US	0.193			NePadd3_cimp	0.202
VIX	0.179			LIBOR_Yen	0.195
Bel_fout	0.177			US_money	0.182
Pana_yy	0.173			Ind_India	0.178
US_dem	0.169			VIX	0.178
US_money	0.159			MR_down	0.174
Pana_deliveries	0.157			US_dem	0.173
Regularization run: 0.05		Regularization run: 0.05		Regularization run: 0.05	
	score		score		score
Pana_fix	0.153	Afra_fix_east	0.170	Ne_cout	0.160
Chi_imp	0.151	Chi_imp	0.145	US_cdem	0.155
US_fout	0.150	Afra_down	0.145	Chi_imp	0.145
US_fexp	0.143	Afra_fix_us	0.140	MR_fix_US	0.140
US_fimp	0.131	Major_imp	0.135	MR_deliveries	0.140
Pana_mdwt	0.129	Afra_fix	0.125	NeUS_cimp	0.135
LIBOR_Yen	0.124	Jap_money	0.120	Brent_forw	0.135
BelUS_fimp	0.122	Brent_forw	0.120	USNe_exp	0.125

Table 8.2: The notable results from the Chow test for structural breaks, with a null hypothesis of no structural breaks. The test was performed for different routes, on notable movements in the data. Included are the time periods that were compared (break dates are considered volatile time periods, while normal dates are a comparable regular time period), the f-value from the test and it associated p-value.

	Break dates	Normal dates	f-value	p-value
TD1	2005-06/2005-10	2006-01/2007-06	20.506	0.000
TD1	2012-08/2013-10	2011-03/2012-07	0.214	1.554
TD3	2012-03/2014-06	2009-01/2012-02	2.385	0.050
TD7	2009-07/2014-04	2004-01/2009-04	2.492	0.035
TD12	2009-07/2011-04	2004-01/2009-07	3.365	0.008
TD12	2013-10/2014-03	2011-06/2013-09	4.401	0.006
TC1	2013-10/2014-03	2011-06/2013-09	1.674	0.179
TC2	2013-10/2014-03	2011-06/2013-09	5.290	0.002
TC2	2014-07/2014-12	2015-01/2016-06	4.860	0.009

Table 8.3: This table contains key information regarding the regimes from all of the route models. It consists of the probabilities of transitioning from one regime to the other, the number of observations, volatility, and much more.

	Domain knowledge models						Parsimonious models					
	TD1	TD3	TD7	TD12	TC1	TC2	TD1	TD3	TD7	TD12	TC1	TC2
Transition π_{12}	0.761	0.258	0.638	0.584	0.392	0.663	0.358	0.096	0.079	0.149	0.579	0.102
Transition π_{21}	0.366	0.479	0.299	0.530	0.449	0.467	0.066	0.584	0.041	0.597	0.593	0.503
LV Regime	**2**	**1**	**2**	**2**	**1**	**2**	**2**	**1**	**2**	**1**	**1**	**1**
LV Observations	93	93	96	78	76	84	120	124	88	114	75	117
HV Observations	51	51	48	66	68	60	24	20	56	30	69	27
LV, Volatility	0.409	0.554	0.824	0.289	0.295	0.362	0.379	0.588	0.500	0.249	0.297	0.266
HV, Volatility	0.961	1.416	1.547	0.360	0.346	0.515	1.388	2.135	1.679	0.500	0.340	0.826
LV Weight	0.68	0.65	0.68	0.52	0.53	0.58	0.85	0.87	0.60	0.80	0.52	0.82
Avg. LV Duration	3.55	4.63	4.25	3.29	3.61	3.53	13.22	10.25	22.00	6.94	3.07	9.75
Avg. LV Return	0.02	0.05	0.00	-0.01	0.00	-0.03	-0.01	0.01	-0.01	0.03	0.03	0.02
HV Weight	0.32	0.35	0.32	0.48	0.47	0.42	0.16	0.13	0.40	0.20	0.48	0.18
Avg. HV Duration	2.17	3.00	2.50	2.59	3.86	2.20	4.00	2.75	18.33	2.50	2.26	2.88
Avg. HV Return	-0.020	-0.081	0.019	0.020	-0.004	0.008	0.098	-0.003	0.036	-0.126	-0.036	-0.155
Avg. LV Rate	43,197	46,799	29,018	25,018	24,486	18,379	43,323	46,073	35,193	25,621	23,950	18,886
Max LV Rate	13,8758	222,142	104,411	76,729	71,255	39,949	138,758	222,142	104,411	76,729	71,255	42,457
Min LV Rate	-3,394	-909	-103,409	4,816	6,676	-3,891	-3,394	-100,510	-16,460	4,816	4,613	3,530
Avg. HV Rate	33,520	32,595	13,621	23,582	23,834	15,766	22,005	15,081	6,119	19,566	24,426	10,373
Max HV Rate	100,635	201,359	91,539	63,373	69,192	42,457	100,635	157,770	91,539	63,373	68,340	40,745
Min HV Rate	-81,723	-100,510	-87,297	-1,760	3,066	-1,794	-81,723	-53,981	-103,409	-1,760	3,066	-3,891

π_{ij}: The transition probability of switching from state i to state j
LV: Low volatility state
HV: High volatility state

Table 8.4: Evaluation of hypothetical coefficient impact. Domain knowledge models and parsimonious models. Percentage correctness of variable coefficients are based on the selection of variables that had an initial hypothesis, i.e. excluding neutral hypothesised variables. Seasonal dummies include the intercept, serving as a seasonal measure for the month of July.

«(*)» next to percentages indicates which regime that has the largest number of observations, i.e. the low volatility regime.

	Percentage correct hypothetical sign of			
	Variable coefficients		**Seasonal dummies**	
	Regime 1	Regime 2	Regime 1	Regime 2
TD1	75%	60%$^{(*)}$	67%	83%$^{(*)}$
TD3	48%$^{(*)}$	48%	75%$^{(*)}$	67%
TD7	70%	39%$^{(*)}$	75%	50%$^{(*)}$
TD12	45%	52%$^{(*)}$	75%	67%$^{(*)}$
TC1	50%$^{(*)}$	39%	75%$^{(*)}$	58%
TC2	41%	31%$^{(*)}$	50%	75%$^{(*)}$
$TD1_{pars}$	60%	40%$^{(*)}$	50%	92%$^{(*)}$
$TD3_{pars}$	25%$^{(*)}$	75%	75%$^{(*)}$	83%
$TD7_{pars}$	100%	67%$^{(*)}$	67%	67%$^{(*)}$
$TD12_{pars}$	50%$^{(*)}$	50%	67%$^{(*)}$	58%
$TC1_{pars}$	55%$^{(*)}$	63%	75%$^{(*)}$	67%
$TC2_{pars}$	29%$^{(*)}$	43%	58%$^{(*)}$	67%

Table 8.5: Estimated coefficients for the TD1 regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	-0.087	-2.084	*	0.088	1.359	
AG_exp	2.392	3.773	***	-3.054	-3.349	***
VLCC_fix_west	-0.335	-6.482	***	0.294	6.102	***
VLCC_due	0.739	15.950	***	0.153	4.129	***
SAPadd3_imp	0.019	0.178		0.042	0.343	
NA_prod	33.738	29.221	***	-4.773	-4.922	***
SA_exp	4.282	12.710	***	2.035	4.444	***
USD_SAR	-61.426	-4.360	***	29.885	1.846	.
US_money	-25.252	-14.222	***	1.362	0.904	
Tadawul	1.467	7.871	***	0.053	0.220	
VLCC_price	0.032	3.323	***	0.027	3.649	***
Major_imp	0.739	1.784		-1.263	-2.226	*
Chi_imp	-1.038	-8.946	***	-0.628	-5.569	***
Ind_imp	0.899	6.537	***	-0.507	-2.460	*
W_prod	56.015	24.341	***	2.144	0.690	
SDR_USD	-5.302	-7.012	***	-1.845	-1.679	.
Ind_China	1.173	13.484	***	-0.177	-1.956	.
Ind_US	-0.010	-0.665		0.029	1.400	
Ind_India	0.053	9.017	***	0.034	6.491	***
LIBOR	-1.512	-8.770	***	-1.503	-7.961	***
Brent_forw	0.042	3.150	**	-0.022	-1.192	
WTI	1.466	6.935	***	1.424	6.483	***
ClarkSea	1.998	13.960	***	0.233	1.442	
VIX	2.307	22.617	***	0.380	3.291	***
Aug	-1.490	-25.470	***	-0.671	-5.962	***
Sep	0.076	1.026		0.041	0.466	
Oct	1.026	18.522	***	0.013	0.135	
Nov	-0.644	-9.633	***	0.343	3.599	***
Dec	0.930	14.220	***	0.221	2.524	*
Jan	0.693	7.747	***	-0.489	-5.113	***
Feb	-2.442	-24.670	***	-0.062	-0.639	
Mar	-0.350	-4.841	***	0.203	2.203	*
Apr	1.212	18.882	***	-0.455	-5.023	***
May	1.353	20.463	***	-0.083	-0.961	
Jun	0.447	6.249	***	-0.153	-1.714	.
R-squared	0.9959			0.8760		
R-squared-adj	0.9843			0.8079		
Approx. num.	51			93		

Significance 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8.6: Estimated coefficients for the TD3 regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	-0.025	-0.209		-0.181	-3.325	***
VLCC_SP	-2.654	-4.437	***	10.474	36.419	***
ClarkSea	1.971	6.650	***	-0.268	-2.066	*
Chi_imp	-0.203	-0.955		3.538	32.726	***
Jap_dem	-0.304	-0.365		-12.985	-31.548	***
VLCC_fix_east	-0.768	-5.047	***	0.024	0.335	
VLCC_fix_jap	0.021	2.972	**	0.048	10.413	***
VLCC_due	0.437	6.207	***	0.460	12.272	***
Jap_imp	2.364	4.700	***	-6.008	-24.551	***
SA_exp	1.396	2.498	*	-2.447	-4.544	***
Jap_money	13.519	3.909	***	6.813	5.069	***
USD_SAR	-117.175	-2.266	*	-200.194	-16.710	***
Yen_USD	3.549	2.375	*	8.157	10.906	***
Jap_CPI	-39.653	-3.895	***	102.439	14.235	***
Ind_Jap	-0.055	-5.170	***	0.145	45.250	***
Tadawul	0.265	0.605		-1.208	-4.228	***
Nikkei	-0.572	-0.730		4.854	15.509	***
VLCC_yy	-1.473	-0.473		24.908	15.067	***
VLCC_price	0.034	2.603	**	0.057	5.788	***
Major_imp	-5.967	-5.205	***	1.551	3.292	***
Ind_imp	-1.060	-3.245	**	-0.011	-0.058	
W_prod	15.060	3.912	***	22.549	9.328	***
Ind_China	-0.621	-3.789	***	-0.627	-6.017	***
Ind_US	0.175	5.126	***	-0.629	-44.324	***
Ind_India	0.039	4.021	***	-0.044	-6.441	***
LIBOR_Yen	0.161	0.250		-9.587	-26.601	***
Brent_forw	-0.067	-2.227	*	-0.360	-23.213	***
WTI	0.427	1.086		-3.147	-18.664	***
VIX	-0.384	-1.864	.	3.773	44.126	***
Aug	0.069	0.416		-1.164	-16.415	***
Sep	0.087	0.575		-1.118	-12.192	***
Oct	0.135	0.757		1.080	13.865	***
Nov	0.162	0.985		1.774	22.573	***
Dec	0.190	1.237		3.031	29.344	***
Jan	-0.610	-3.748	***	1.294	10.852	***
Feb	0.152	0.807		-1.232	-16.085	***
Mar	-0.129	-0.732		1.734	19.369	***
Apr	-0.536	-2.754	**	-0.976	-13.063	***
May	0.360	1.854	.	-0.379	-3.917	***
Jun	0.435	2.158	*	-0.968	-11.621	***
R-squared	0.8166			0.9978		
R-squared-adj	0.6840			0.9892		
Approx. num.	93			51		

Significance 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8.7: Estimated coefficients for the TD7 regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	-1.735	-21.604	***	0.776	3.961	***
Afra_demo_price	7.310	24.416	***	-0.021	-0.057	
Afra_price	-28.322	-39.997	***	4.828	2.358	*
Eur_dem	8.998	16.301	***	2.349	1.892	.
NS_exp	-3.639	-23.957	***	3.013	7.830	***
Afra_fix_sum	0.087	1.374		-0.056	-0.362	
Eur_imp	5.338	32.953	***	-1.010	-1.885	.
Ge_imp	4.513	21.278	***	-1.966	-2.847	**
NS_prod	-22.028	-47.721	***	4.323	3.963	***
USD_Pound	-1.445	-1.362		-12.353	-6.200	***
Eur_CPI	-134.752	-20.670	***	43.073	1.998	*
Ind_Eur	0.340	32.673	***	-0.006	-0.188	
Afra_mdwt	-0.892	-25.861	***	-0.619	-5.034	***
Afra_yy	0.314	30.495	***	-0.552	-5.393	***
Afra_SP	7.164	28.082	***	-0.789	-0.785	
Chi_imp	3.564	21.405	***	-0.025	-0.101	
Ind_imp	8.488	59.274	***	1.190	2.676	**
W_prod	24.104	14.858	***	-1.436	-0.233	
Ind_China	2.068	20.349	***	-0.469	-2.089	*
Ind_US	0.415	31.165	***	0.036	0.898	
Ind_India	0.103	28.528	***	-0.008	-0.598	
LIBOR	-3.230	-18.322	***	1.377	2.997	**
Brent_forw	-0.365	-28.984	***	-0.187	-5.011	***
WTI	2.224	12.093	***	1.076	2.052	*
ClarkSea	-2.884	-23.580	***	0.513	1.368	
VIX	-1.454	-13.256	***	-0.307	-1.473	
Aug	3.264	26.641	***	-1.209	-3.912	***
Sep	1.972	14.716	***	-0.934	-2.689	**
Oct	1.186	11.996	***	0.172	0.576	
Nov	5.376	48.962	***	-1.699	-5.883	***
Dec	4.971	40.347	***	-0.216	-0.786	
Jan	-1.639	-17.869	***	-1.048	-5.006	***
Feb	1.608	16.963	***	-1.103	-4.229	***
Mar	2.261	20.384	***	-1.140	-4.286	***
Apr	3.887	38.295	***	-1.616	-5.353	***
May	-1.573	-19.866	***	-0.716	-2.693	**
Jun	2.224	20.185	***	-1.080	-3.780	***
R-squared	0.9989			0.8543		
R-squared-adj	0.9946			0.7683		
Approx. num.	48			96		

Significance 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 8.8: Estimated coefficients for the TD12 regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	-0.231	-3.412	***	0.263	11.489	***
US_dem	-2.685	-2.401	*	0.184	0.589	
US_fdem	0.306	3.704	***	-0.522	-20.785	***
Eur_fdem	-1.056	-4.453	***	-0.396	-4.098	***
Bel_fexp	-0.566	-7.314	***	0.306	16.732	***
US_fexp	-0.205	-2.223	*	-0.230	-9.948	***
USBel_fexp	2.806	2.943	**	-3.685	-10.200	***
Pana_fix_US	-0.003	-0.430		-0.033	-14.348	***
Pana_fix	0.355	5.693	***	0.057	2.990	**
Bel_fimp	-0.021	-0.230		-0.598	-21.522	***
Padd3_fimp	1.835	3.723	***	-0.513	-2.797	**
US_fout	-0.577	-2.723	**	0.059	0.797	
Bel_fout	-0.029	-0.328		-0.248	-11.067	***
BelPadd3_fimp	0.064	2.374	*	-0.037	-4.089	***
US_money	-10.230	-7.217	***	4.823	7.261	***
US_CPI	4.819	0.725		-11.734	-5.908	***
Ind_OECD	-0.027	-2.068	*	0.000	0.037	
Pana_yy	-0.028	-2.000	*	-0.060	-13.022	***
Pana_order	-0.584	-2.552	*	1.023	12.191	***
Pana_price	0.084	4.704	***	0.087	13.197	***
Pana_demo_price	-0.423	-2.281	*	-0.059	-0.633	
Major_imp	-2.185	-4.363	***	2.305	13.506	***
Chi_imp	0.143	1.354		-0.283	-7.547	***
Ind_imp	-0.513	-2.884	**	0.583	10.664	***
OPEC_prod	3.469	2.974	**	0.726	1.600	
SDR_USD	5.333	6.154	***	2.788	8.995	***
Ind_China	0.004	0.063		-0.461	-13.335	***
Ind_US	0.111	6.721	***	0.018	2.382	*
Ind_India	-0.016	-2.613	**	0.018	11.188	***
LIBOR	-0.156	-1.046		1.110	21.773	***
Brent_forw	-0.084	-4.982	***	-0.046	-9.830	***
WTI	0.182	0.876		0.971	11.977	***
ClarkSea	0.614	3.203	**	0.119	2.369	*
VIX	0.120	1.052		0.472	15.135	***
Aug	-0.379	-4.312	***	-0.407	-13.055	***
Sep	0.255	2.955	**	0.254	10.496	***
Oct	0.553	5.859	***	-0.104	-4.357	***
Nov	0.081	0.970		0.358	13.207	***
Dec	0.592	6.118	***	0.227	8.365	***
Jan	0.350	3.715	***	0.222	6.204	***
Feb	-0.605	-5.379	***	0.130	4.137	***
Mar	0.587	6.789	***	-0.222	-7.235	***
Apr	0.505	4.695	***	-0.635	-19.187	***
May	0.195	2.431	*	-0.122	-4.303	***
Jun	0.114	1.398		-0.237	-7.514	***
R-squared	0.9436			0.9876		
R-squared-adj	0.8402			0.9693		
Approx. num.	66			78		

Significance 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 8.9: Estimated coefficients for the TC1 regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	0.095	2.551	*	-0.101	-3.401	***
SA_cexp	0.170	6.300	***	-0.097	-3.123	**
Afra_fix_east	0.321	15.160	***	0.110	4.765	***
Afra_fix_us	0.036	3.140	**	0.195	13.921	***
Afra_fix	0.018	0.220		-0.559	-5.452	***
Jap_cimp	-0.029	-1.193		-0.320	-11.150	***
SA_prod	-0.841	-2.880	**	-0.868	-2.814	**
SA_cout	-0.174	-1.410		-0.502	-4.048	***
Jap_money	0.320	0.104		17.597	6.624	***
USD_SAR	27.550	3.839	***	8.800	0.719	
Yen_USD	6.475	16.467	***	1.285	2.593	**
Jap_CPI	24.332	8.323	***	-29.048	-6.716	***
Ind_Jap	0.002	0.750		0.024	14.000	***
Tadawul	-1.215	-9.067	***	-0.321	-2.587	**
Nikkei	0.513	2.679	**	-1.575	-7.733	***
LR2_mdwt	-0.229	-4.693	***	-0.403	-8.681	***
LR2_order	-1.928	19.418	***	0.534	5.604	***
Afra_price	3.366	5.563	***	-0.853	-2.487	*
Afra_SP	-0.348	-1.668	.	-1.166	-6.112	***
Afra_demo_price	-0.252	-2.047	*	-1.323	-10.799	***
Major_imp	-0.339	-1.384		-0.726	-2.651	**
Chi_imp	-0.164	-2.570	*	-0.049	-0.919	
W_prod	10.654	7.923	***	9.123	7.340	***
SDR_USD	3.750	10.526	***	9.647	14.064	***
Ind_China	0.175	4.214	***	0.037	0.818	
Ind_US	0.059	6.556	***	0.003	0.282	
Ind_India	-0.007	-2.063	*	-0.004	-1.464	
LIBOR_Yen	1.237	7.967	***	1.299	5.613	***
Bunker_Jap	-0.999	-7.228	***	1.042	7.442	***
Brent_forw	-0.009	-1.084		-0.011	-1.217	
WTI	-0.358	-3.508	***	-0.453	-4.473	***
ClarkSea	0.785	9.665	***	0.423	4.559	***
VIX	0.372	5.860	***	-0.146	-2.664	**
Aug	-0.330	-5.482	***	0.260	7.552	***
Sep	0.018	0.410		0.298	6.795	***
Oct	-0.191	-4.456	***	-0.442	-11.030	***
Nov	-0.071	-1.449		-0.545	-14.182	***
Dec	0.128	2.461	*	-0.024	-0.485	
Jan	-0.140	-2.820	**	-0.213	-4.949	***
Feb	-0.317	-5.776	***	0.110	2.536	*
Mar	-0.273	-5.124	***	0.235	5.047	***
Apr	-0.229	-5.105	***	-0.020	-0.460	
May	-0.209	-3.835	***	0.296	7.746	***
Jun	-0.016	-0.356		-0.523	-9.314	***
R-squared	0.9666			0.9803		
R-squared-adj	0.9230			0.9434		
Approx. num.	76			68		

Significance 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8.10: Estimated coefficients for the TC2 regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	0.592	10.121	***	0.121	2.484	*
Eur_cdem	-5.447	-15.425	***	-3.346	-10.287	***
Ne_cexp	-0.117	-0.978		-0.219	-1.752	.
US_cexp	0.459	7.020	***	-0.201	-2.539	*
USNe_exp	1.752	8.914	***	-2.345	-14.404	***
MR_fix_US	0.048	1.453		0.351	8.244	***
Ne_cimp	-0.869	-8.946	***	-0.064	-0.830	
US_cimp	0.451	3.295	***	-0.236	-2.514	*
US_prod	2.314	7.016	***	-0.259	-0.531	
Padd1_refuti	-1.574	-11.781	***	-1.369	-8.266	***
Ne_cout	-0.856	-8.127	***	-1.318	-8.473	***
US_cout	-2.587	-3.669	***	0.690	1.511	
NeUS_cimp	0.174	6.736	***	-0.172	-8.677	***
NePadd3_cimp	-5.853	-18.518	***	-0.995	-3.954	***
US_money	-0.711	-0.420		-1.455	-1.547	
US_CPI	-28.646	-7.073	***	27.112	6.455	***
Ind_OECD	-0.052	-4.522	***	-0.018	-2.815	**
MR_3tc	3.431	9.568	***	-0.010	-0.021	
MR_yy	-0.216	-8.143	***	0.096	5.112	***
MR_order	1.245	3.260	**	-2.821	-8.011	***
MR_demo_price	0.701	5.040	***	0.514	3.482	***
MR_price	-5.318	-7.088	***	3.458	4.870	***
MR_SP	-0.125	-0.441		-1.705	-10.113	***
Major_imp	-4.387	-11.003	***	1.898	6.413	***
Chi_imp	-0.110	-1.599		0.215	3.226	**
Ind_imp	-0.320	-2.492	*	1.069	9.186	***
OPEC_prod	-8.965	-11.236	***	5.051	6.506	***
SDR_USD	-5.988	-9.870	***	2.000	3.771	***
Ind_China	-0.518	-8.683	***	-0.290	-7.207	***
LIBOR	0.435	4.657	***	-0.359	-3.753	***
Brent_forw	0.085	9.713	***	-0.233	-26.523	***
VIX	-0.178	-3.179	**	-0.231	-3.689	***
Aug	0.147	2.194	*	-0.330	-6.900	***
Sep	-0.242	-2.985	**	-0.019	-0.362	
Oct	0.571	10.142	***	-0.492	-9.031	***
Nov	0.408	4.453	***	0.037	0.575	
Dec	0.916	12.374	***	0.136	1.978	*
Jan	-0.062	-0.718		0.158	2.570	*
Feb	0.089	1.244		-0.496	-7.981	***
Mar	0.271	3.240	**	0.236	2.984	**
Apr	-0.372	-3.414	***	-0.186	-2.735	**
May	-0.363	-4.336	***	0.007	0.136	
Jun	-0.807	-12.525	***	0.315	4.387	***
R-squared	0.9892			0.9685		
R-squared-adj	0.9638			0.9354		
Approx. num.	60			84		

Significance 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

seasonal dummies do follow, almost, the same seasonal pattern throughout the year. So, do these differences make sense? Yes, this is well in line with theory presented, considering the elasticity characteristics of the supply curve. It is thus proven that freight rates have more evident seasonal effects in high volatility regimes.

Additionally, we present seasonal plots across all years in a 3D format in Appendix A.4 to see whether we can observe periods where the seasonal components seemed abnormal. Occurrences of abnormality and spikes in seasonality should thus be mostly related to the high volatility regime considering our discussion above. Our brief take on these plots is that dirty tanker routes clearly exhibit a higher degree of abnormality than clean tanker routes.

8.4 Checking the Assumptions

We continued to perform diagnostic tests to look for indications of misspecification in the models. The set of assumptions, as presented in Section 7.7, were checked against the pooled residuals, which should give an approximate representation of the underlying regression models. The results in Table 8.11, and the plots in Figure D.1 and D.2 from the Appendix were consulted during the diagnostic.

The linear relationship between the dependent variable and explanatory variables was evaluated by studying the plots of residuals versus fitted values, in D.1 and D.2. The observations appeared to be symmetrically distributed around the horizontal line. Some minor fluctuations in the line were discovered for all models.

An investigation into the presence of autocorrelation in the error terms was made. With no serial correlation, a Durbin-Watson (DW) statistic of 2 is expected. For the domain models, the DW statistic showed no apparent signs of autocorrelation. The associated DW statistics diverges only slightly from a value of 2. As for the Breusch-Godfrey (BG) test, a noteworthy exception appeared at $order = 1$ and $order = 5$. TD1 reject the null hypothesis of no serial correlation at a 10 % significance. Further inspection by addressing the ACF plots was required, see Figure D.1. TD1 showed a barely significant positive correlation at lag 5 and negative correlation at lag 9. There were however no clear patterns, and no initial apparent signs of serial correlation.

For the parsimonious models, both TD3_{pars} and TD7_{pars} were identified to have a DW statistic noticeably greater than 2. This indicates that successive error terms are negatively correlated. From the BG test we observe significant statistics at 1 % for all orders for TD7_{pars}. The ACF plot validates this, and significant correlations exist at lag 1, 3, 4, 5 and 6. The pattern seems to shift from positive to negative for these values. Misspecification might be of issue for TD7_{pars}. Overall, the (mostly) lack of correlation suggested that the forecasts were good.

To assess heteroscedasticity in the errors, we first looked to the Breusch-Pagan (BP) test. No route models showed evidence of rejecting the null hypothesis of homoscedastic errors. This was further evaluated against the plots of fitted values versus residuals. The residuals were approximately evenly distributed along the horizontal axis. No signs of the typical funnel shape for heteroscedastic errors were seen.

To evaluate normality, the Jarque-Bera test was employed. We were able to reject the null hypothesis for both TD12 and TD7_{pars} with a 1 % significance, as well as for TC2 with a 10 % significance. The latter is not considered much of an issue, as normality is not a strong requirement. The prior are not likely to originate from a gaussian distribution. The Q-Q plots in D.1 and D.2 supported the evidence of non-normality for TD12 and TD7_{pars}. The residuals were not distributed along the diagonal line, and appeared to have fat tails. It was noted that TD7_{pars} appears to have repeated issues. Deviations from normality is not necessarily a cause for concern as analyses are quite robust.

8.5 Model assessment

In this section we will analyse the results obtained from in-sample and out-of-sample tests to evaluate predictive power. The latter is performed with the three objectives in mind, and will be compared to benchmark models.

In-Sample Comparison

We first evaluate the models' performance on the training data. How well does it manage to fit the observations? We investigate the in-sample results in Table 8.14, for our variable rich models and our parsimonious models.

We first consider the regular *domain knowledge models*. What is initially apparent, is that weighted R^2_{adj} appears to be abnormally high. By looking specifically at TD1 in Table 8.5, we see very large R^2_{adj} values. The normal regime has a value of 0.81, while the volatile regime has the significantly higher value of 0.98. The same is true for the other routes, with weighted R^2_{adj} in the upper percentile. Does the same apply for the other metrics in Table 8.14? We see very low residual errors, for example, the MASE, which is in the approximate range of 0.10 - 0.15. The is most likely a result of overfitting, whereby the models are fit too well to the training data. We can already postulate that these models will not generalize well to never-before-seen data.

Table 8.11: Results from the statistical tests performed to evaluate the underlying assumptions. The table contains statistics from the Breusch-Pagan test for heteroscedasticity, the Durbin-Watson test for serial correlation, the Jarque-Bera test for normality, and the Breusch-Godfrey test for higher order serial correlation, for each route model.

	B-P	D-W	J-B	B-G (1)	B-G (2)	B-G (3)	B-G (4)	B-G (5)	B-G (6)
TD1	0.952	2.004	3.073	3,052*	4.050	4.874	4.850	9,278*	9.290
TD3	0.022	1.709	3.325	0.911	1.224	1.359	1.543	1.594	2.247
TD7	0.003	2.204	2.966	1.606	2.452	2.465	2.536	2.977	3.515
TD12	1.847	1.950	68,998***	2.112	2.476	2.762	3.196	3.264	8.740
TC1	0.001	1.803	3.248	0.865	1.322	1.528	1.632	1.491	3.394
TC2	0.619	1.861	4,757*	1.237	2.286	3.484	4.121	4.130	4.652
$TD1_{pars}$	0.042	2.169	0.123	0.501	2.504	2.679	2.908	2.927	3.300
$TD3_{pars}$	0.005	2.538	0.495	0.375	1.073	1.559	4.085	5.363	8.945
$TD7_{pars}$	0.109	3.028	71,905***	28,625***	31,583***	32,324***	40,624***	41,395***	41,349***
$TD12_{pars}$	1.599	2.048	0.752	2,736*	3.426	5.782	6.414	7.126	7.325
$TC1_{pars}$	0.001	1.830	0.951	0.018	0.022	2.736	2.724	2.674	2.727
$TC2_{pars}$	0.258	2.333	0.746	0.291	0.878	1.075	1.816	3.919	6.421

Significance: *** 0.01, ** 0.05, * 0.1

We then continue to look at the *parsimonious regime models* for the same routes, in Table 8.14. These weighted R^2_{adj} values seem far more reasonable, with most models being around 0.40. In other words, these values do not appear as inflated as for the regular regime models. Similarly, higher residual errors are now observed. $TC1_{pars}$ and $TD12_{pars}$ do appear to have the lowest RMSE, MAE, MAPE and MASE, and notably higher R^2_{adj} values. Seeing them deviate that much from the other models, can indicate that we will see problems with these two models' out-of-sample performance. Overall, however, the parsimonious models thus seem like better candidates, with less likeliness of misspecification.

From the initial look at the models' in-sample performance it was evident that the parsimonious models showed the most promise. The domain knowledge models score better on metrics, but these numbers are so high that one comes to think that the models might be suffering from overfitting or misspecification (see, further, Chapter 9).The question now, is how well these models will generalise to new data.

Out-of-Sample Comparison

The out-of-sample comparison of forecasts is performed in conjunction with the evaluation of the three objectives.

For the benchmark models, we included one minimum requirement model, the random walk with a drift component. A well-defined model is expected to outperform this basic measure, based on random values. The remaining three models were set up to be challenging to beat. These models include a fitted linear model with all the same carefully selected variables as our parsimonious regime model; a model based on the average historic values in the training sample; and an automatically configured best fit ARIMA(p, d, q) model (with p, d, q determined by the best fit for the given route). In other words, we evaluate our regime models against tough contenders in the task of forecasting the tanker freight rate.

The forecasting period is set from August 2014 to June 2017. A comprehensive summary of prediction performance for each model is presented in Table 8.12. The accuracy is measured using the RMSE, MAE, MAPE, MdAPE and MASE. The best forecasting model is the one with the lowest values. Most interesting is the MASE, which is the more robust performance metric. We do expect too see a rise in error metrics when comparing out-of-sample to in-sample metrics. The Diebold-Mariano test for forecasting accuracy is also considered.

As we were looking to model volatile periods, and irregularities, we did not expect to significantly outperform models with a more straight-forward approach - models which do not necessarily account for such behavior, like the average, and linear model benchmarks. Modelling extraordinary characteristics is a difficult task, and thus we might assume that simpler models will perform better on average. However, the *potential* from getting it right, is much greater for models that can predict uncommon movements.

I Considering a parsimonious model

We evaluate the explanatory power of the parsimonious regime models against the domain knowledge regime models for different routes. We look to uncover whether a regime model with far less predictors will fare better or worse than a regime model with more variables.

We refer to the the performance metrics for the domain knowledge regime models and the parsimonious regime models which can be seen in Table 8.12[60]. As lower values indicate a better performing model, it is immediately evident that the parsimonious models

[60]It must be noted that the MAPE metric is abnormally high for all models, as a consequence of the metric being used on return values. Percentage errors are unreliable and can have an extremely skewed distribution at values close to zero (Hyndman, 2006).

Table 8.12: Out-of-sample performance measures for all route models considered in this thesis, including benchmark models. All models are «scored» compared to one parsimonious model and one variable rich domain knowledge model. The following can be assessed from the table: (i) Domain and parsimonious models are compared to see if a model with fewer variables perform better; (ii) The next four benchmarks are used to evaluate the out-of-sample forecast accuracy of the prior two regime models (with a focus on the parsimonious one); and (iii) The last global benchmark is used to study if route-specific models outperform a generic model based on global variables.

						~Best score	
	RMSE	MAE	MAPE	MdAPE	MASE	Domain know.	Pars.
Parsimonious regime models							
$TD1_{pars}$	0.380	0.323	142.3	96.4	0.525		
$TD3_{pars}$	0.446	0.381	318.9	90.8	0.397		
$TD7_{pars}$	0.670	0.543	774.6	119.5	0.436		
$TD12_{pars}$	0.259	0.206	2035.1	122.2	0.598		
$TC1_{pars}$	0.334	0.244	318.3	121.6	0.717		
$TC2_{pars}$	0.415	0.276	160.9	84.3	0.681		
Domain knowledge regime models							
TD1	0.737	0.562	313.0	111.6	0.914		✓
TD3	0.695	0.588	654.7	156.1	0.611		✓
TD7	1.025	0.849	1628.6	190.3	0.682		✓
TD12	0.340	0.265	2900.8	166.0	0.771		✓
TC1	0.377	0.263	387.8	101.2	0.771		✓
TC2	0.528	0.424	299.5	122.4	1.047		✓
Benchmark models							
RW-TD1	0.713	0.514	203.671	132.076	0.835	÷	✓
RW-TD3	0.633	0.477	377.586	127.428	0.496	÷	✓
RW-TD7	0.865	0.674	971.685	147.806	0.541	÷	✓
RW-TD12	0.257	0.198	3161.424	133.151	0.577	÷	÷
RW-TC1	0.396	0.319	695.615	160.417	0.934	✓	✓
RW-TC2	0.472	0.386	565.985	124.220	0.954	÷	✓
Fit-TD1	0.409	0.321	155.3	94.0	0.523	÷	÷
Fit-TD3	0.537	0.429	276.4	122.6	0.446	÷	✓
Fit-TD7	0.744	0.593	1086.8	118.8	0.476	÷	✓
Fit-TD12	0.198	0.166	2098.1	96.7	0.483	÷	÷
Fit-TC1	0.240	0.195	667.1	80.7	0.573	÷	÷
Fit-TC2	0.417	0.328	222.5	114.1	0.810	÷	✓
Mean-TD1	0.479	0.383	100.2	100.4	0.623	÷	✓
Mean-TD3	0.419	0.342	101.5	100.9	0.356	÷	÷
Mean-TD7	0.538	0.414	97.8	100.5	0.332	÷	÷
Mean-TD12	0.210	0.169	139.6	100.4	0.491	÷	÷
Mean-TC1	0.267	0.211	104.5	99.8	0.619	÷	÷
Mean-TC2	0.359	0.284	107.1	101.7	0.703	÷	✓
Arima-TD1	0.466	0.381	133.7	116.2	0.619	÷	✓
Arima-TD3	0.490	0.421	294.3	123.4	0.438	÷	✓
Arima-TD7	0.506	0.416	404.1	106.6	0.334	÷	÷
Arima-TD12	0.212	0.178	1740.2	100.6	0.518	÷	÷
Arima-TC1	0.267	0.211	100.0	100.0	0.620	÷	÷
Arima-TC2	0.347	0.293	266.2	121.4	0.723	÷	✓
Generic variables benchmark regime models							
Global-TD1	0.878	0.615	263.305	143.179	1.000	✓	✓
Global-TD3	0.508	0.395	310.076	110.551	0.411	÷	✓
Global-TD7	1.176	0.982	1481.228	213.937	0.788	✓	✓
Global-TD12	0.258	0.215	5063.729	118.803	0.625	÷	✓
Global-TC1	0.378	0.304	752.205	139.098	0.891	✓	✓
Global-TC2	0.518	0.398	543.786	88.500	0.983	✓	✓

Table 8.13: Diebold-Mariano results for forecasting accuracy for the parsimonious regime models versus comparison models.

		Benchmark models				
Pars. model	Domain. know.	RW	Fit	Mean	Arima	Global
$TD1_{pars}$	3,202*** ✓	2,092** ✓	0.574 ÷	1,43* ✓	1,391* ✓	1,6* ✓
$TD3_{pars}$	2,598*** ✓	2,127** ✓	1,763** ✓	-0.464 ÷	0.706 ÷	0.935 ÷
$TD7_{pars}$	2,592*** ✓	1,728** ✓	1.212 ÷	-1.908 ÷	-1.775 ÷	3,305*** ✓
$TD12_{pars}$	1,694** ✓	-0.053 ÷	-2.183 ÷	-1.547 ÷	-1.401 ÷	-0.053 ÷
$TC1_{pars}$	0.918 ✓	1,409* ✓	-2.408 ÷	-2.017 ÷	-2.018 ÷	2,106** ✓
$TC2_{pars}$	1,397* ✓	0.533 ÷	0.026 ÷	-0.700 ÷	-0.657 ÷	0.919 ÷

Significance: *** 0.01, ** 0.05, * 0.1

yield more precise results. A MASE below 1 also means that the forecast is better than a naive forecast. Some of the domain knowledge models have a high MASE, even close to or above 1. This is especially true for TD1 and TC2. Considering the parsimonious models, approximately four out of six models are likely to have about half as much error as the naive forecast. We also see that the models for both clean routes, TC1 and TC2, seem to have a higher MASE value than other parsimonious models.

When comparing the in-sample performance metrics in Table 8.14 with the out-of-sample performance metrics in Table 8.12, we see a notable discrepancy for the domain knowledge models. These models have residual metrics which point towards an improbably excellent model in-sample, while performing close to a naive forecasting approach out-of-sample. This is a very likely sign of an overfitted model. For the parsimonious regime models, the difference is much more in line with what is expected. The models get reasonable results, and performs slightly better in-sample. The two models $\mathrm{TD12_{pars}}$ and $\mathrm{TC1_{pars}}$ are notable, though. They provide significantly better results in-sample than out-of-sample, and differentiate themselves from the other parsimonious models. We do not expect these two models to generalise well, when comparing them to the benchmark models in the next section.

Furthermore, we evaluate the results from the Diebold-Mariano test for predictive accuracy. The forecasts of the domain knowledge model is compared to the corresponding parsimonious model for each route, under the null hypothesis of equal predictive accuracy. The predictions for all dirty tanker routes reject the null hypothesis with a 5% significance, signaling a greater accuracy for the parsimonious model in the one-sided test. For the clean routes, we fail to reject the null hypothesis in the Diebold-Mariano test, and no significant difference was therefore detected.

However, the combined results of the performance metrics and the Diebold-Mariano test illustrates that the parsimonious models perform reasonably well for the 1-month-ahead forecasts, compared to the knowledge models. They also behave more reasonably, and do not appear to have the same degree of overfitting. *The parsimonious models will therefore be used for further assessments*, and are considered an improvement of the domain knowledge models.

II Assessing the forecasting capabilities of our regime-model

As previously mentioned, it is imperative that the forecasting models perform adequately on never-before-seen data. The models are consequently compared to benchmark models to assess the adequacy of the predictions. Figures 8.4 and 8.5 show a comparison of the actual values for each route versus the 1-month-ahead forecast of the parsimonious regime models. Tables 8.12 and 8.13 contain the key results for the comparison.

Earlier, we saw that the parsimonious models performed better, and therefore these are the most interesting to evaluate[61].

$\mathbf{TD1_{pars}}$ From plots in Figures 8.4 and 8.5, the model appears to capture the freight rate's movements very well. This model performs roughly 50 % better than a naive model on the new data, according to the MASE. When compared to the benchmark model, random walk with a drift component (RW), we also see good results. For the more challenging benchmarks, the difference becomes far less. TD1 yields better metrics than the Mean model and the Arima model, but does fail to beat the simple regression model, Fit.

Similar results are observed when evaluating the Diebold-Mariano statistics. TD1 provides better forecasting accuracy than 3/4 of the benchmark models, with a significance of 10 %.

$\mathbf{TD3_{pars}}$ The plot in Figure 8.5 appears equally good as TD1. The directions look to be mostly correct, and the deviations look to be small. The model outperforms the minimum benchmark, and also provides *slightly* better metrics than Arima and Fit, while being unable to beat Mean. The DM statistics indicate that the model's forecasts are more accurate than both RW and Fit with a 5 % significance.

$\mathbf{TD7_{pars}}$ The plot in Figure 8.5 does have most of the same characteristics as the prior models. From a visual inspection, the model does, however, seem to miss the mark more often, but this can also be attributed to what looks to be a time series with more movements. The MAE and MASE of 0.670 and 0.436, respectively, yields better results than both the RW and Fit benchmark models. Both Arima and Mean appear to have very low metrics for this time series, and are not outperformed. These findings are confirmed by the DM test, whereby only the RW is beaten, with a 5 % significance.

$\mathbf{TD12_{pars}}$ For this model, we postulated poorer results based on the in-sample model review. This is not inherently clear from the plots in Figures 8.4 and 8.5. We do, however, see lower performance when comparing it to the benchmark models. TD12 fails to yield better results than all of the benchmarks. The difference is not necessarily great, but the failure to

[61] The variable rich domain knowledge model does not provide adequate results. It only outperforms the minimum requirement benchmark in one instance.

Table 8.14: Performance measures for route specific regime models, in-sample

Route	RMSE	MAE	MAPE	MdAPE	MASE	$R^2_{adj,weighted}$
TD1	0.119	0.091	158.038	29.266	0.148	0.870
TD3	0.188	0.140	162.014	27.495	0.146	0.792
TD7	0.254	0.182	457.526	24.337	0.146	0.844
TD12	0.061	0.042	53.783	16.921	0.122	0.910
TC1	0.050	0.038	59.024	14.483	0.111	0.933
TC2	0.059	0.043	76.188	15.243	0.107	0.947
$TD1_{pars}$	0.277	0.215	351.127	83.982	0.350	0.379
$TD3_{pars}$	0.450	0.352	331.427	84.158	0.366	0.325
$TD7_{pars}$	0.811	0.532	984.047	72.304	0.427	0.392
$TD12_{pars}$	0.138	0.100	114.060	47.888	0.289	0.623
$TC1_{pars}$	0.108	0.083	117.228	34.117	0.243	0.799
$TC2_{pars}$	0.183	0.141	2606.474	61.999	0.349	0.423

beat the minimum requirement RW can be considered troublesome.

$TC1_{pars}$ As was the case with the prior model, this model was also expected to perform less than adequate on never-before-seen data. We see this come to fruition when it does not give any better results in the metrics nor the DM test, apart from beating the RW.

$TC2_{pars}$ For this regime model, the plot in Figure 8.5 does look promising. Although the time series appears to have less spiky movements, we see the model providing better residual measures than both RW, Fit, Mean and Arima. Being able to contend with established time series models is promising. We do not see a significantly better accuracy in the DM test, though.

About three years of data was included for the test period. This is not necessarily long enough to cover the cyclical behavior which characterizes the shipping freight rates. The impact of accounting for several states could thereby also potentially be impaired, as few high volatility occurrences might happen. It would have been interesting and favorable to evaluate the performance on more test data.

Although, from these observations, we frequently see the parsimonious regime models being ranked in the top two places[62] for the performance metrics, compared to the benchmark models. So, while no model is consistently better on all of the routes, the regime models are amongst the better. This in turn means that they can be assumed to generalise well to new data.

III Evaluating the route impact

Is there a benefit to specify models specifically for certain routes? To evaluate the benefit of creating route-specific models, generic benchmark models were established based on *global variables*. These regime models were set up as models which could be applicable to all routes, with no distinct route characteristics for the variables (as seen in Section 7.8). The global benchmark model, Global, was used to forecast the 1-month-ahead freight rate for the same routes as our other regime models. For this section, we only consider the parsimonious models[63]. In the following, we present the results from these forecasts.

From the Diebold-Mariano tests of forecast accuracy, we saw evidence of some routes being better predicted by route specific models (see Table 8.13). For the regime model $TD1_{pars}$, we were able to reject the null hypothesis with a 10% significance The forecasts from $TD7_{pars}$ and $TC1_{pars}$ were even more promising, with a rejection of the null hypothesis at a significance level of 1% and 5% respectively.

Similar results can be discerned by comparing the performance metrics from the different route models with the global benchmark model in Table 8.12. The parsimonious models appeared to outperform the generic benchmark model based on global variables, for *all* models, $TD1_{pars}$, $TD3_{pars}$, $TD7_{pars}$, $TD12_{pars}$, $TC1_{pars}$ and $TC2_{pars}$.

Now, this improved forecasting accuracy could possibly be attributed to our previous findings, of parsimonious regime models outperforming models with more variables. The generic benchmark model based on global variables did consist of approximately two and a half times as many variables as the parsimonious models. This makes the model fall somewhere in-between the domain regime models and the parsimonious regime models, in terms of the number of variables. We are therefore careful to make assertions regarding the distinct route impact. However, the majority of the models did give indications of better results in the performance metrics, and this includes the domain knowledge models[63]. We therefore still *propose* that a route-specific model could improve forecasting.

[62] Rank ~ 1, 2, 3, 2 and 1 for the routes, respectively.

[63] The domain knowledge regime model was also considered, but did not outperform the global benchmark within a 5% significance. We were, however, able to identify four domain knowledge models that performed better, TD1, TD7, TC1 and TC2 in regards to the residual performance metrics.

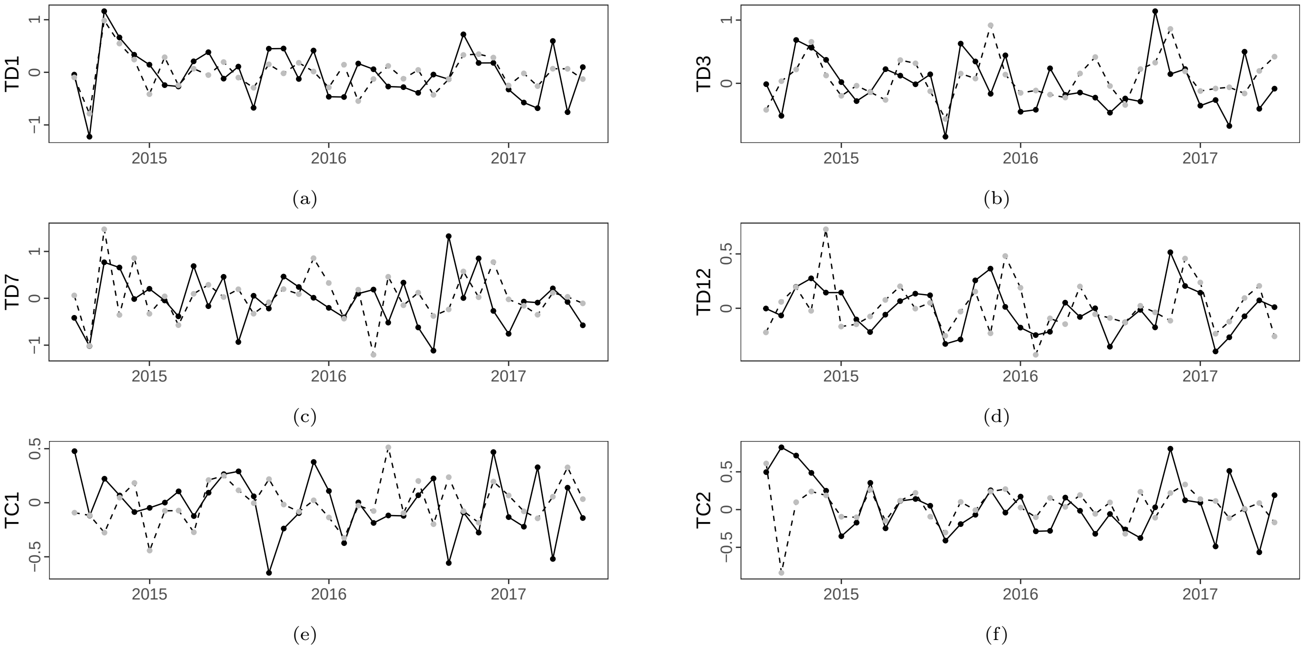


Figure 8.4: Comparing the forecasted returns (stapled line) from the parsimonious regime model, with the actual values (solid line), for each route.

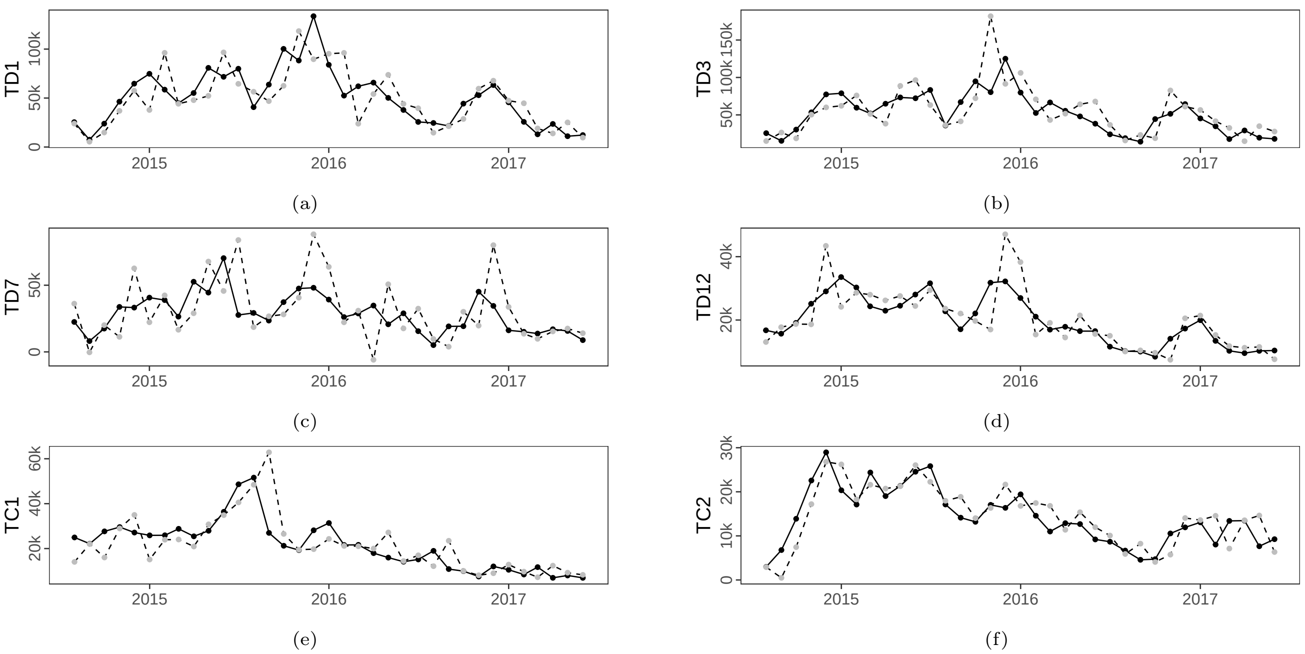


Figure 8.5: Comparing the forecasted freight rates (stapled line) from the parsimonious regime model, with the actual values (solid line), for each route.

Table 8.15: Estimated coefficients for the TD1 parsimonious regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	0.784	54.055	***	-0.057	-0.558	
ClarkSea	8.240	115.246	***	0.799	3.442	***
SA_exp	10.757	83.909	***	0.451	0.799	
Chi_imp	-1.480	-20.789	***	-0.694	-3.717	***
VLCC_due	1.778	98.778	***	0.135	2.194	*
VIX	3.982	117.799	***	-0.129	-0.753	
Aug	-1.795	-104.983	***	-0.221	-1.455	
Sep	0.332	12.834	***	0.111	0.779	
Oct	-0.638	-27.157	***	0.159	1.126	
Nov	-0.974	-32.346	***	0.306	2.083	*
Dec	-0.705	-32.064	***	0.110	0.768	
Jan	-1.179	-56.956	***	-0.128	-0.873	
Feb	-2.151	-77.935	***	0.202	1.448	
Mar	-1.115	-54.141	***	0.123	0.778	
Apr	-0.271	-11.275	***	-0.184	-1.297	
May	0.629	25.553	***	0.086	0.621	
Jun	-0.558	-24.902	***	0.015	0.100	
R-squared	1.000			0.354		
R-squared-adj	0.999			0.255		
Approx. num.	24			120		

Significance 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8.16: Estimated coefficients for the TD3 parsimonious regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	-0.172	-1.164		-0.812	-53.098	***
VIX	-0.314	-1.174		3.759	139.219	***
Jap_money	15.522	2.844	**	28.083	69.892	***
VLCC_SP	-1.653	-1.624		10.990	322.296	***
Ind_imp	-0.385	-0.847		-4.010	-60.302	***
Aug	-0.249	-1.163		-0.616	-33.650	***
Sep	0.337	1.537		-0.509	-24.570	***
Oct	0.163	0.753		2.294	108.716	***
Nov	0.791	3.690	***	3.460	159.442	***
Dec	0.138	0.660		1.878	82.722	***
Jan	0.075	0.361		-1.150	-56.382	***
Feb	0.182	0.841		-0.675	-38.770	***
Mar	-0.072	-0.337		0.969	52.923	***
Apr	-0.120	-0.579		0.430	15.368	***
May	0.031	0.147		3.874	177.711	***
Jun	0.473	2.224	*	1.501	64.412	***
R-squared	0.310			1.000		
R-squared-adj	0.216			1.000		
Approx. num.	124			20		

Significance 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8.17: Estimated coefficients for the TD7 parsimonious regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	0.266	0.469		-0.135	-1.052	
Chi_imp	2.457	1.807	.	0.188	0.863	
Ind_India	0.076	1.470		0.056	3.810	***
LIBOR	-1.136	-0.861		0.552	1.357	
ClarkSea	0.325	0.210		0.580	1.873	.
Afra_demo_price	3.656	1.361		0.171	0.470	
Eur_dem	6.331	1.070		-2.727	-2.721	**
WTI	2.117	0.985		-0.722	-1.599	
Aug	-0.145	-0.164		-0.285	-1.661	.
Sep	-1.844	-1.974	*	0.218	1.218	
Oct	0.704	0.794		0.660	3.506	***
Nov	-0.352	-0.412		0.153	0.821	
Dec	0.730	0.779		0.973	4.780	***
Jan	-1.030	-1.186		-0.213	-1.194	
Feb	-0.815	-0.950		-0.010	-0.057	
Mar	0.442	0.523		0.032	0.183	
Apr	-0.298	-0.365		-0.448	-2.253	*
May	-1.788	-1.901	.	0.652	3.190	**
Jun	0.139	0.168		-0.090	-0.501	
R-squared	0.425			0.638		
R-squared-adj	0.160			0.540		
Approx. num.	56			88		

Significance 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8.18: Estimated coefficients for the TD12 parsimonious regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	0.056	0.961		-0.408	-77.057	***
ClarkSea	0.664	4.871	***	1.164	41.568	***
Pana_price	0.040	2.230	*	0.044	15.000	***
Ind_US	-0.005	-0.329		0.362	150.792	***
WTI	-0.307	-1.760	.	-0.575	-17.683	***
Bel_fexp	0.065	1.172		-0.925	-58.176	***
Chi_imp	-0.149	-1.412		0.441	24.781	***
Brent_forw	-0.008	-0.597		-0.132	-66.200	***
Bel_fimp	-0.216	-2.988	**	-0.642	-58.908	***
US_dem	1.151	1.352		-6.944	-60.384	***
Aug	-0.181	-2.161	*	0.106	6.644	
Sep	-0.030	-0.387		0.730	54.485	***
Oct	0.176	2.320	*	0.395	37.226	***
Nov	-0.054	-0.682		-0.173	-16.178	***
Dec	0.348	4.488	***	0.601	71.548	***
Jan	-0.015	-0.173		0.595	70.000	***
Feb	-0.178	-1.932	.	0.281	24.658	***
Mar	-0.118	-1.442		1.038	97.925	***
Apr	-0.115	-1.427		-0.203	-24.410	***
May	0.078	1.010		0.512	55.630	***
Jun	-0.160	-1.911	.	0.372	40.424	***
R-squared	0.607			1.000		
R-squared-adj	0.524			0.999		
Approx. num.	114			30		

Significance 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8.19: Estimated coefficients for the TC1 parsimonious regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	0.121	1.912	.	0.099	1.970	*
Afra_fix_east	0.130	2.964	**	0.246	6.685	***
W_prod	17.204	6.034	***	1.309	0.464	
SDR_USD	3.195	2.826	**	3.978	5.230	***
LR2_order	-0.637	-3.058	**	-1.176	-5.356	***
ClarkSea	0.832	4.247	***	0.471	3.432	***
Ind_Jap	0.013	2.867	**	0.014	3.512	***
Brent_forw	-0.093	-4.194	***	-0.039	-2.543	*
Ind_US	0.006	0.247		-0.052	-2.802	**
Jap_cimp	-0.126	-2.470	*	0.068	1.434	
SA_cout	0.128	0.328		-0.629	-3.370	***
Tadawul	-1.208	-5.347	***	0.734	3.180	**
Afra_fix_us	-0.004	-0.177		0.065	2.016	*
Aug	0.063	0.796		-0.477	-5.499	***
Sep	-0.040	-0.469		0.186	2.420	*
Oct	0.036	0.399		-0.546	-7.909	***
Nov	-0.456	-5.336	***	-0.002	-0.024	
Dec	-0.207	-2.187	*	0.164	2.262	*
Jan	-0.295	-2.807	**	-0.369	-4.788	***
Feb	-0.338	-4.021	***	-0.156	-1.643	
Mar	-0.264	-2.738	**	0.043	0.481	
Apr	-0.669	-5.906	***	-0.007	-0.090	
May	0.281	2.141	*	-0.036	-0.562	
Jun	0.034	0.400		-0.175	-2.144	*
R-squared	0.851			0.877		
R-squared-adj	0.782			0.817		
Approx. num.	75			69		

Significance 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 8.20: Estimated coefficients for the TC2 parsimonious regime-switching model

	Regime 1			Regime 2		
	Coef	t-value		Coef	t-value	
(Intercept)	0.079	1.040		-1.220	-24.740	***
USNe_exp	-0.677	-1.884	.	-5.739	-32.025	***
MR_order	-0.735	-1.182		6.676	12.197	***
Ne_cimp	-0.210	-1.512		-0.360	-3.909	***
Ne_cout	-0.560	-1.962	*	-1.088	-5.928	***
Brent_forw	-0.095	-3.918	***	0.125	6.505	***
OPEC_prod	4.043	2.963	**	4.882	3.855	***
MR_3tc	0.171	0.258		-2.059	-3.441	***
Aug	-0.319	-2.979	**	1.704	22.752	***
Sep	0.011	0.105		0.854	9.348	***
Oct	-0.068	-0.653		0.092	1.485	
Nov	0.111	1.071		0.970	18.412	***
Dec	0.158	1.495		2.565	37.178	***
Jan	-0.042	-0.417		1.123	10.031	***
Feb	-0.124	-1.179		1.061	19.188	***
Mar	-0.031	-0.296		1.079	13.298	***
Apr	-0.176	-1.662	.	0.629	10.528	***
May	-0.067	-0.679		2.810	21.063	***
Jun	-0.089	-0.858		1.001	15.205	***
R-squared	0.401			0.996		
R-squared-adj	0.293			0.984		
Approx. num.	117			27		

Significance 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

9 Discussion

Modelling Improvements and Next Steps

We saw abnormally good in-sample results for several of the domain knowledge models. This could potentially be a problem of overfitting, or misspecification. A model is overfit when it adapts too much to the training data, as previously mentioned. As we saw a large discrepancy between the in-sample results and the out-of-sample results, this leads us to believe that such a problem could have occurred. Issues were found in the residuals of some models, notably TD7_{pars}. If the residuals misbehave, the results are not necessarily reliable, and thus cannot be trusted.

We did see the parsimonious models provide more realistic in-sample results. The domain regime models consisted of far more variables than the parsimonious models. An issue can therefore also stem from multicollinearity, whereby one explanatory variable can be linearly predicted based on others. A more comprehensive Variance Inflation Factor (VIF) analysis could potentially be performed to evaluate this, *while keeping in mind that we are dealing with a multi-regime model*[64]. In short, the VIF test is based on the regression of a single predictor against all the other predictors. If the R^2 is high, there is an indication that the relationship of this predictor to the dependent variable is already accounted for in one or several of the other predictors.

The problem of omitted variables could potentially also be an issue. We made a good effort to *consider* as many potentially relevant variables as possible. However, a great challenge lies in finding variables that might describe the volatile regime with better accuracy. Having variables that can explain more of the irregular movements would be beneficial. Such variables are not easy to come by, and there could therefore be room for exploring more candidate predictors.

During the transformation of variables to adhere to the need for stationarity, there is a loss of information. Some variable transformation could potentially have been avoided. This is especially true for the dependent variables, whereby four out of six variables already rejected the null hypothesis of the presence of a unit root (see, further, Table 6.3). However, they were all transformed with logarithmic differencing to unify the modelling approaches, and make them easier to work with.

The predicted probabilities of being in a regime one-month ahead, appeared less than accurate. The infrequent, high volatility regime, appeared to get too much of an influence in the weighting of the forecasted values. The extreme impact of this can possibly be seen by looking at the minimum freight rates for the regimes, with some values being more than -100,000 (see Table 8.3). This was a consequence of too moderate distinction between the two regimes in the transition matrix, which was used to calculate the step-ahead probabilities. This can be seen when estimating the smoothed probability for the test period. When we compared the smoothed probabilities, estimated based on seeing the actual state values, to the forecasted probabilities, the smoothed probabilities were almost close to binary. The certainty of which state the observation was from, was far more precise. In other words, the forecasted probabilities was far less decisive in the indication of the current state. The regime model thus had a bias towards predictions from the high volatility regime. Better probability predictions could improve the forecasts.

The inclusion of more frequent data points could provide a significant benefit. With more observations, the regime model could potentially be more efficient at locating structural breaks. We would have liked to include data on a weekly, or even daily frequency. However, it was not possible to gather explanatory variables to the extent we have done with such frequency.

For the selection of lags, we could have expanded the scope to include more lags than six. We are certain that we were not able to capture the full relationship between some dependent variables and independent variables with only a six months lag. Endogenous supply variables may even have a lead time of several years before they directly affect the freight rates. The impact of, e.g., the orderbook and newbuilding contracting might thus not yield any results to the fleet capacity until vessels are delivered and readily available in the market. As such, an even greater lag analysis could be performed.

[64] Note, VIF tests based on simple OLS regression were performed on all variables (see, further, Section 8.2).

10 Conclusion

An in-depth analysis of the freight market for oil tankers has been performed. Regime-switching regression models were then developed in order to provide predictions of one-month ahead freight rates for six tanker routes.

The existing literature on freight rates in the shipping field was reviewed. There exists a lot of attempts at modelling the freight rates, and a clear increase in the level of sophistication is seen when comparing earlier work to more recent efforts. Several studies were also found to investigate the determinants, the driving factors, of the shipping markets on both a microeconomic and a macroeconomic level. Additionally, shipping literature also exist in the context of multi-regimes. However, these studies are mainly focused on assessing risk and volatility. The results from these studies are nevertheless conclusive, and confirm the presence of distinctive regimes in the freight rates - whereby freight rate characteristics indicate a normal, low volatility state and a more irregular, high volatile state. Besides empirical shipping literature, we were able to find inspiration in other research fields, such as oil and electricity price modelling. This lead us to explore a combination of Markov regime-switching and multiple regression, applied to specific tanker routes in the oil tanker market.

We further considered the theoretical foundation of the shipping market. The four shipping markets and common stakeholders were explored. We discussed different contract types and industry measures, and concluded that the freight rate quoted on a Time Charter-Equivalent basis would be best suited for the modelling to come. The freight rate mechanism and shipping cycles were then assessed, which were key to our choice of building a two-regime model - a model which could potentially correspond with the elastic and inelastic part of the supply curve. A look into the tanker market specifically was done, and we considered cargo, vessels, and market characteristics - on which the market appears to operate under nearly perfect competition.

We moved on to take a specific look at the tanker trade, and the different routes we would be taking an in-depth look at, as well as model. After addressing a set of criteria, we landed on six routes, TD1, TD3, TD7, TD12, TC1, and TC2. The routes span the globe and cover different vessels and cargo.

After having gained insight into the oil tanker market and various routes, the key driving factors of the tanker market were analysed. We divided the determining factors into three major groups, based on supply, demand, and economic and non-fundamental factors. Determinants such as fleet size, orderbook, vessel prices, oil demand, oil import and export, crude oil prices, refinery output and a range of global indices were assessed. The intricacies and relationship between the determining factors were evaluated at great length.

Moreover, we began the process of collecting data. We also noted a few stylized facts about the freight rates, such as a tendency to mean-revert, evidence of seasonality and frequent rate jumps. The data was fetched from a collection of sources, and some initial descriptive statistics were presented.

Subsequently, the modelling methodology was outlined - how we would approach the modelling of route-specific oil tanker freight rates using Markov regime-switching multiple regression. We established three objectives to study the model benefits:

i) to figure out if a parsimonious model with fewer variables would outperform a variable rich model;

ii) to see how well the model forecasted;

iii) to see if incorporating route-specific variables would yield a benefit.

We then addressed how we would get the data set stationary, determine an appropriate number of lags and how we would model seasonality. Moreover, the method for variable selection was discussed. As we already had a solid theoretical foundation of important variables, we also wanted to assess this importance statistically. A procedure was set up based on stability selection with randomized LASSO. For this approach, one takes subsamples of variables and subsamples of data, and creates regression models based on the selection algorithm LASSO with a random tuning parameter. This technique was run thousands of times, and the results were aggregated. The most frequently appearing variables in the regression models were ranked the highest.

We moved on to the actual model formulation. We examined a method of discovering structural breaks using the Chow test, and presented the regime-switching model and the approach to estimating it based on the two-step expectation-maximization algorithm. We also noted how we would set up the model to be used for forecasting, by weighting predicted values from two models with the probability of being in each regime.

Insight into the underlying assumptions for OLS were presented. As our regime-switching model consisted of two multiple regression models, we wanted to assess their residuals. Approaches for evaluating linearity, autocorrelation, heteroscedasticity, and normality in these residuals were thus presented. Moving on, we laid out how we would assess the models' performance. The problem of overfitting was introduced, as well as the performance metrics to be used. A set of benchmark models were be established for comparison, and the Diebold-Mariano (DM) test for forecasting accuracy would be employed. The approach to studying the three main objectives were further embarked upon. All objectives would be considered with the performance metrics and the DM test. To consider the benefit of a parsimonious model, we would create parsimonious variants of our regime-switching models, consisting of a smaller set of variables, for each route. To assess the forecasting capabilities of the regime-switching model, it would be compared to benchmark models on out-of-sample data, for each route. To evaluate the benefit of incorporating route-specific variables in modelling, we compared the route-specific regime-switching models with a generic regime-switching model based only on global variables, for each route.

The actual model development was then performed. The initial preparation was handled, transforming the data to stationary and setting up lags. An analysis into the variable selection was subsequently performed. We found route-specific variables to be good contenders as they ranked high across all routes. We briefly looked into the presence of structural breaks by performing a Chow test, and found indications of volatile states, most evident in TD3 and TD7. The regime-switching models were later evaluated, and specifically the independent variable's impact on the freight rates, in accordance with the theoretical assumptions. Furthermore, we were able to replicate the theoretical seasonal components for each route well by introducing seasonal dummy variables. Findings of seasonality were in line with the findings of Kavussanos and Alizadeh (2002).

The different regimes were subsequently examined. We were able to construct important characteristics of each regime, such as volatility magnitude, durations and occurrences of volatility clusters – whereby the presumably lower volatility regime where found to hold the highest weighting of number of observations, the longest duration, and the lowest volatility, across all routes. Results did also somewhat indicate that volatility increases with an increasing vessel size. The regime findings can potentially support shipping portfolio managers and vessel operators, as expressed by Abouarghoub et al. (2014). One can use the improved understanding of the dynamics in a low volatility and high volatility state, in addition to the transition probabilities, duration of regimes, and level of returns to ones advantage during operations, hedging or speculation.

The model residuals were later checked for linearity, autocorrelation, heteroscedasticity, and normality. The residuals of most models were well-behaved, with the notable exception of TD7_{pars}, which had some repeated issues.

The in-sample performance of the regime models was then considered. A great discrepancy was found between the domain knowledge models, and the parsimonious models. The prior had very high R^2 and low error metrics. The latter showed more well-behaved results, in line with that is to be expected. It was therefore assumed that the prior models suffered from overfitting, or misspecification.

We moved on to consider the out-of-sample performance. The performance on never-before-seen data was done with the three objectives in mind.

Firstly the parsimonious regime models were evaluated against the variable rich domain regime models. The parsimonious models were found to outperform variable rich models in most metrics. This was also mostly true for the DM test, whereby four out of six forecasts were found to be more accurate with a 5% significance. Fewer variables therefore seemed to improve the modelling effort. Top performing variables in the parsimonious models include, secondhand prices, import and export factors, Chinese crude imports, vessel fixtures, and the ClarkSea index, amongst others.

Secondly, the forecasting capabilities of the parsimonious regime-switching models (as these were shown to be better) were evaluated against benchmark models. The benchmark models were set up to be challenging contenders, with the exception of the random walk. As we were attempting to model uncommon behaviour in the freight rates, and not simply the normal behaviour, the benchmark models were expected to outperform the regime models on regular observations. The regime models were, however, seen to generalise well to the new data. They provided results well above a naive forecasting method, and challenged the benchmark models on most routes. The regime models consistently scored in the higher tier.

Lastly, we evaluated the impact of constructing regime models with route-specific variables. The parsimonious regime models were evaluated against a generic benchmark regime model with global variables. By studying the performance metrics, all regime models were found to outperform the generic benchmark to various degrees. Three out of these were also found to have better forecasts with a 10% significance in the DM test. Accounting for route characteristics can therefore be considered beneficial. Finally, we discussed the findings, how the models could be improved and potential next steps. Different measures of handling the overfitted domain knowledge regime models were considered. The issue of omitted variables was brought up, and it was proposed that more high volatility predictors could be added to the model. We looked at the regime-bias in the forecasting probabilities, the benefit of more data and the value of increasing the scope for selecting lags.

Our opening statements for the thesis claimed that the tanker market exhibits high volatility. Risk is close to a necessity for speculation and high returns. Since we are attempting to model irregularities, and not just the normal behaviour, we are expected to miss more.

The irregularities and volatile conditions are more challenging to predict. Thus, the rewards of getting them right, are also potentially higher. Our models might miss, but when they hit, the returns can be significant. With higher frequency data, further research is highly recommended into the benefit of predicting the great, abnormal changes of the freight rates.

Appendix A Data

A.1 Data List

Table A.1: An overview of the input data used throughout the thesis, along with description, unit, format (e.g. average during the period, or start/end of month), and routes applicable for each variable.

Data period: August 2002 - June 2017 (Start from February 2002 if including lags)
Frequency: Monthly
# of variables: 169

Variable	Description	Unit	Format	Source	Route(s)
Dependent Variable					
TD1	TCE Ras Tanura – LOOP - VLCC route	$/day	Average	Clarksons SIN	-
TD3	TCE Ras Tanura – Chiba - VLCC route	$/day	Average	Clarksons SIN	-
TD7	TCE Sullom Voe – Wilhelmshaven - Aframax route	$/day	Average	Clarksons SIN	-
TD12	TCE Antwerp – Houston - LR1 route	$/day	Average	Clarksons SIN	-
TC1	TCE Ras Tanura – Chiba - LR2 route	$/day	Average	Clarksons SIN	-
TC2	TCE Rotterdam – New York - MR route	$/day	Average	Clarksons SIN	-
Panel A: Supply					
Fleet Size:					
VLCC_mdwt	Fleet size, VLCC	mDWT	SoM	Clarksons SIN	TD1,TD3
Afra_mdwt	Fleet size, Aframax	mDWT	SoM	Clarksons SIN	TD7
Pana_mdwt	Fleet size, Panamax	mDWT	SoM	Clarksons SIN	TD12
LR2_mdwt	Fleet size, LR2	mDWT	SoM	Clarksons SIN	TC1
MR_mdwt	Fleet size, MR	mDWT	SoM	Clarksons SIN	TC2
VLCC_yy	Fleet growth, VLCC	%	Yr/Yr	Clarksons SIN	TD1,TD3
Afra_yy	Fleet growth, Aframax	%	Yr/Yr	Clarksons SIN	TD7
Pana_yy	Fleet growth, Panamax	%	Yr/Yr	Clarksons SIN	TD12
LR2_yy	Fleet growth, LR2	%	Yr/Yr	Clarksons SIN	TC1
MR_yy	Fleet growth, MR	%	Yr/Yr	Clarksons SIN	TC2
VLCC_new	Newbuilding contracting, VLCC	mDWT	Sum	Clarksons SIN	TD1,TD3
Afra_new	Newbuilding contracting, Aframax	mDWT	Sum	Clarksons SIN	TD7
Pana_new	Newbuilding contracting, Panamax	mDWT	Sum	Clarksons SIN	TD12
LR2_new	Newbuilding contracting, LR2	mDWT	Sum	Clarksons SIN	TC1
MR_new	Newbuilding contracting, MR	mDWT	Sum	Clarksons SIN	TC2
VLCC_order	Orderbook, VLCC	mDWT	SoM	Clarksons SIN	TD1,TD3
Afra_order	Orderbook, Aframax	mDWT	SoM	Clarksons SIN	TD7
Pana_order	Orderbook, Panamax	mDWT	SoM	Clarksons SIN	TD12
LR2_order	Orderbook, LR2	mDWT	SoM	Clarksons SIN	TC1
MR_order	Orderbook, MR	mDWT	SoM	Clarksons SIN	TC2
VLCC_order_fleet	Orderbook as percentage of fleet, VLCC	%	SoM	Clarksons SIN	TD1,TD3
Afra_order_fleet	Orderbook as percentage of fleet, Aframax	%	SoM	Clarksons SIN	TD7
Pana_order_fleet	Orderbook as percentage of fleet, Panamax	%	SoM	Clarksons SIN	TD12
LR2_order_fleet	Orderbook as percentage of fleet, LR2	%	SoM	Clarksons SIN	TC1
MR_order_fleet	Orderbook as percentage of fleet, MR	%	SoM	Clarksons SIN	TC2
VLCC_deliveries	Vessel deliveries, VLCC	mDWT	Sum	Clarksons SIN	TD1,TD3
Afra_deliveries	Vessel deliveries, Afrmax	mDWT	Sum	Clarksons SIN	TD7
Pana_deliveries	Vessel deliveries, Panamax	mDWT	Sum	Clarksons SIN	TD12
LR2_deliveries	Vessel deliveries, LR2	mDWT	Sum	Clarksons SIN	TC1
MR_deliveries	Vessel deliveries, MR	mDWT	Sum	Clarksons SIN	TC2
VLCC_down	Down adj. of fleet (Demolitions & Removals), VLCC	mDWT	Sum	Clarksons SIN	TD1,TD3
Afra_down	Down adj. of fleet (Demolitions & Removals), Aframax	mDWT	Sum	Clarksons SIN	TD7,TC1
Pana_down	Down adj. of fleet (Demolitions & Removals), Panamax	mDWT	Sum	Clarksons SIN	TD12
LR2_down	Down adj. of fleet (Demolitions), LR2	mDWT	Sum	Clarksons SIN	
MR_down	Down adj. of fleet (Demolitions & Removals), MR	mDWT	Sum	Clarksons SIN	TC2
Fleet Age:					
VLCC_age	Average age of fleet, VLCC	Yr	SoM	Clarksons SIN	TD1,TD3
Afra_age	Average age of fleet, Aframax	Yr	SoM	Clarksons SIN	TD7,TC1
Pana_age	Average age of fleet, Panamax	Yr	SoM	Clarksons SIN	TD12
MR_age	Average age of fleet, MR	Yr	SoM	Clarksons SIN	TC2
Vessel Prices:					
VLCC_price	Newbuilding price quote, VLCC c.320k dwt	$m	EoM	Clarksons SIN	TD1,TD3
Afra_price	Newbuilding price quote, Afrmax c.115k dwt	$m	EoM	Clarksons SIN	TD7,TC1
Pana_price	Newbuilding price quote, Panamax c.75k dwt	$m	EoM	Clarksons SIN	TD12
MR_price	Newbuilding price quote, MR c.50k dwt	$m	EoM	Clarksons SIN	TC2
VLCC_SP	10Yr old secondhand price quote (S&P), VLCC c.300k dwt	$m	EoM	Clarksons SIN	TD1,TD3
Afra_SP	10Yr old secondhand price quote (S&P), Aframax c.105k dwt	$m	EoM	Clarksons SIN	TD7,TC1
Pana_SP	5Yr old secondhand price quote (S&P), Panamax c.73k dwt	$m	EoM	Clarksons SIN	TD12
MR_SP	10Yr old secondhand price quote (S&P), MR c.37k dwt	$m	EoM	Clarksons SIN	TC2
VLCC_demo_price	Demolition price, VLCC	$/ldt	EoM	Clarksons SIN	TD1,TD3
Afra_demo_price	Demolition price, Aframax	$m	EoM	Clarksons SIN	TD7,TC1
Pana_demo_price	Demolition price, Panamax	$m	EoM	Clarksons SIN	TD12
MR_demo_price	Demolition price, MR	$m	EoM	Clarksons SIN	TC2

Panel B: Demand

Oil Demand:					
US_dem	Oil demand (total products), United States	mbbl/d	Average	JODI	TD1,TD12,TC2
Jap_dem	Oil demand (total products), Japan	mbbl/d	Average	JODI	TD3,TC1
Eur_dem	Oil demand (total products), EU-4 (Germ., France, UK, Italy)	mbbl/d	Average	JODI	TD7,TD12,TC2
US_fdem	Fuel Oil demand, United States	mbbl/d	Average	JODI	TD12
Eur_fdem	Fuel Oil demand, EU-4 (Germany, France, UK, Italy)	mbbl/d	Average	JODI	TD12
Jap_cdem	Clean oil demand (Naphtha, Gasoline, Kerosene, Diesel), Japan	mbbl/d	Average	JODI	TC1
Eur_cdem	Clean oil demand (Naph., Gaso., Kero., Dies.), EU-4	mbbl/d	Average	JODI	TC2
US_cdem	Clean oil demand (Naph., Gaso., Kero., Dies.), United States	mbbl/d	Average	JODI	TC2
Oil Import:					
US_sea_imp	Crude imports (*Seaborne*), United States	mbbl/d	Average	Clarksons SIN	TD1
US_imp	Crude imports, United States	mbbl/d	Average	JODI	TD1
SAUS_imp	Crude imports, United States from Saudi Arabia	mbbl/d	Average	EIA	TD1
SAPADD3_imp	Crude imports, Saudi Arabia to US PADD3	mbbl/d	Average	EIA	TD1
PADD3_imp	Crude imports, PADD3	mbbl/d	Average	EIA	TD1
Jap_imp	Crude imports (*Seaborne*), Japan	mbbl/d	Average	Clarksons SIN	TD3
Eur_imp	Crude imports, EU-4 (Germany, France, UK, Italy)	mbbl/d	Average	Clarksons SIN	TD7
Ge_imp	Crude imports, Germany	mbbl/d	Average	JODI	TD7
Major_imp	Crude imports, US, EU-4, Japan	mbbl/d	Average	Clarksons SIN	All
Chi_imp	Crude imports (*Seaborne*), China	mbbl/d	Average	Clarksons SIN	All
Ind_imp	Crude imports (*Seaborne*), India	mbbl/d	Average	Clarksons SIN	All
Bel_fimp	Dirty products imports, Belgium	mbbl/d	Average	JODI	TD12
US_fimp	Dirty products imports, United States	mbbl/d	Average	JODI	TD12
PADD3_fimp	Dirty products imports, United States PADD3	mbbl/d	Average	EIA	TD12
BelUS_fimp	Dirty products imports (*unfinished oils*), Belgium to US	mbbl/d	Average	EIA	TD12
BelPADD3_fimp	Dirty products imports (*unfinished oils*), Belgium to US PADD3	mbbl/d	Average	EIA	TD12
Jap_cimp	Clean products imports, Japan	mbbl/d	Average	JODI	TC1
Ne_cimp	Clean products imports, Netherlands	mbbl/d	Average	JODI	TC2
US_cimp	Clean products imports, United States	mbbl/d	Average	JODI	TC2
NeUS_cimp	Clean products imports (mainly Gasoline), Netherlands to US	mbbl/d	Average	EIA	TC2
NePADD1_cimp	Clean products imports, Netherlands to US PADD1	mbbl/d	Average	EIA	TC2
Oil Export:					
AG_exp	Crude exports, Arabian Gulf (incl. Red Sea from 2014)	mbbl/d	Average	Clarksons SIN	TD1,TD3
NS_exp	Crude exports, North Sea (UK and Norway)	mbbl/d	Average	JODI	TD7
SA_exp	Crude exports, Saudi Arabia	mbbl/d	Average	JODI	TD1,TD3
Bel_fexp	Dirty products exports, Belgium	mbbl/d	Average	JODI	TD12
US_fexp	Dirty products exports, United States	mbbl/d	Average	JODI	TD12
USBel_fexp	Dirty products exports, United States to Belgium	mbbl/d	Average	EIA	TD12
SA_cexp	Clean products exports, Saudi Arabia	mbbl/d	Average	JODI	TC1
Ne_cexp	Clean products exports, Netherlands	mbbl/d	Average	JODI	TC2
US_cexp	Clean products exports, United States	mbbl/d	Average	JODI	TC2
USNe_exp	Products exports (total products), United States to Netherlands	mbbl/d	Average	EIA	TC2
Vessel Fixtures:					
VLCC_fix	VLCC single voyage fixtures (total)	#	Sum	Clarksons SIN	TD1,TD3
VLCC_fix_west	VLCC single voyage fixtures, Arabian Gulf - West	#	Sum	Clarksons SIN	TD1
VLCC_fix_east	VLCC single voyage fixtures, Arabian Gulf - East	#	Sum	Clarksons SIN	TD3
VLCC_fix_jap	VLCC single voyage fixtures, Arabian Gulf - Japan	#	Sum	Clarksons SIN	TD3
VLCC_due	VLCCs due this month, Arabian Gulf	#	Sum	Clarksons SIN	TD1,TD3
Afra_fix	Aframax single voyage fixtures (total)	#	Sum	Clarksons SIN	TD7,TC1
Afra_fix_sum	Aframax sing. v.fix, sum of AG-Cont, Baltic-Med, FarEast-Cont	#	Sum	Clarksons SIN	TD7
Afra_fix_east	Aframax single voyage fixtures, Arabian Gulf - East	#	Sum	Clarksons SIN	TC1
Afra_fix_US	Aframax single voyage fixtures, Arabian Gulf - US	#	Sum	Clarksons SIN	TC1
Pana_fix	Panamax single voyage fixtures (total)	#	Sum	Clarksons SIN	TD12
Pana_fix_US	Panamax single voyage fixtures, Mediterranean - US	#	Sum	Clarksons SIN	TD12
MR_fix	Handysize single voyage fixtures, sum of UKC-USG, UKC-USAC	#	Sum	Clarksons SIN	TC2
MR_fix_US	Handysize single voyage fixtures (total)	#	Sum	Clarksons SIN	TC2
Crude Oil Production:					
W_prod	Crude oil production, World total	mbbl/d	Average	Clarksons SIN	All
ME_prod	Crude oil production, Middle East	mbbl/d	Average	Clarksons SIN	TD1,TD3
NA_prod	Crude oil production, North America	mbbl/d	Average	Clarksons SIN	TD1
US_prod	Crude oil production, United States	mbbl/d	Average	JODI	TD1,TD12,TC2
SA_prod	Crude oil production, Saudi Arabia	mbbl/d	Average	JODI	TD1,TD3,TC1
NS_prod	Crude oil production, North Sea (UK and Norwegian shelf)	mbbl/d	Average	Clarksons SIN	TD7
OPEC_prod	Crude oil production, OPEC	mbbl/d	Average	Clarksons SIN	All
Refinery Output:					
US_fout	Fuel Oil refinery output, United States	mbbl/d	Average	JODI	TD12
Bel_fout	Fuel Oil refinery output, Belgium	mbbl/d	Average	JODI	TD12
SA_cout	Clean products refinery output, Saudi Arabia	mbbl/d	Average	JODI	TC1
Ne_cout	Clean products refinery output, Netherlands	mbbl/d	Average	JODI	TC2
US_cout	Clean products refinery output, United States	mbbl/d	Average	JODI	TC2
Refinery Utilisation:					
PADD3_refuti	Refinery utilisation, PADD3	%	Average	EIA	TD12
PADD1_refuti	Refinery utilisation, PADD1	%	Average	EIA	TC2

Panel C: Economic & Non-fundamental

Gross Domestic Product:					
GDP_w	Weighted geometric mean of real GDP indices	Index	-	Quandl	All
Time Charter Rate:					
VLCC_1tc	1Yr Time charter rate, VLCC c.310k dwt	$/day	Average	Clarksons SIN	TD1,TD3

Afra_1tc	1Yr Time charter rate, Aframax c.110k dwt	$/day	Average	Clarksons SIN	TD7
Pana_1tc	1Yr Time charter rate, Panamax c.80k dwt	$/day	Average	Clarksons SIN	TD12
LR2_1tc	1Yr Time charter rate, LR2 c.115k dwt	$/day	Average	Clarksons SIN	TC1
MR_1tc	1Yr Time charter rate, MR c.48k dwt	$/day	Average	Clarksons SIN	TC2
MR_3tc	3Yr Time charter rate, MR c.37k dwt	$/day	Average	Clarksons SIN	TC2
Exchange Rate:					
USD_SAR	Exchange rate, USD/SAR	Ratio	EoM	Bloomberg	TD1,TD3,TC1
USD_Pound	Exchange rate, USD/Pound	Ratio	Average	Clarksons SIN	TD7
USD_Eur	Exchange rate, USD/Eur	Ratio	Average	Clarksons SIN	TD7,TD12,TC2
Yen_USD	Exchange rate, Yen/USD	Ratio	Average	Clarksons SIN	TD3,TC1
SDR_USD	IMF index, SDR is based on USD, Euro, Yen, Pound, Yuan	Index	Average	Clarksons SIN	All
Euro_index	Index, Euro	Index	Average	Clarksons SIN	All
USD_index	Index, USD	Index	Average	FRED	All
Consumer Price Index & Money Supply:					
US_CPI	Consumer price index, United States	Index	-	BLS	TD1,TD12,TC2
Jap_CPI	Consumer price index, Japan	Index	-	FRED/OECD	TD3,TC1
Eur_CPI	Consumer price index, 19 European countries	Index	-	FRED/Eurostat	TD7
US_money	M1 money stock/supply, United States	$Bn	-	FRED	TD1,TD12,TC2
Jap_money	M1 money stock/supply, Japan	YenTn	-	FRED/OECD	TD3,TC1
Interest Rate:					
LIBOR	Interest rate, 3-month USD based LIBOR	%	EoM	FRED	All
LIBOR_Yen	Interest rate, 3-month Yen based LIBOR	%	EoM	FRED	All
LIBOR_Eur	Interest rate, 3-month Eur based LIBOR	%	EoM	FRED	All
Industrial Production:					
Ind_US	Industrial production, United States	%	Yr/Yr	Clarksons SIN	All
Ind_Jap	Industrial production, Japan	%	Yr/Yr	Clarksons SIN	TD3,TC1
Ind_Eur	Industrial production, Europe	%	Yr/Yr	Clarksons SIN	TD7
Ind_OECD	Industrial production, OECD	%	Yr/Yr	Clarksons SIN	TD12,TC2
Ind_China	Industrial production, China	%	Yr/Yr	Clarksons SIN	All
Ind_India	Industrial production, India	%	Yr/Yr	Clarksons SIN	All
Crude Oil & Oil Products Price:					
Brent	Oil price, Brent crude	$/bbl	Average	Clarksons SIN	All
Brent_forw	Brent crude forward curve (6month − 1month)	$/bbl	EoM	Bloomberg	All
WTI	Oil price, WTI crude	$/bbl	EoM	EIA	All
Dubai	Oil price, Dubai crude	$/bbl	Average	IMF	All
Oil_price_index	Index, Brent, WTI, Dubai	Index	Average	IMF	All
US_gaso	Gasoline conventional spot FOB, US Gulf Coast	$/gallon	-	EIA	TC2
Bunker Price:					
Bunker_Jap	Bunker price 380CST, Japan	$/tonne	Average	Clarksons SIN	All
Bunker_Phil	Bunker price 380CST, Philadelphia US	$/tonne	Average	Clarksons SIN	All
Shipping Index:					
ClarkSea	ClarkSea shipping index, weighted of major vessel classes	$/day	Average	Clarksons SIN	All
ClarkAve	Clarksons Average Tanker Earnings	$/day	Average	Clarksons SIN	All
BDTI	Baltic Exchange Dirty Tanker Index (BDTI)	Index	Average	Clarksons SIN	All
BCTI	Baltic Exchange Clean Tanker Index (BCTI)	Index	Average	Clarksons SIN	All
Stock Index:					
Tadawul	Tadawul stock index, Saudi Arabia	Index	EoM	Bloomberg	TD1,TD3,TC1
Nikkei	Nikkei stock index, Japan	Index	EoM	Bloomberg	TD3,TC1
SP500	S&P500 stock index, United States	Index	EoM	Yahoo	All
MSCI_w	Stock index, MSCI World	Index	EoM	Bloomberg	All
MSCI_e	Stock index, MSCI Emerging	Index	EoM	Bloomberg	All
VIX	S&P500 volatility index, VIX	Index	EoM	Yahoo	All

Comments from data sources:
Panel A:
*MR tankers (SIN): Most MR time series' include chemical tankers.
*NB, S&P and Demo. prices (SIN): Long run historical series based on last quotes for each vessel class. Size (dwt) may vary over the time series period.
*S&P prices (SIN): Between October 2008 and January 2010, Clarksons Research did not publish benchmark values and users should be aware that this was a period of transition in the Sale and Purchase markets, characterised by spells of rapidly changing price levels, low levels of sales activity and a wide spread of price ideas. During this period, the data should be treated with caution as confidence limits will vary over time and between sectors.
*Demolition prices (SIN): Series based on highest quoted demolition price of vessels within an appropriate age range each month.
Panel B:
*Oil demand (JODI): Deliveries or sales to domestic consumption plus Refinery Fuel plus International Marine and Aviation Bunkers.
*Clean products (JODI): Naphtha, Gasoline, Kerosene, Diesel.
*Fuel Oil/Dirty products (JODI): Heavy residual oil/boiler oil, including bunker oil.
*Dirty products (EIA): Sum of Distillate Fuel Oil, Residual Fuel Oil, Lubricants and Petroleum Coke.
*Crude imports PADD3 (EIA): Excluding SPR (strategic petro. reserves.)
*Unfinished oils (EIA): partial refining of crude oil, and include naphthas and lighter oils, kerosene and gas oils, and residuum.
*Imp.&Exp. (JODI): Goods having physically crossed the national boundaries, excluding transit trade, intl. marine and aviation bunkers.
*Fixtures (SIN): Note that weeks do not fall discretely into months. Consequently, the monthly timeseries does not equal the sum of weekly data points during the month. The weekly frequency represents fixtures reported in the seven days prior to the datum specified.
*Refinery output (JODI): Refinery output of finished products only.
Panel C:
*GDP (Quandl): Weighted g.mean of real GDP indices for various countries, weights equal to each country's share of world *oil consumption.*
*1Yr TC rate (SIN): Long run historical series based on average TC rates for each vessel class. Size (dwt) may vary over the time series period.
*CPI (BLS&FRED): US - all items in U.S. city average. Japan and Europe - all items.
*Money supply (FRED): M1 includes funds that are readily accessible for spending.
*LIBOR (FRED): LIBOR 3-month interest rate is the average interest rate at which a selection of banks in London are prepared to lend to one another in USD/Yen/Eur with a maturity of 3 months.
*MSCI World & MSCI Emerging (Bloomberg): World - a broad global equity benchmark that represents large and mid-cap equity performance across 23 developed markets countries. Emerging - large and mid cap representation across 24 emerging markets countries.

A.2 Data List per Route

Table A.2: Brief overview of data input per route - Variable abbreviations as used in the empirical work (see also Appendix A.1)

TD1	TD3	TD7	TD12	TC1	TC2	Common (All)
US_dem	Jap_dem	Eur_dem	US_dem	Jap_cdem	US_dem	Major_imp
AG_exp	AG_exp	NS_exp	US_fdem	Jap_dem	Eur_dem	Chi_imp
VLCC_fix_west	VLCC_fix_east	Afra_fix_sum	Eur_fdem	SA_cexp	Eur_cdem	Ind_imp
VLCC_fix	VLCC_fix_jap	Afra_fix	Eur_dem	Afra_fix_east	US_cdem	OPEC_prod
VLCC_due	VLCC_fix	Eur_imp	Bel_fexp	Afra_fix_US	Ne_cexp	W_prod
US_sea_imp	VLCC_due	Ge_imp	US_fexp	Afra_fix	US_cexp	SDR_USD
US_imp	Jap_imp	NS_prod	USBel_fexp	Jap_cimp	USNe_exp	Euro_index
SAUS_imp	ME_prod	USD_Pound	Pana_fix_US	SA_prod	MR_fix_US	USD_index
SAPADD3_imp	SA_prod	USD_Eur	Pana_fix	SA_cout	MR_fix	GDP_w
PADD3_imp	SA_exp	Eur_CPI	Bel_fimp	Jap_money	Ne_cimp	Ind_China
ME_prod	Jap_money	Ind_Eur	US_fimp	USD_SAR	US_cimp	Ind_US
NA_prod	USD_SAR	Afra_1tc	PADD3_fimp	Yen_USD	US_prod	Ind_India
US_prod	Yen_USD	Afra_age	US_prod	Jap_CPI	PADD1_refuti	LIBOR
SA_prod	Jap_CPI	Afra_down	PADD3_refuti	Ind_Jap	Ne_cout	LIBOR_Yen
SA_exp	Ind_Jap	Afra_deliveries	US_fout	LR2_1tc	US_cout	LIBOR_Eur
USD_SAR	VLCC_1tc	Afra_mdwt	Bel_fout	Tadawul	NeUS_cimp	Bunker_Phil
US_CPI	Tadawul	Afra_yy	BelUS_fimp	Nikkei	NePADD1_cimp	Bunker_Jap
US_money	Nikkei	Afra_new	BelPADD3_fimp	Afra_age	US_money	Brent_forw
VLCC_1tc	VlCC_age	Afra_order	US_money	Afra_down	USD_Eur	Brent
Tadawul	VLCC_down	Afra_order_fleet	USD_Eur	LR2_deliveries	US_CPI	WTI
VLCC_age	VLCC_deliveries	Afra_price	US_CPI	LR2_mdwt	Ind_OECD	Dubai
VLCC_down	VLCC_mdwt	Afra_SP	Ind_OECD	LR2_yy	MR_1tc	Oil_price_index
VLCC_deliveries	VLCC_yy	Afra_demo_price	Pana_1tc	LR2_new	MR_3tc	ClarkSea
VLCC_mdwt	VLCC_new		Pana_age	LR2_order	US_gaso	ClarkAve
VLCC_yy	VLCC_order		Pana_down	LR2_order_fleet	MR_age	BDTI
VLCC_new	VLCC_order_fleet		Pana_deliveries	Afra_price	MR_down	BCTI
VLCC_order	VLCC_price		Pana_mdwt	Afra_SP	MR_deliveries	MSCI_w
VLCC_order_fleet	VLCC_SP		Pana_yy	Afra_demo_price	MR_mdwt	MSCI_e
VLCC_price	VLCC_demo_price		Pana_new		MR_yy	VIX
VLCC_SP			Pana_order		MR_new	SP500
VLCC_demo_price			Pana_order_fleet		MR_order	
			Pana_price		MR_order_fleet	
			Pana_SP		MR_price	
			Pana_demo_price		MR_SP	
					MR_demo_price	

A.3 Data Sources

Below, we have copied in a brief description of each source that we have used, in accordance to how they are presented on their respective webpages. We have, to the extent possible, tried to find reliable and publicly available sources. However, subscriptions are needed to access the following sources: *Clarksons SIN and WFR*, and *Bloomberg Terminal*. Clarksons Shipping Intelligence Network (SIN) is the most widely used data source for actors in the shipping market, and holds information and data that could not be reached elsewhere. We are extremely grateful for Clarksons Platou's contribution by granting us access to the Holy Grail of shipping intelligence.

Shipping Intelligence Network (SIN) & World Fleet Register (WFR) by Clarksons Research Services Limited (CRSL):

«Shipping Intelligence Network (SIN) provides access to the comprehensive range of data collected and published by Clarksons Research, including the latest information on the shipping markets at a glance, easily downloadable versions of our wide range of market reports, comprehensive fleet and orderbook listings and thousands of timeseries and graphs of key commercial indicators all updated regularly by our leading industry analysts.» *Note that SIN also gathers data from publicly available sources (e.g., EIA, IEA, OECD).

"...World Fleet Register (WFR) is the market leading online vessel reference tool from Clarksons Research. Updated daily, WFR provides comprehensive, authoritative and timely information on over 150,000 ships...»

URL: sin.clarksons.net & clarksons.net/wfr2/

Joint Organisations Data Initiative (JODI) - JODI Oil World Database:

«The Joint Organisations Data Initiative is a concrete outcome of the producer-consumer energy dialogue. The initiative relies on the combined efforts of producing and consuming countries and the seven JODI partner organisations to build the timely, comprehensive, and sustainable energy data provision architecture which is a prerequisite for stable energy commodity markets. JODI Oil partners include APEC, Eurostat, IEA, OLADE, OPEC and UNSD»

URL: jodidata.org/oil/

U.S. Energy Information Administration (EIA):

«The U.S. Energy Information Administration (EIA) is the statistical and analytical agency within the U.S. Department of Energy. EIA collects, analyzes, and disseminates independent and impartial energy information to promote sound policymaking, efficient markets, and public understanding of energy and its interaction with the economy and the environment.

EIA's data and analyses are widely used by federal and state agencies, business & industry, media, researchers, consumers, financial, international, students, and educators.»

URL: eia.gov

Quandl - Financial, Economic and Alternative Data:

«Quandl delivers financial, economic and alternative data to over 250,000 people worldwide. Quandl offers essential financial and economic data alongside a suite of unique, alpha-generating alternative datasets. Quandl's customers include the world's top hedge funds, asset managers and investment banks.»

URL: quandl.com

Bloomberg Professional Services - Bloomberg Terminal:

«The Bloomberg Terminal brings together real-time data on every market, breaking news, in-depth research, powerful analytics, communications tools ... »

URL: bloomberg.com/professional/

Federal Reserve Economic Data (FRED):

«This site offers a wealth of economic data and information to promote economic education and enhance economic research. The widely used database FRED is updated regularly and allows 24/7 access to regional and national financial and economic data.»

URL: fred.stlouisfed.org

U.S. Bureau of Labor Statistics (BLS):

«The Bureau of Labor Statistics (BLS) of the U.S. Department of Labor is the principal federal agency responsible for measuring labor market activity, working conditions, and price changes in the economy. Its mission is to collect, analyze, and disseminate essential economic information to support public and private decision making. As an independent statistical agency, BLS serves its diverse user communities by providing products and services that are accurate, objective, relevant, timely, and accessible.»

URL: www.bls.gov/home.htm

OECD Data:

«...We measure productivity and global flows of trade and investment. We analyse and compare data to predict future trends.»

URL: data.oecd.org

Eurostat:

«Eurostat is the statistical office of the European Union situated in Luxembourg. Its mission is to provide high quality statistics for Europe.»

URL: ec.europa.eu/eurostat/web/main/home

International Monetary Fund (IMF):

«The IMF publishes a range of time series data on IMF lending, exchange rates and other economic and financial indicators. Manuals, guides, and other material on statistical practices at the IMF, in member countries, and of the statistical community at large are also available.»

URL: imf.org/en/data

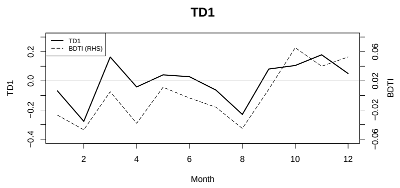
Yahoo Finance:

«At Yahoo Finance, you get free stock quotes, up-to-date news, portfolio management resources, international market data, social interaction and mortgage rates ...»

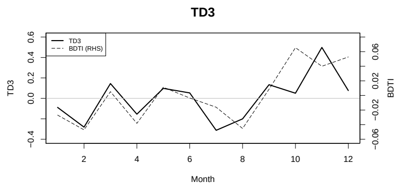
URL: finance.yahoo.com

A.4 Route Seasonality

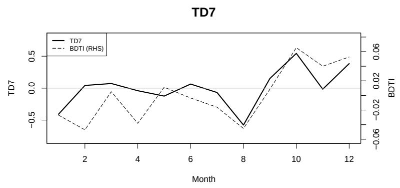
Seasonal components for each individual route are given in Figure A.1, alongside seasonal components for their respective Baltic Exchange index.



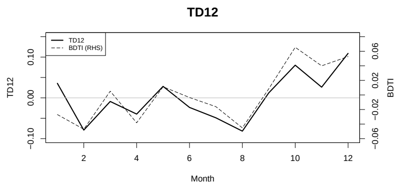
(a) Seasonal components TD1



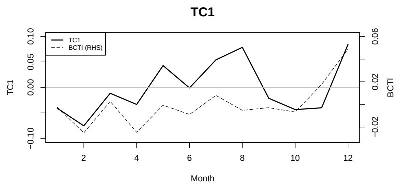
(b) Seasonal components TD3



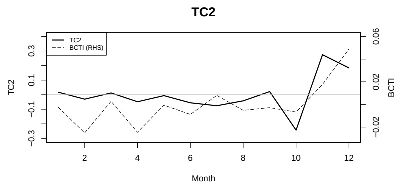
(c) Seasonal components TD7



(d) Seasonal components TD12



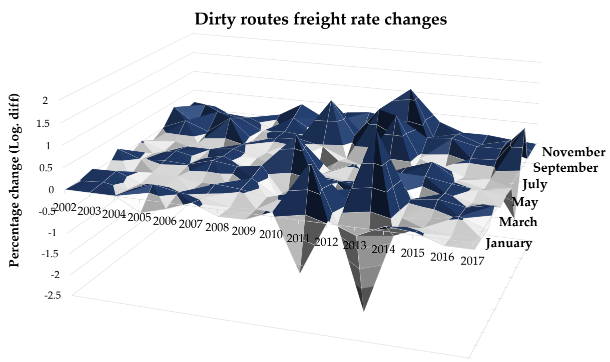
(e) Seasonal components TC1



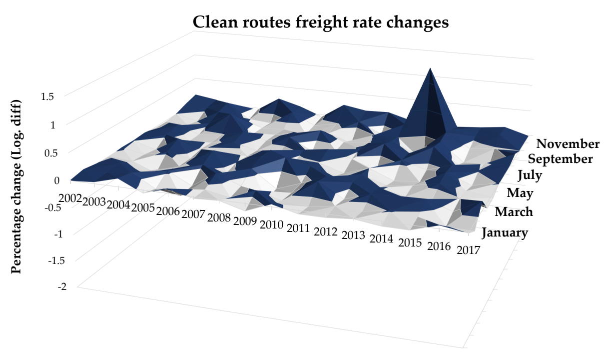
(f) Seasonal components TC2

Figure A.1: Individual seasonality components for each route. Calculated as the mean percentage change within a month for all observations over the entire time series period.

In Figure A.2, freight rate changes are represented on an average level for dirty tanker (TD1, TD3, TD7, TD12) and clean tanker (TC1, TC2) routes. Serves as an illustration of the existence of higher seasonal irregularity in dirty tanker routes versus clean tanker routes. *Seasonal components from Figure 7.1 thus equals the average of each month across all years of the plots in Figure A.2.



(a) Dirty routes

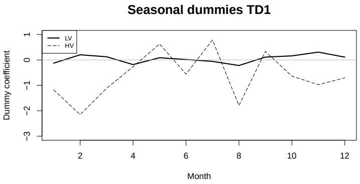


(b) Clean routes

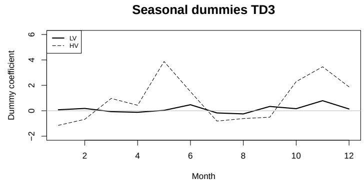
Figure A.2: Freight rate changes on average level for our selection of dirty tanker (TD1, TD3, TD7, TD12) and clean tanker (TC1, TC2) routes. Blue and gray levels indicate positive and negative freight rate changes, respectively.

A.5 Seasonality from Modelling

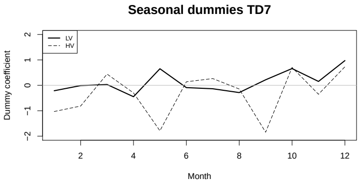
Seasonal dummy coefficients for each individual route are given in Figure A.3, alongside seasonal components for their respective Baltic Exchange index.



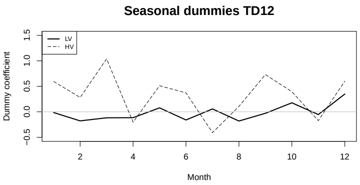
(a) Seasonal components TD1



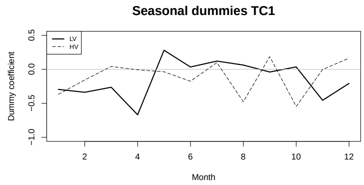
(b) Seasonal components TD3



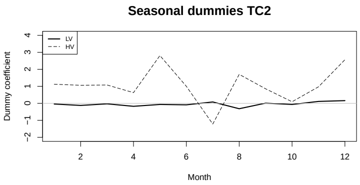
(c) Seasonal components TD7



(d) Seasonal components TD12



(e) Seasonal components TC1



(f) Seasonal components TC2

Figure A.3: Individual seasonality dummy coefficients for each route. HV: high volatility regime (stapled line). LV: low volatility regime. See, further, Chapter 8.

A.6 BDTI and BCTI

In Table A.3, a complete list of the routes included in the dirty (BDTI), and clean (BCTI) Baltic Exchange index is given. Routes in bold are relevant for this thesis.

Table A.3: Route constituents of the Baltic Exchange indices for dirty and clean tanker routes (Kavussanos and Visvikis, 2016).

Baltic Exchange Indices		
Route	Vessel	Description
Panel A: Baltic Dirty Tanker Index (BDTI)		
TD1	VLCC	Ras Tanura (Saudi Arabia) - LOOP (US Gulf)
TD2	VLCC	Ras Tanura (Saudi Arabia) - Singapore
TD3	VLCC	Ras Tanura (Saudi Arabia) - Chiba (Japan)
TD6	Suezmax	Novorossiysk (Russia) - Augusta (USA)
TD7	Aframax	Sullom Voe (UK) - Wilhelmshaven (Germany)
TD8	Aframax	Mena al Ahmadi (Kuwait) - Singapore
TD9	Panamax	Puerto La Cruz (Venezuela) - Corpus Christi (USA)
TD12	Panamax	Antwerp (Belgium) - Houston (USA)
TD14	Aframax	Syria - Sydney (Australia)
TD15	VLCC	Bonny Offshore (Nigeria) - Ningbo (China)
TD17	Aframax	Primorsk (Russia) - Wilhelmshaven (Germany)
TD18	Handysize	Tallinn (Estionia) - Amsterdam (Netherlands)
TD19	Aframax	Ceyhan (Turkey) - Lavera (France)
TD20	Suezmax	West Africa - UK Continent/Rotterdam
TD21	Panamax	Mamonal (Colombia) - Houston (USA)
Panel B: Baltic Clean Tanker Index (BCTI)		
TC1	Aframax (LR2)	Ras Tanura (Saudi Arabia) - Chiba/Yokohama (Japan)
TC2	Handysize (MR)	Rotterdam (Netherlands) - New York (USA)
TC5	Panamax (LR1)	Ras Tanura (Saudi Arabia) - Yokohama (Japan)
TC6	Handysie (MR)	Skikda (Syria) - Lavera (France)
TC8	Panamax (LR1)	Jubail (Saudi Arabia) - Rotterdam (Netherlands)
TC9	Handysize (MR)	Primorsk (Russia) - Le Havre (France)
TC14	Handysize (MR)	Houston (USA) - Amsterdam (Netherlands)
TC15	Aframax (LR2)	MED - Far East (Trial)
TC16	Panamax (LR1)	Amsterdam (Netherlands) - Lome Offshore (Togo) (Trial)

A.7 Clarksons TCE Freight Rate Calculations

Tanker TCE freight rates in the SIN database is based on the calculation breakdown seen in Table A.4 (Clarksons Research Services Limited, 2017). In Section 6.2, a more detailed discussion of the calculations is given.

Table A.4: Breakdown of the calculation method of TCE tanker freight rates in Clarksons SIN (Clarksons Research Services Limited, 2017)

Tanker TCE Calculations		
Element	Description	Unit
Earnings (E):		
E = (R−C)/D	Average Earnings	$/day
Revenue (R):		
V*($*WS/100)*(1−CM)	Freight Revenue	$
Costs (C):		
(D−dl*SM)*L−dfo*FO−d$	Laden Fuel Oil Cost	$
(D−dl*SM)*B−dfo*FO−d$	Ballast Fuel Oil Cost	$
(D−dl*SM)*L−ddo*DO−d$	Laden mdo Cost	$
(D−dl*SM)*B−ddo*DO−d$	Ballast mdo Cost	$
P−dfo*FO−d$	Port Fuel Cost	$
P−ddo*DO−d$	Port mdo Cost	$
P−d$	Port Charges	$
C−d$	Canal Charges	$
Voyage Time (D):		
(D−db+D−dl)*SM+D−dp	Total Voyage Time	days
Key items:		
Earnings:		
E	Earnings	$/day
R	Revenue	$
C	Costs	$
D	Voyage Time	days
Revenue:		
V	Cargo Loaded	tonnes
WS	Worldscale Rate	rate
$	Worldscale Basic	$/tonne
CM	Commission	2.50%
Costs:		
D−dl	Days Laden	days
D−db	Days Ballast	days
D−dp	Days in Port	days
L−dfo, B−dfo	FO Cons.	m.t./day
L−ddo, B−ddo	DO Cons.	m.t./day
P−dfo, P−ddo	Port Cons.	tonnes
FO−d$	Cost FO	$/tonne
DO−d$	Cost DO	$/tonne
SM	Sea Margin	Fixed 5%
P−d$	Port Charges	$
C−d$	Canal Charges	$
Voyage time:		
D	Voyage Time	days
General:		
L	Laden	
B	Ballast	
P	In Port	
FO	Fuel Oil	
DO	Diesel Oil	

Appendix B Statistics and methods

B.1 $\mathbf{R^2}$ & $\mathbf{R^2_{adj}}$

The R^2 and R^2_{adj} are commonly used to evaluate the goodness of fit of a regression model. To define these, we first need a few other definitions.

The total sum of squares is the sum of the squared mean deviations of the dependent variable (Alexander, 2008b).

$$TSS = \sum_{t=1}^{T} (y_t - \bar{y})^2$$

The amount of variation in Y that is captured by the model, is measured by the explained sum of squares (Alexander, 2008b).

$$ESS = \sum_{t=1}^{T} (\bar{y}_t - \bar{y})^2$$

The regression R^2 is the square of the correlation between the fitted value $\hat{y}$ and y (Alexander, 2008a). The value lies between 0 and 1, where values closer to 1 signifies a good model fit.

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

An issue with the standard R^2 is that it does not take into account the parsimony of the model. R^2 will increase as one adds more variables to the model, but this does not imply a good fit. R^2_{adj} solves this issue by accounting for the degrees of freedom. It is therefore preferable to use for comparing models.

$$R^2_{adj} = 1 - \frac{RSS/(n-p-1)}{TSS/(n-1)}$$

B.2 Residual sum of squares

The sum of squared residuals, which is frequently called residual sum of squares (RSS), measures the deviation between the data and an estimation model. It can be written as

$$RSS = \sum_{t=1}^{T} e_t^2 = \sum_{t=1}^{T} (y_t - \bar{y}_t)^2$$

where Y is the observation, $\bar{Y}$ is the predicted value, i.e. the model, and e is the residual.

B.3 Akaike information criterion (AIC)

The Akaike criteria is used for model selection, and is a degree-of-freedom-penalized version of the mean-squared residual (Diebold, 2017). A lower AIC value is better.

$$AIC = e^{\left(\frac{2K}{T}\right)} \frac{\sum_{t=1}^{T} e_t^2}{T}$$

B.4 LASSO

The LASSO procedure attempts to force most features to zero, while only selecting a few best candidates. As seen in Doreswamy and Vastrad (2013) the LASSO estimator uses ℓ_1 penalized least squares basis to get a solution to the following optimization problem

$$\hat{\beta}_{LASSO} = \arg\min_{\beta} \|\boldsymbol{Y} - \boldsymbol{X\beta}\|_2^2 + \lambda \sum_{k=1}^{p} |\boldsymbol{\beta_k}|$$

where the added $\lambda \sum_{k=1}^{p} |\boldsymbol{\beta_k}|$ equals the ℓ_1-norm penalty, and λ can be considered a regularization parameter.

B.5 Stability selection with rand. LASSO

Large datasets might require methods to improve their understanding and interpretation. Stability selection is a way of achieving this. Stability selection is based on subsampling in combination with selection algorithms.

Stability selection enhances existing variable selection techniques. For variable selection methods, we usually have a tuning parameter, λ, which controls the amount of regularization applied. We would normally want to know, study and figure out this value.

For every value of $\lambda \in \Lambda$, one obtains a structure estimate $\hat{S}^\lambda = \{k; \hat{\beta}_k^\lambda \neq 0\}$, and a set of non-zero β coefficients. One then wishes to figure out if there exists a $\lambda \in \Lambda$ which makes $\hat{S}^\lambda$ equal to S with high probability.

Stability paths can be considered the probability for each variable to be selected, when randomly resampling from the dataset (Meinshausen and Bühlmann, 2009), as described below from the selection probability

$$\hat{\Pi}_k^\lambda = P^*(K \subseteq \hat{S}^\lambda(I))$$

where, I is a random subsample of size $\lfloor n/2 \rfloor$.

Normally, in variable selection, one would simply select one element from the set of models

$$\{\hat{S}^\lambda; \lambda \in \Lambda\}$$

With stability selection one performs subsampling many times, selecting the variables that frequently occur. Variables with a high selection probability are kept. A set of stable variables is defined as

$$\hat{S}^{stable} = \{k : \max_{\lambda \in \Lambda} \hat{\Pi}_k^\lambda \geqslant \pi_{thr}\}$$

where π_{thr} is regarded as a cutoff between 0 and 1, and with the set of regularization parameters Λ.

When estimating with stability selection, the choice of an initial regularization parameter Λ is not crucial (Meinshausen and Bühlmann, 2009). The problem of proper regularization is addressed with a generic subsampling approach.

Table B.1: The automatically estimated Arima models to be used as benchmarks for each route.

Route	Best fit Arima model
TD1	Arima(1,0,1)
TD3	Arima(3,0,1)
TD7	Arima(0,0,1)
TD12	Arima(2,0,2)
TC1	Arima(0,0,0)
TC2	Arima(1,0,2)

Randomized LASSO changes the penalty of λ to a randomly chosen value in the range $[\lambda, \lambda/\alpha]$ Randomized LASSO with a weakness $\alpha \in (0, 1]$ can be defined as follows

$$\hat{\beta}^{\lambda,W} = arg \min_{\beta \in \mathbb{R}} ||y - x\beta||_2^2 + \lambda \sum_{k=1}^{p} \frac{|\beta_k|}{W_k}$$

with W_k as i.i.d random variables in the range $[\alpha, 1]$, for $k = (1, \ldots, p)$, and $\hat{\beta}^{\lambda,W}$ as the randomized LASSO estimator (Meinshausen and Bühlmann, 2009).

B.6 ARIMA

An $ARIMA(p, d, q)$ is given by the following equation

$$(1 - \sum_{k=1}^{p} \alpha_k L^k)(1 - d)X_t = (1 + \sum_{k=1}^{q} \beta_k L^k)\epsilon_t$$

whereby $\alpha_1, \ldots, \alpha_p$ and $\beta_1, \ldots, \beta_p$ are related the autoregressive and moving average terms. L is the lag operator, X_t the variable, and $\epsilon_1, \ldots, \epsilon_k$ are the errors. As noted by Nau (2017):

- p is the number of autoregressive terms,
- d is the number of nonseasonal differences needed for stationarity, and
- q is the number of lagged forecast errors in the prediction equation.

The ARIMA models that will be considered in this thesis are listed in Table B.1.

B.7 Random walk

A random walk with a drift component is simply defined as

$$y_t = y_{t-1} + a + \epsilon_t$$

where ϵ_t is a white noise term with zero mean, and variance equal to one.

B.8 Mean model

A forecast based on the average of past data can be written as in the equation below, where $y_1, \ldots, y_T$ represents the historical data, and $\hat{y}_{T+h|T}$ is notation for the estimate of y_{T+h} based on this data.

$$\hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T$$

B.9 Multiple regression

The linear regression model can be written as

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_n x_{ni} + u_i$$

by which the variable y is the dependent variable, the x variables are the explanatory variables, and u_i is a stochastic error term (Gujarati, 2011).

To estimate the coefficients of the linear regression model, the method of ordinary least squares is appropriate. Here, one tries to minimize the sum of squared residuals by tuning the coefficients.

$$\hat{\beta} = \arg \min_{\beta} RSS$$

The result is estimated parameter values for the linear model.

Statistical Tests

B.10 Augmented Dickey-Fuller (ADF) test

The augmented Dickey-Fuller test is based on the regression

$$\Delta X_t = \alpha + \beta X_{t-1} + \gamma_1 \Delta X_{t-1} + \cdots + \gamma_q \Delta X_{t-q} + \epsilon_t$$

This test is carried out similarly to a Dickey-Fuller test, whereby the test statistic is the t ratio on the estimated coefficient $\hat{\beta}$. However, the critical values are dependent upon the number of lags, q. (Alexander, 2008a).

Durbin-Watson statistic (DW)

The Durbin-Watson test statistic can be used to identify first order autocorrelation in the residuals. For this statistic, T is the number of observations and e_t is the residual at time t.

$$d = \frac{\sum_{t=2}^{T} (e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

The value of the statistic d lies in the range 0 to 4. Having a d equal to 2 signifies no autocorrelation.

B.11 Breusch-Godrey (BG) test

The Breusch-Godfrey test is a method of assessing the model residuals for serial correlation. It uses the errors from the model and derives a test statistic from these.

The model is fitted initially to obtain a set of sample residuals.

$$y_t = \beta_1 + \beta_2 x_{2t} + \cdots + \beta_n x_{nt} + u_t$$

The residuals are then regressed on the independent variables and on the lagged residuals.

$$u_t = \rho_0 + \rho_1 u_{t-1} + \cdots + \rho_k u_{t-k} + v_t$$

The test statistic is then estimated by multiplying the R^2 from the second regression with the number of observations in the set. The test statistic can then be evaluated against the applicable chi-squared distribution. Rejection of the null hypothesis indicates serial correlation in the residuals.

$$H_0 : \rho_0 + \rho_1 + \rho_2 + \cdots + \rho_k = 0$$
$$H_1 : \rho_0 \cup \rho_1 \cup \rho_2 \cup \cdots \cup \rho_k \neq 0$$

B.12 Jarque-Bera test (JB)

The Jarque-Bera test is utilised to test for normality in a data set. The formula for the test is as such:

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] \sim \chi^2_2$$

Here S is the skewness coefficient, n is the sample size, and K the kurtosis effect. The closer the statistic is to zero, the closer the sample is to normal (Gujarati, 2011). This can be seen if inputting the normality values of skewness and kurtosis, $S = 0$ and $K = 3$ into the formula, which results in a statistic of zero.

B.13 Breusch-Pagan (BP) test

The Breusch-Pagan test is a test for heteroskedasticity in the residuals of a regression model. Under the null hypothesis, the residuals are assumed to have a constant variance, i.e. being homoskedastic (Gujarati, 2011).

Consider the following estimated linear regression.

$$y_t = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_n x_{nt} + u_t$$

Obtain $\hat{u}_t^2$, and the predicted $\hat{Y}_t$ values. Proceed to fit the squared residuals to the predicted values. The R^2 is then retained for the chi-squared statistic, TR^2.

$$\hat{u}_t^2 = \delta_0 + \delta_1 \hat{Y}_t$$

$\hat{Y}_n$ represents the predicted value of y from the original regression.

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1t} + \cdots + \hat{\beta}_n x_{nt}$$

The chi-squared statistic is estimated based on R^2 and the number of observations, with degrees of freedom equal to 1. The null hypothesis can then be evaluated.

$$TR^2 \sim \chi^2(m)$$

B.14 Diebold-Mariano (DM) test

Diebold and Mariano propose a test which allows one to evaluate the significance of apparent predictive superiority (Diebold, 2013). The null hypothesis is of equal expected loss. It relies on assumptions made on the loss difference of the forecast error. With a quadratic loss this would equate to $L(e_t) = e_t^2$. The loss difference between two forecasts can be written as $d_{12t} = L(e_{1t}) - L(e_{2t})$. The test only requires this difference to be covariance stationary, i.e.:

$$Assumpt. \begin{cases} E(d_{12t}) = \mu, \quad \forall t \\ Cov(d_{12t}, d_{12(t-\tau)}) = \gamma(\tau), \forall t \\ 0 < Var(d_{12t}) = \sigma^2 < \inf . \end{cases}$$

The null hypothesis corresponds to $E(d_{12t}) = 0$, which entails:

$$DM_{12} = \frac{\bar{d}_{12}}{\hat{\sigma}_{\bar{d}_{12}}} \xrightarrow{d} N(0,1)$$

where $\bar{d}_{12} = \frac{1}{T}\sum_{t=1}^{T} d_{12t}$ and $\hat{\sigma}_{\bar{d}_{12}}$ is an estimate of the standard deviation of $\bar{d}_{12}$.

Appendix C Oil Trade Details

Tanker Fleet Productivity

The tables below (Tables C.1 − C.3) represent a more detailed breakdown of the components that are included in the tanker fleet productivity table from Section 3.5 (Table 3.5).

Top 10 Countries in Oil Trade

An overview of top 10 countries engaged in oil trade based on various measures is given in Table C.4.

Table C.1: Tanker fleet productivity: Tonnes carried per dwt of tankers. Includes tankers above 10 k.dwt (Clarksons Research Services Limited, 2017). *NA: data not available.*

Year	Crude oil	Seaborne trade Oil products [m.tonnes]	Total oil	Crude	Tanker fleet Products [m.dwt, mid year]	Total fleet	Ratio: Tonnes carried per dwt of Crude	Products	Total
1990	1,132.75	415.41	1,548.16	200.62	NA	NA	5.64	NA	NA
1995	1,454.73	444.29	1,899.02	212.18	NA	NA	6.86	NA	NA
2000	1,676.18	561.91	2,238.09	220.82	53.56	274.38	7.60	10.50	8.16
2005	1,878.43	712.63	2,591.06	241.36	68.59	309.95	7.78	10.39	8.36
2010	1,871.87	881.52	2,753.39	295.57	116.56	412.13	6.33	7.56	6.68
2015	1,872.01	1,021.83	2,893.85	344.51	139.01	483.52	5.43	7.35	5.98
2017	2,003.58	1,098.06	3,101.64	382.07	156.00	538.07	5.24	7.04	5.76

Table C.2: Tanker fleet productivity: Tonne-miles carried per dwt of tankers. Includes tankers above 10 k.dwt (Clarksons Research Services Limited, 2017). *NA: data not available.*

Year	Crude oil	Seaborne trade Oil products [b.tonne-miles]	Total oil	Crude	Tanker fleet Products [m.dwt, mid year]	Total fleet	Ratio: 1k. tonne-miles carried per dwt of Crude	Products	Total
1990	5,345.1	1,262.7	6,607.8	200.62	NA	NA	26.64	NA	NA
1995	6,931.2	1,371.6	8,302.8	212.18	NA	NA	32.67	NA	NA
2000	8,050.4	1,571.9	9,022.5	220.82	53.56	274.38	36.46	29.35	35.07
2005	8,603.4	2,122.5	10,725.9	241.36	68.59	309.95	35.65	30.94	34.61
2010	8,686.8	2,600.5	11,287.3	295.57	116.56	412.13	29.39	22.31	27.39
2015	9,087.6	2,905.7	11,993.3	344.51	139.01	483.52	26.38	20.90	24.80
2017	10,101.6	3,141.7	13,243.3	382.07	156	538.07	26.44	20.14	24.61

Table C.3: Tanker fleet productivity: Average haul in miles - dividing tonne-miles by tonnes. Includes tankers above 10 k.dwt (Clarksons Research Services Limited, 2017).

Year	Crude oil	Average haul Oil products [miles]	Total oil
1990	4,718.7	3,039.7	4,268.2
1995	4,764.6	3,087.2	4,372.2
2000	4,802.8	2,797.4	4,031.3
2005	4,580.1	2,978.4	4,139.6
2010	4,640.7	2,950.0	4,099.4
2015	4,854.5	2,843.6	4,144.4
2017	5,041.8	2,861.1	4,269.8

Table C.4: Top 10 countries in oil trade. As of January 2018 (Joint Organisations Data Initiative(JODI)). *Remark: Chinese numbers must be treated with caution. Russia only included in the crude production entry.*

Rank #	**Crude prod.** (mbbl/d)	**Oil dem.** (mbbl/d)	**Crude imp.** (mbbl/d)	**Crude exp.** (mbbl/d)	**Products imp.** (mbbl/d)	**Products exp.** (mbbl/d)
1	Russia (10.4)	U.S. (20.8)	China (8.5)	Saudi Arabia (7.2)	Singapore (2.3)	U.S. (4.9)
2	Saudi Arabia (10.0)	China (11.5)	U.S. (8.0)	Iraq (3.8)	Netherlands (2.3)	Netherlands (2.4)
3	U.S. (9.9)	India (4.5)	India (4.8)	Canada (2.9)	U.S. (2.2)	Singapore (1.9)
4	Iraq (4.4)	Japan (4.3)	Japan (3.4)	Nigeria (2.0)	Japan (1.2)	Saudi Arabia (1.9)
5	China (3.8)	Korea (3.0)	Korea (3.2)	Angola (1.4)	Mexico (1.0)	China (1.6)
6	Canada (3.3)	Germany (2.5)	Thailand (2.1)	U.S. (1.4)	Korea (0.8)	India (1.5)
7	Brazil (2.6)	Canada (2.2)	Germany (1.9)	Brazil (1.3)	India (0.8)	Korea (1.4)
8	Nigeria (2.0)	Brazil (2.1)	Spain (1.6)	Norway (1.2)	Belgium (0.8)	Italy (0.6)
9	Mexico (1.9)	Saudi Arabia (2.1)	Italy (1.3)	Mexico (1.2)	Germany (0.8)	Belgium (0.5)
10	Venezuela (1.8)	Mexico (1.8)	France (1.2)	UK (0.7)	Saudi Arabia (0.8)	UK (0.5)

Appendix D Results and Plots

Table D.1: ADF test. The results from the augmented Dickey-Fuller tests performed on all transformed variables. The null hypothesis of a unit root is rejected, and all time series are shown to be stationary with a 5 percent significance. Most variables were transformed with logarithmic differencing.

Variable	ADF-stat	Variable	ADF-stat	Variable	ADF-stat
TD1	-6,903***	ME_prod	-7,181***	MSCI_w	-5,108***
TD3	-7,292***	NA_prod	-7,177***	MSCI_e	-5,682***
TD7	-7,058***	US_prod	-6,341***	VIX	-6,59***
TD12	-6,773***	SA_prod	-6,229***	SP500	-5,249***
TC1	-7,6***	NS_prod	-7,481***	VLCC_age	-8,275***
TC2	-8,012***	OPEC_prod	-5,441***	Afra_age	-3,813**
US_dem	-6,328***	W_prod	-5,996***	Pana_age	-4,532***
Jap_dem	-8,224***	Padd3_refuti	-8,057***	MR_age	-8,114***
Eur_dem	-8,161***	Padd1_refuti	-7,336***	VLCC_down	-3,61**
US_fdem	-7,829***	USGC_marg	-5,173***	Afra_down	-3,683**
Eur_fdem	-7,544***	NWE_marg	-7,532***	Pana_down	-4,785***
Jap_cdem	-8,64***	US_fout	-8,801***	LR2_down	-12,16***
Eur_cdem	-8,081***	Bel_fout	-7,993***	MR_down	-3,688**
US_cdem	-6,206***	SA_cout	-7,126***	VLCC_deliveries	-8,532***
AG_exp	-6,836***	Ne_cout	-8,789***	Afra_deliveries	-7,091***
NS_exp	-8,839***	US_cout	-7,851***	Pana_deliveries	-6,893***
Bel_fexp	-7,636***	USD_SAR	-7,569***	LR2_deliveries	-7,135***
US_fexp	-7,18***	Yen_USD	-5,682***	MR_deliveries	-7,54***
USBel_fexp	-5,393***	USD_Pound	-5,852***	VLCC_mdwt	-4,542***
SA_cexp	-7,507***	USD_Eur	-5,268***	VLCC_yy	-4,655***
Ne_cexp	-7,845***	SDR_USD	-5,595***	Afra_mdwt	-6,935***
US_cexp	-6,587***	Euro_index	-5,289***	Afra_yy	-6,173***
USNe_exp	-7,189***	USD_index	-5,077***	Pana_mdwt	-4,321***
SA_exp	-5,886***	GDP_w	-4,823***	Pana_yy	-3,685**
VLCC_fix_west	-8,266***	US_CPI	-6,788***	LR2_mdwt	-6,517***
VLCC_fix	-8,421***	Jap_CPI	-6,256***	LR2_yy	-4,03***
VLCC_due	-7,426***	Eur_CPI	-4,38***	MR_mdwt	-7,547***
VLCC_fix_east	-8,175***	Ind_US	-3,898**	MR_yy	-3,631**
VLCC_fix_jap	-8,539***	Ind_Jap	-4,682***	VLCC_new	-5,196***
Afra_fix_sum	-7,418***	Ind_Eur	-4,131***	Afra_new	-7,356***
Afra_fix	-6,028***	Ind_OECD	-4,186***	Pana_new	-4,651***
Pana_fix_US	-4,141***	Ind_China	-6,648***	LR2_new	-7,45***
Pana_fix	-7,675***	Ind_India	-5,797***	MR_new	-3,87**
Afra_fix_east	-8,518***	LIBOR	-4,252***	VLCC_order	-4,758***
Afra_fix_us	-4,897***	LIBOR_Yen	-11,485***	VLCC_order_fleet	-4,739***
MR_fix_US	-7,126***	LIBOR_Eur	-3,71**	Afra_order	-7,815***
MR_fix	-6,415***	Jap_money	-3,632**	Afra_order_fleet	-8,4***
US_sea_imp	-7,279***	US_money	-4,402***	Pana_order	-3,634**
US_imp	-7,336***	VLCC_1tc	-5,254***	Pana_order_fleet	-3,629**
SAUS_imp	-5,923***	Afra_1tc	-4,615***	LR2_order	-3,581**
SAPadd3_imp	-5,996***	Pana_1tc	-4,442***	LR2_order_fleet	-3,491**
PADD3_imp	-8,162***	LR2_1tc	-4,847***	MR_order	-3,55**
Jap_imp	-7,859***	MR_1tc	-5,245***	MR_order_fleet	-3,467**
Eur_imp	-8,316***	MR_3tc	-3,861**	VLCC_price	-6,991***
Ge_imp	-7,976***	US_gaso	-6,837***	VLCC_SP	-4,864***
Bel_fimp	-7,775***	Bunker_Jap	-5,563***	VLCC_demo_price	-4,914***
US_fimp	-8,082***	Bunker_Phil	-5,638***	Afra_price	-3,511**
Padd3_fimp	-8,039***	Brent_forw	-5,292***	Afra_SP	-4,229***
Jap_cimp	-6,439***	Brent	-5,831***	Afra_demo_price	-4,877***
Ne_cimp	-6,848***	WTI	-5,875***	Pana_price	-8,607***
US_cimp	-8,142***	Dubai	-6,354***	Pana_SP	-4,175***
Major_imp	-6,692***	Oil_price_index	-6,297***	Pana_demo_price	-4,9***
Chi_imp	-8,113***	ClarkSea	-5,803***	MR_demo_price	-4,865***
Ind_imp	-8,172***	ClarkAve	-7,028***	MR_price	-3,549**
BelUS_fimp	-4,052***	BDTI	-7,318***	MR_SP	-4,461***
BelPadd3_fimp	-8,833***	BCTI	-6,829***	VLCC_speed	-5,988***
NeUS_cimp	-8,595***	Tadawul	-4,319***	Crude_speed	-5,827***
NePadd3_cimp	-3,686**	Nikkei	-5,698***	Afra_speed	-6,45***

Significance *** 0.01 ** 0.05 * 0.1

Table D.2: An overview of the different lags applied to the time series for the specific routes; TD1, TD3, TD7, TD12, TC1, and TC2. All variables are lagged between 1 and 6. Note, not all variables are relevant to each route - see Table A.1 for which variables that are applicable to each route.

Variable	TD1	TD3	TD7	TD12	TC1	TC2
US_dem	4	5	5	5	6	6
Jap_dem	1	1	1	1	6	6
Eur_dem	3	3	3	3	5	5
US_fdem	1	1	1	1	1	1
Eur_fdem	1	1	1	1	1	1
Jap_cdem	3	3	3	3	3	3
Eur_cdem	1	1	1	1	1	1
US_cdem	6	6	6	6	6	6
AG_exp	1	1	1	1	1	1
NS_exp	6	6	6	6	6	6
Bel_fexp	4	4	4	4	4	4
US_fexp	1	1	1	1	1	1
USBel_fexp	1	1	1	1	1	1
SA_cexp	1	1	1	1	1	1
Ne_cexp	1	1	1	1	1	1
US_cexp	1	1	1	1	1	1
USNe_exp	1	1	1	1	1	1
SA_exp	1	1	1	1	1	1
VLCC_fix_west	1	1	1	1	1	1
VLCC_fix	1	1	1	1	1	1
VLCC_due	1	1	1	1	1	1
VLCC_fix_east	5	5	5	5	5	5
VLCC_fix_jap	2	2	2	2	2	2
Afra_fix_sum	1	1	1	1	1	1
Afra_fix	1	1	1	1	1	1
Pana_fix_US	1	1	1	1	1	1
Pana_fix	1	1	1	1	1	1
Afra_fix_east	1	1	1	1	1	1
Afra_fix_us	4	4	4	4	4	4
MR_fix_US	1	1	1	1	1	1
MR_fix	1	1	1	1	1	1
US_sea_imp	1	1	1	1	1	1
US_imp	4	4	4	4	4	4
SAUS_imp	1	1	1	1	1	1
SAPadd3_imp	4	4	4	4	4	4
PADD3_imp	1	1	1	1	1	1
Jap_imp	1	1	1	1	1	1
Eur_imp	1	1	1	1	1	1
Ge_imp	1	1	1	1	1	1
Bel_fimp	1	1	1	1	1	1
US_fimp	1	1	1	1	1	1
Padd3_fimp	1	1	1	1	1	1
Jap_cimp	1	1	1	1	1	1
Ne_cimp	1	1	1	1	1	1
US_cimp	1	1	1	1	1	1
Major_imp	1	1	1	1	1	1
Chi_imp	1	6	1	2	1	1
Ind_imp	1	3	1	2	1	1
BelUS_fimp	6	6	6	6	6	6
BelPadd3_fimp	6	6	6	6	6	6
NeUS_cimp	1	1	1	1	1	1
NePadd3_cimp	6	6	6	6	6	6
ME_prod	1	1	1	1	1	1
NA_prod	1	1	1	1	1	1
US_prod	1	1	1	1	1	1

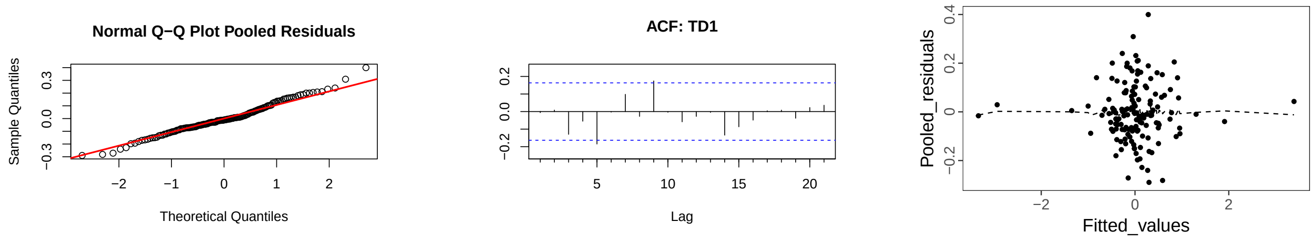
Variable	TD1	TD3	TD7	TD12	TC1	TC2
SA_prod	1	1	1	1	6	6
NS_prod	1	1	1	1	1	1
OPEC_prod	1	1	1	6	6	6
W_prod	1	6	6	6	6	6
Padd3_refuti	6	6	6	6	6	6
Padd1_refuti	6	6	6	6	6	6
US_fout	1	1	1	1	1	1
Bel_fout	6	6	6	6	6	6
SA_cout	1	1	1	1	1	1
Ne_cout	2	2	2	2	2	2
US_cout	2	2	2	2	2	2
USD_SAR	1	1	1	1	1	1
Yen_USD	6	6	6	6	1	1
USD_Pound	1	1	1	1	1	1
USD_Eur	1	1	1	1	1	1
SDR_USD	1	1	1	1	1	1
Euro_index	1	1	1	1	1	1
USD_index	1	1	6	6	6	6
GDP_w	1	6	6	6	6	6
US_CPI	1	1	1	6	6	6
Jap_CPI	1	1	1	1	1	1
Eur_CPI	6	6	6	6	6	6
Ind_US	3	1	5	5	6	6
Ind_Jap	2	2	2	2	5	5
Ind_Eur	2	2	2	2	2	2
Ind_OECD	4	4	4	4	6	6
Ind_China	6	6	6	6	6	6
Ind_India	4	2	5	6	3	6
LIBOR	1	1	1	1	1	1
LIBOR_Yen	1	1	1	1	1	1
LIBOR_Eur	1	1	1	1	1	1
Jap_money	6	6	6	6	1	1
US_money	1	1	1	1	1	1
VLCC_1tc	1	1	1	1	1	1
Afra_1tc	1	1	1	1	1	1
Pana_1tc	1	1	1	1	1	1
LR2_1tc	1	1	1	1	1	1
MR_1tc	1	1	1	1	1	1
MR_3tc	1	1	1	1	1	1
US_gaso	1	6	6	6	6	6
Bunker_Jap	1	6	6	6	6	6
Bunker_Phil	1	6	6	6	6	6
Brent_forw	1	1	1	1	1	1
Brent	1	6	6	6	6	6
WTI	1	6	6	6	6	6
Dubai	1	6	6	6	6	6
Oil_price_index	1	6	6	6	6	6
ClarkSea	1	1	1	1	1	1
ClarkAve	1	1	1	1	1	1
BDTI	1	1	1	1	1	1
BCTI	1	1	1	1	1	1
Tadawul	1	1	1	1	1	1
Nikkei	6	6	6	6	1	1
MSCI_w	6	1	1	1	1	1

Variable	TD1	TD3	TD7	TD12	TC1	TC2
MSCI_e	1	1	1	1	1	1
VIX	1	1	6	6	1	6
SP500	1	1	1	1	1	1
VLCC_age	1	1	1	1	1	1
Afra_age	6	6	6	6	6	6
Pana_age	6	6	6	6	6	6
MR_age	6	6	6	6	6	6
VLCC_down	1	1	1	1	1	1
Afra_down	1	1	1	1	1	1
Pana_down	1	1	1	1	1	1
LR2_down	1	1	1	1	1	1
MR_down	1	1	1	1	1	1
VLCC_deliveries	4	4	4	4	4	4
Afra_deliveries	1	1	1	1	1	1
Pana_deliveries	1	1	1	1	1	1
LR2_deliveries	5	5	5	5	5	5
MR_deliveries	2	2	2	2	2	2
VLCC_mdwt	6	6	6	6	6	6
VLCC_yy	6	6	6	6	6	6
Afra_mdwt	6	6	6	6	6	6
Afra_yy	6	6	6	6	6	6
Pana_mdwt	6	6	6	6	6	6
Pana_yy	1	1	1	1	1	1
LR2_mdwt	1	1	1	1	1	1
LR2_yy	1	1	1	1	1	1
MR_mdwt	1	1	1	1	1	1
MR_yy	1	1	1	1	1	1
VLCC_new	6	1	1	1	1	1
Afra_new	1	1	1	1	1	1
Pana_new	6	6	6	6	6	6
LR2_new	1	1	1	1	1	1
MR_new	4	4	4	4	4	4
VLCC_order	1	1	1	1	1	1
VLCC_order_fleet	1	1	1	1	1	1
Afra_order	1	1	1	1	1	1
Afra_order_fleet	1	1	1	1	1	1
Pana_order	1	1	1	1	1	1
Pana_order_fleet	1	1	1	1	1	1
LR2_order	1	1	1	1	1	1
LR2_order_fleet	1	1	1	1	1	1
MR_order	1	1	1	1	1	1
MR_order_fleet	1	1	1	1	1	1
VLCC_price	1	1	1	1	1	1
VLCC_SP	1	1	1	1	1	1
VLCC_demo_price	1	6	1	1	1	1
Afra_price	1	1	1	1	1	1
Afra_SP	1	1	1	1	1	1
Afra_demo_price	3	3	3	3	1	1
Pana_price	1	1	1	1	1	1
Pana_SP	1	1	1	1	1	1
Pana_demo_price	1	1	1	1	1	1
MR_demo_price	1	1	1	1	1	1
MR_price	1	1	1	1	1	1
MR_SP	1	1	1	1	1	1

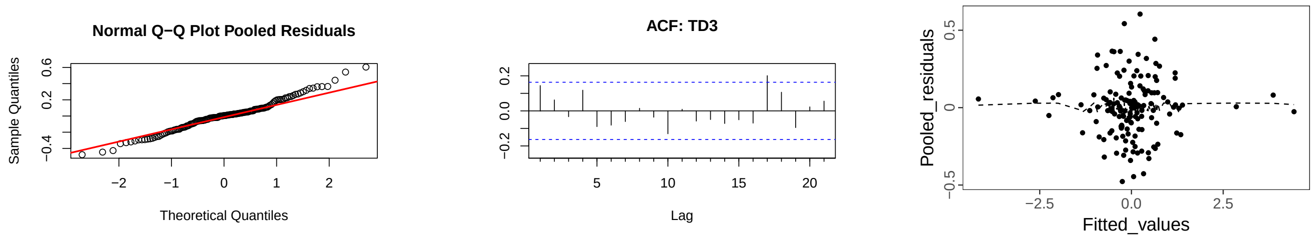
Table D.3: Parsimonious search. The ranked variables which are the basis for creating our parsimonious models. These variables are based on a search from the variables included in the original domain knowledge regime models. They have been ranked by the method of stability selection with randomized lasso. The fraction of included observations was tuned to 0.3 and 0.05 in order to look for variables that perform better in certain short time periods, i.e. we were attempting to narrow down variables for the high volatility regime.

TD1		TD3		TD7	
Regularization run: 0.3		Regularization run: 0.3		Regularization run: 0.3	
	score		score		score
ClarkSea	0.688	VLCC_SP	0.323	Chi_imp	0.559
SA_exp	0.653	Ind_imp	0.240	Ind_India	0.490
Chi_imp	0.611	Jap_money	0.210	LIBOR	0.437
VLCC_due	0.595	Ind_Jap	0.157	ClarkSea	0.366
VIX	0.509	VIX	0.150	Afra_demo_price	0.365
VLCC_price	0.431			Ind_imp	0.364
W_prod	0.425			Ind_US	0.355
SAPadd3_imp	0.411			WTI	0.346
Ind_China	0.411			Eur_dem	0.335
Ind_India	0.396			Ind_Eur	0.326
USD_SAR	0.391			NS_exp	0.312
VLCC_fix_west	0.376			USD_Pound	0.310
US_money	0.354			Eur_imp	0.288
Regularization run: 0.05		Regularization run: 0.05		Regularization run: 0.05	
	score		score		score
VLCC_fix_west	0.310	Ind_imp	0.210	Ind_imp	0.315
Chi_imp	0.276	VIX	0.205	NS_exp	0.255
SAPadd3_imp	0.273	VLCC_SP	0.170	ClarkSea	0.255
Ind_India	0.272	Jap_imp	0.165	Chi_imp	0.250
VLCC_due	0.267	ClarkSea	0.160	Ind_Eur	0.245
Major_imp	0.266			VIX	0.230
VIX	0.260			Afra_SP	0.230
ClarkSea	0.258			NS_prod	0.215
Ind_imp	0.257			Ind_US	0.215
Brent_forw	0.252			Afra_mdwt	0.210
US_money	0.243			WTI	0.205
SA_exp	0.238			USD_Pound	0.200
WTI	0.236			Ind_India	0.200

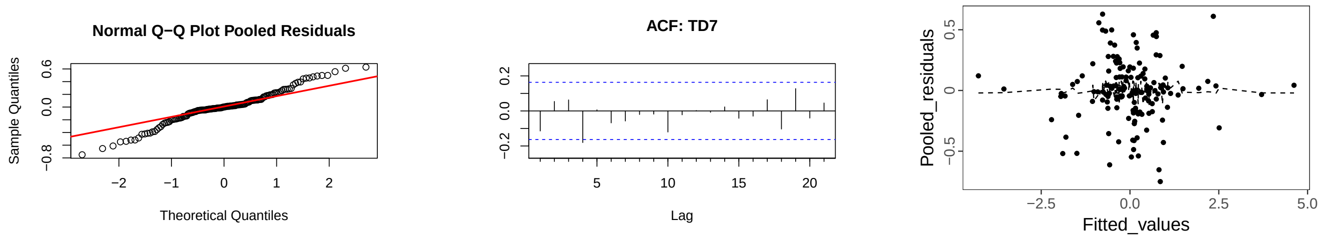
TD12		TC1		TC2	
Regularization run: 0.3		Regularization run: 0.3		Regularization run: 0.3	
	score		score		score
ClarkSea	0.745	Afra_fix_east	0.687	USNe_exp	0.330
Pana_price	0.385	W_prod	0.643	MR_order	0.255
Ind_US	0.342	SDR_USD	0.636	Ne_cimp	0.244
Chi_imp	0.312	LR2_order	0.597	Ne_cout	0.236
WTI	0.298	ClarkSea	0.554	MR_3tc	0.213
US_fdem	0.281	Ind_US	0.468	Brent_forw	0.186
US_dem	0.272	Yen_USD	0.448	US_CPI	0.172
Bel_fimp	0.259	Jap_cimp	0.434	Eur_cdem	0.168
Brent_forw	0.259	SA_cout	0.430		
LIBOR	0.243	Tadawul	0.406		
Ind_India	0.239	Afra_fix_us	0.400		
SDR_USD	0.223	Ind_Jap	0.381		
Major_imp	0.220	Chi_imp	0.377		
Regularization run: 0.05		Regularization run: 0.05		Regularization run: 0.05	
	score		score		score
ClarkSea	0.270	Ind_Jap	0.275	Brent_forw	0.200
Chi_imp	0.230	Brent_forw	0.255	USNe_exp	0.190
Bel_fexp	0.220	W_prod	0.245	Chi_imp	0.185
Brent_forw	0.210	SA_cout	0.240	OPEC_prod	0.180
BelPadd3_fimp	0.210	LR2_order	0.235	MR_fix_US	0.165
US_dem	0.200	Jap_cimp	0.225	MR_3tc	0.165
US_fdem	0.190	Ind_US	0.225	VIX	0.160
Padd3_fimp	0.190	Major_imp	0.220	Ne_cout	0.160
Ind_India	0.190	Ind_India	0.220	Major_imp	0.160
Ind_China	0.185	ClarkSea	0.220	Ind_OECD	0.155
Bel_fout	0.180	WTI	0.215	Ind_China	0.155
US_fout	0.175	Afra_price	0.215		
US_money	0.165	Tadawul	0.210		



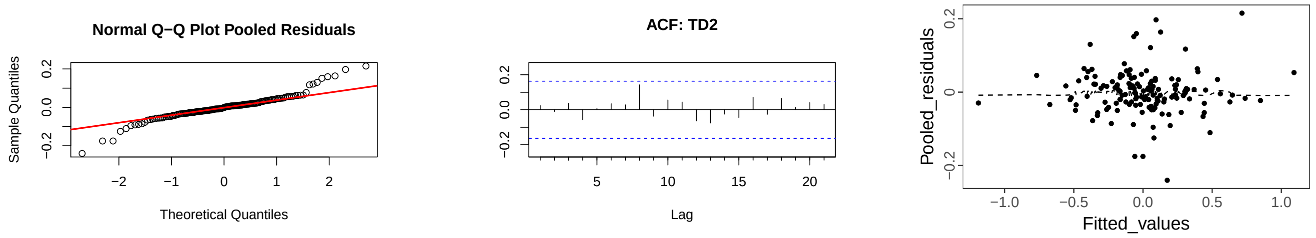
(a) Q-Q plot TD1 (b) ACF plot TD1 (c) Residual plot TD1



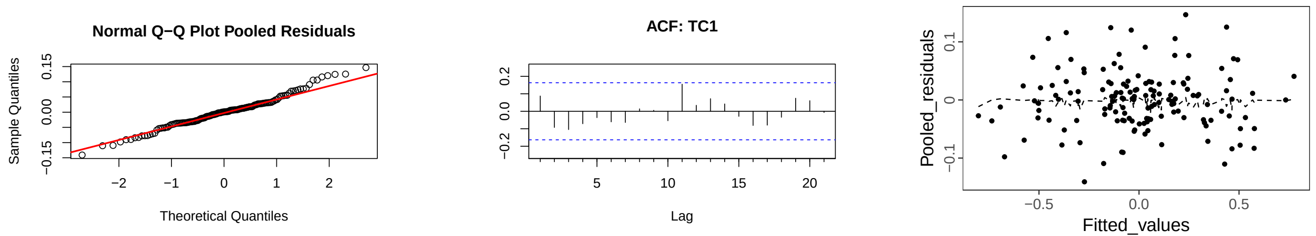
(d) Q-Q plot TD3 (e) ACF plot TD3 (f) Residual plot TD3



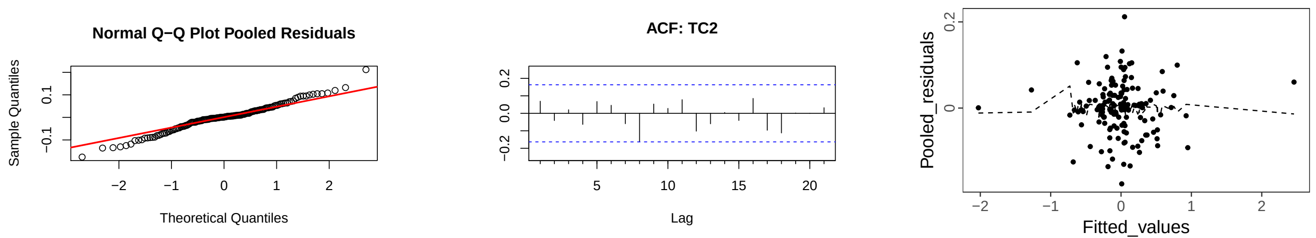
(g) Q-Q plot TD7 (h) ACF plot TD7 (i) Residual plot TD7



(j) Q-Q plot TD12 (k) ACF plot TD12 (l) Residual plot TD12

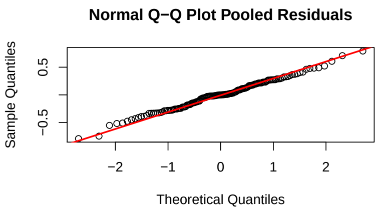


(m) Q-Q plot TC1 (n) ACF plot TC1 (o) Residual plot TC1

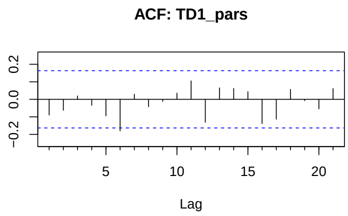


(p) Q-Q plot TC2 (q) ACF plot TC2 (r) Residual plot TC2

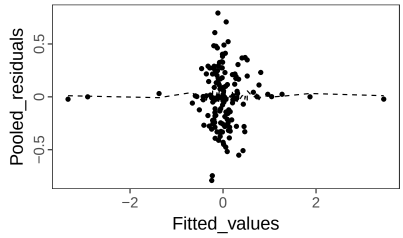
Figure D.1: Residual plots to validate the assumptions for the **domain knowledge regime models**.



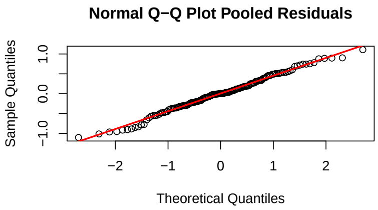
(a) Q-Q plot $TD1_{\text{pars}}$



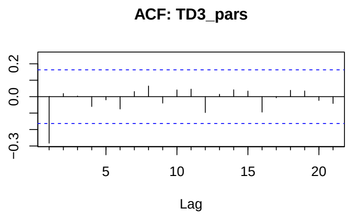
(b) ACF plot $TD1_{\text{pars}}$



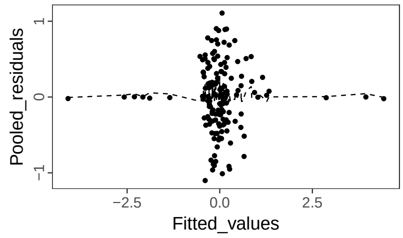
(c) Residual plot $TD1_{\text{pars}}$



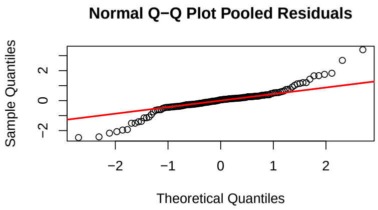
(d) Q-Q plot $TD3_{\text{pars}}$



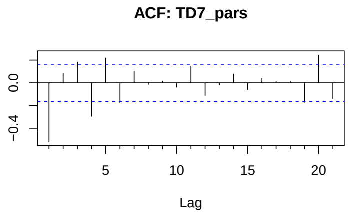
(e) ACF plot $TD3_{\text{pars}}$



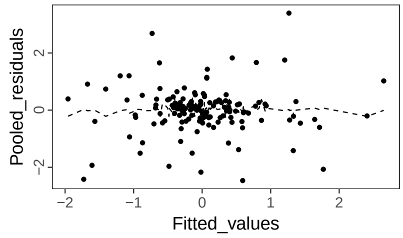
(f) Residual plot $TD3_{\text{pars}}$



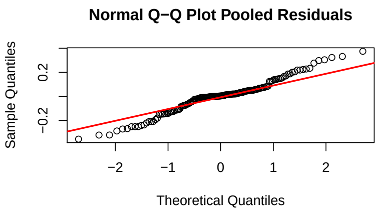
(g) Q-Q plot $TD7_{\text{pars}}$



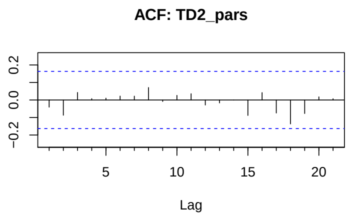
(h) ACF plot $TD7_{\text{pars}}$



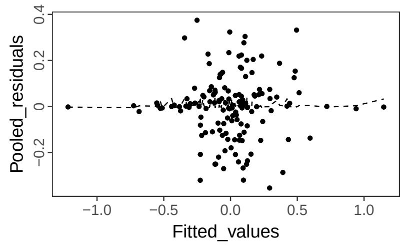
(i) Residual plot $TD7_{\text{pars}}$



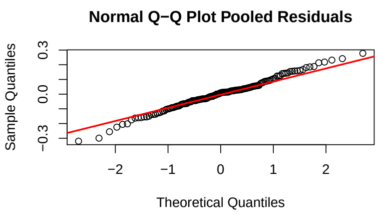
(j) Q-Q plot $TD12_{\text{pars}}$



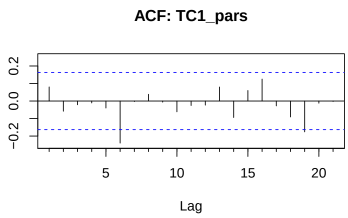
(k) ACF plot $TD12_{\text{pars}}$



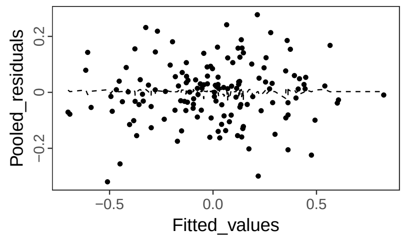
(l) Residual plot $TD12_{\text{pars}}$



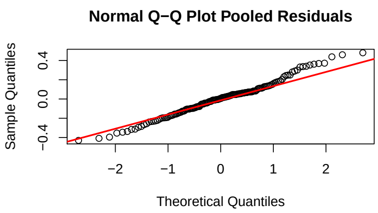
(m) Q-Q plot $TC1_{\text{pars}}$



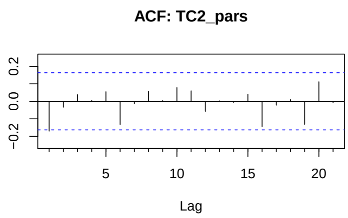
(n) ACF plot $TC1_{\text{pars}}$



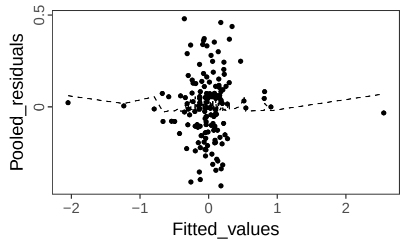
(o) Residual plot $TC1_{\text{pars}}$



(p) Q-Q plot $TC2_{\text{pars}}$



(q) ACF plot $TC2_{\text{pars}}$



(r) Residual plot $TC2_{\text{pars}}$

Figure D.2: Residual plots to validate the assumptions for the **parsimonious regime models**.

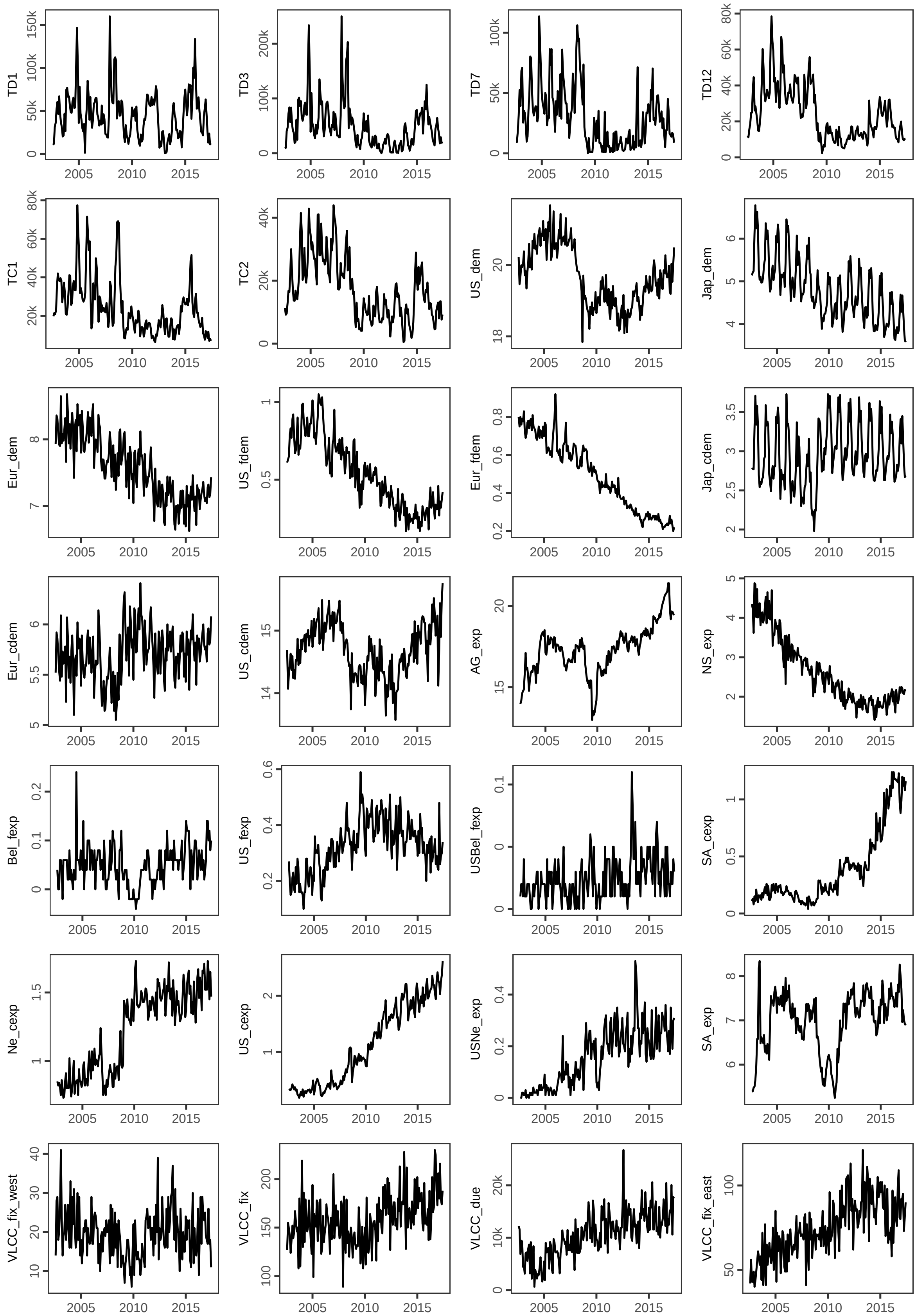


Figure D.3: Displaying all time series considered for this thesis 1/7.

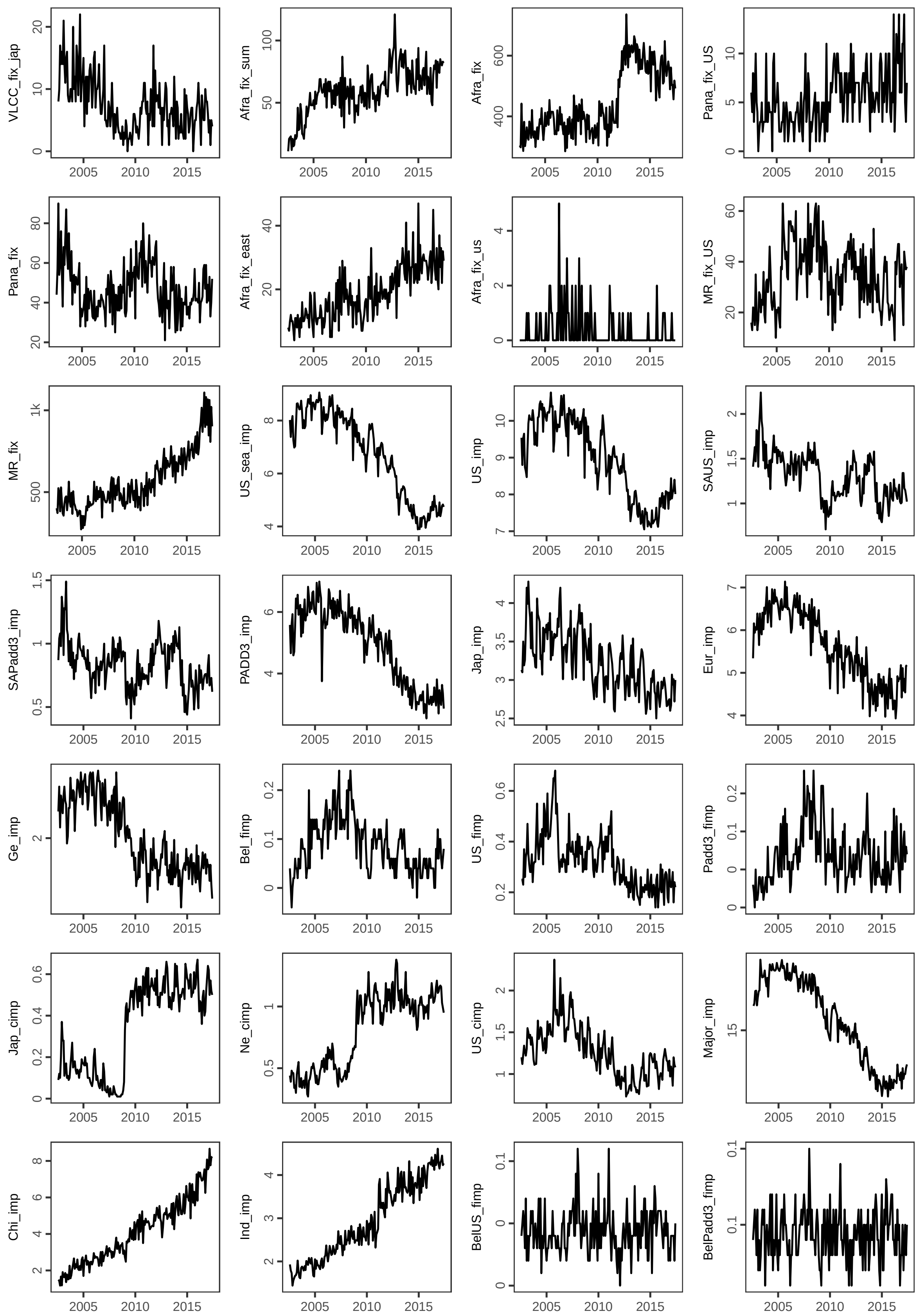


Figure D.4: Displaying all time series considered for this thesis 2/7.

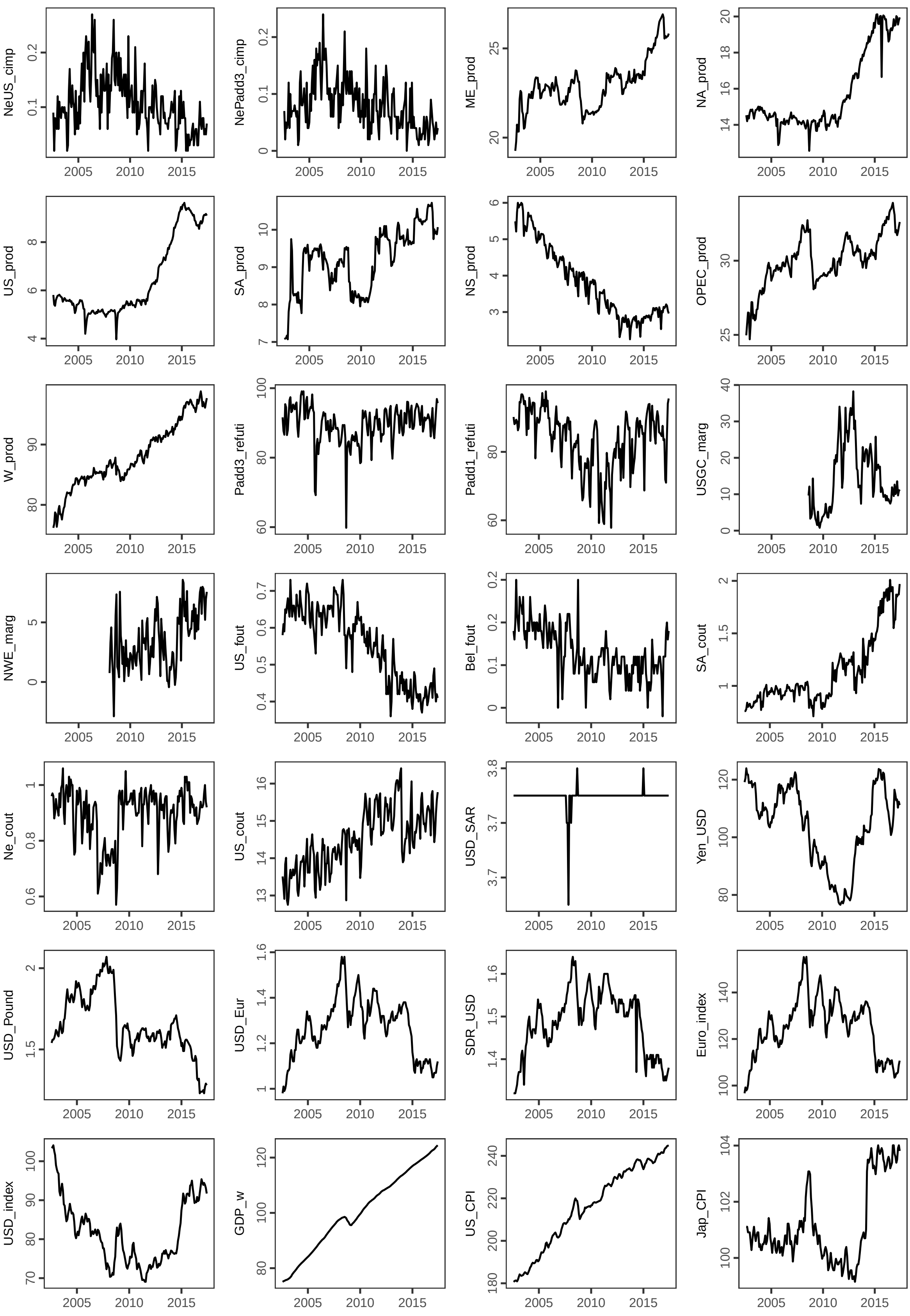


Figure D.5: Displaying all time series considered for this thesis 3/7.

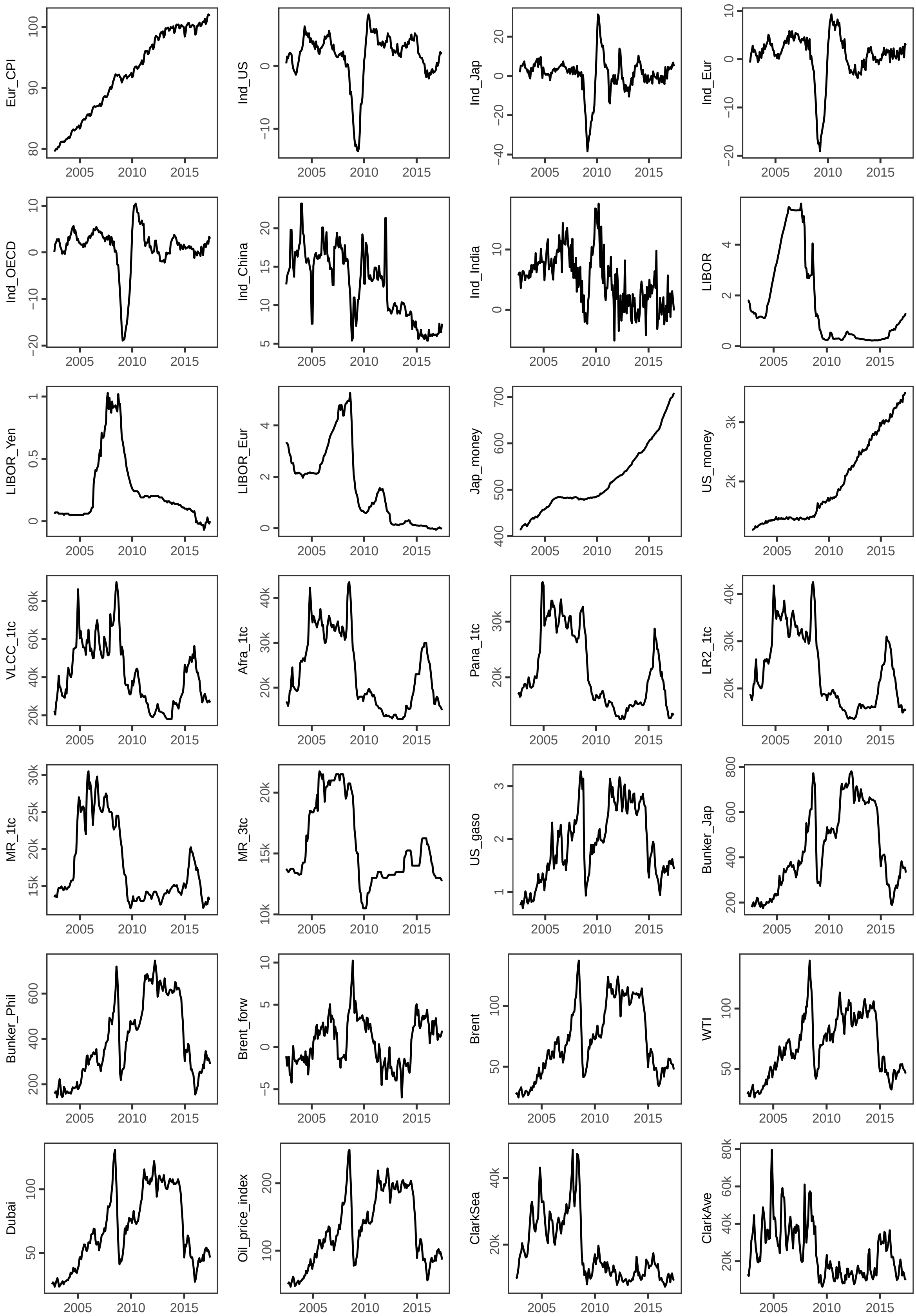


Figure D.6: Displaying all time series considered for this thesis 4/7.

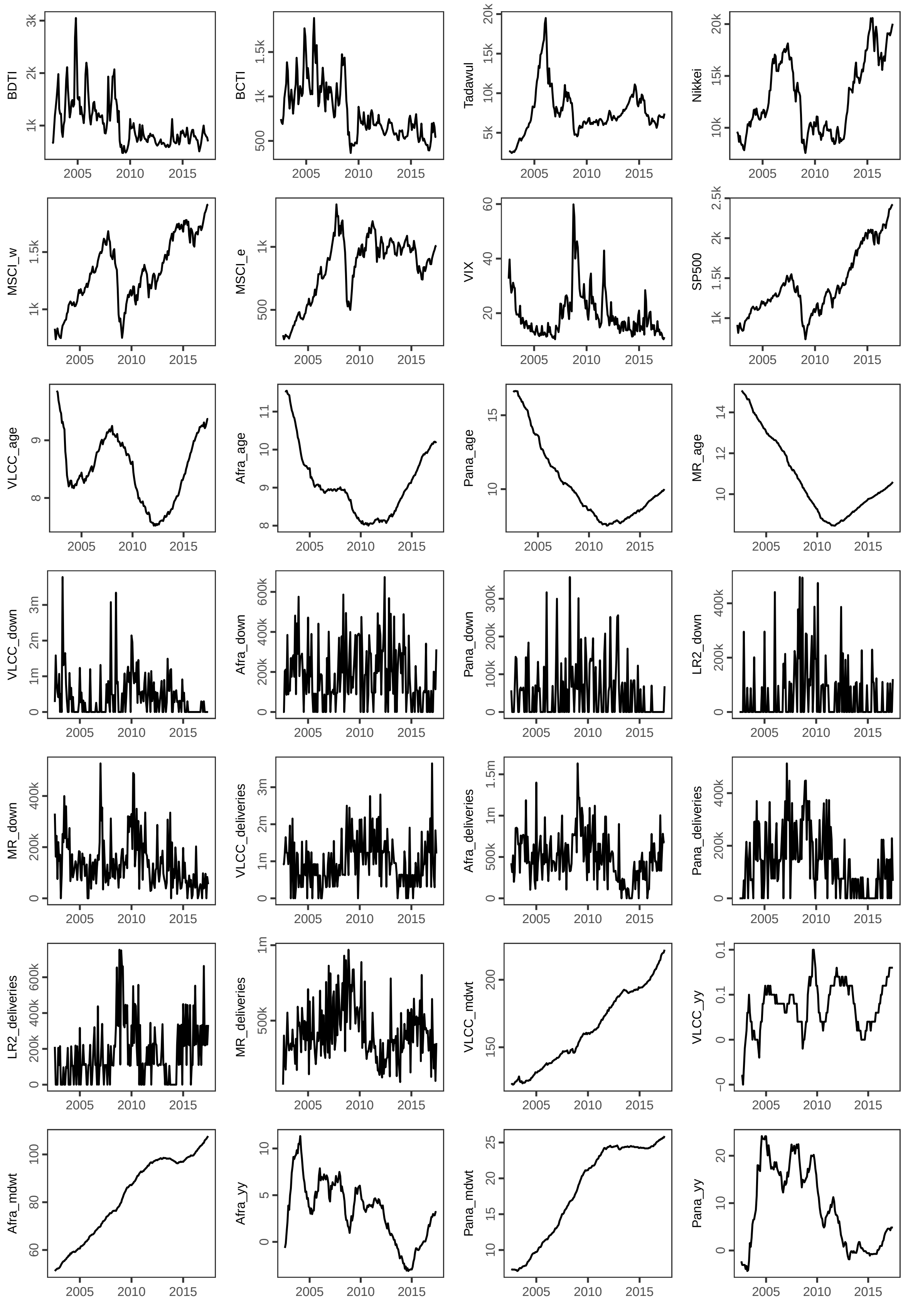


Figure D.7: Displaying all time series considered for this thesis 5/7.

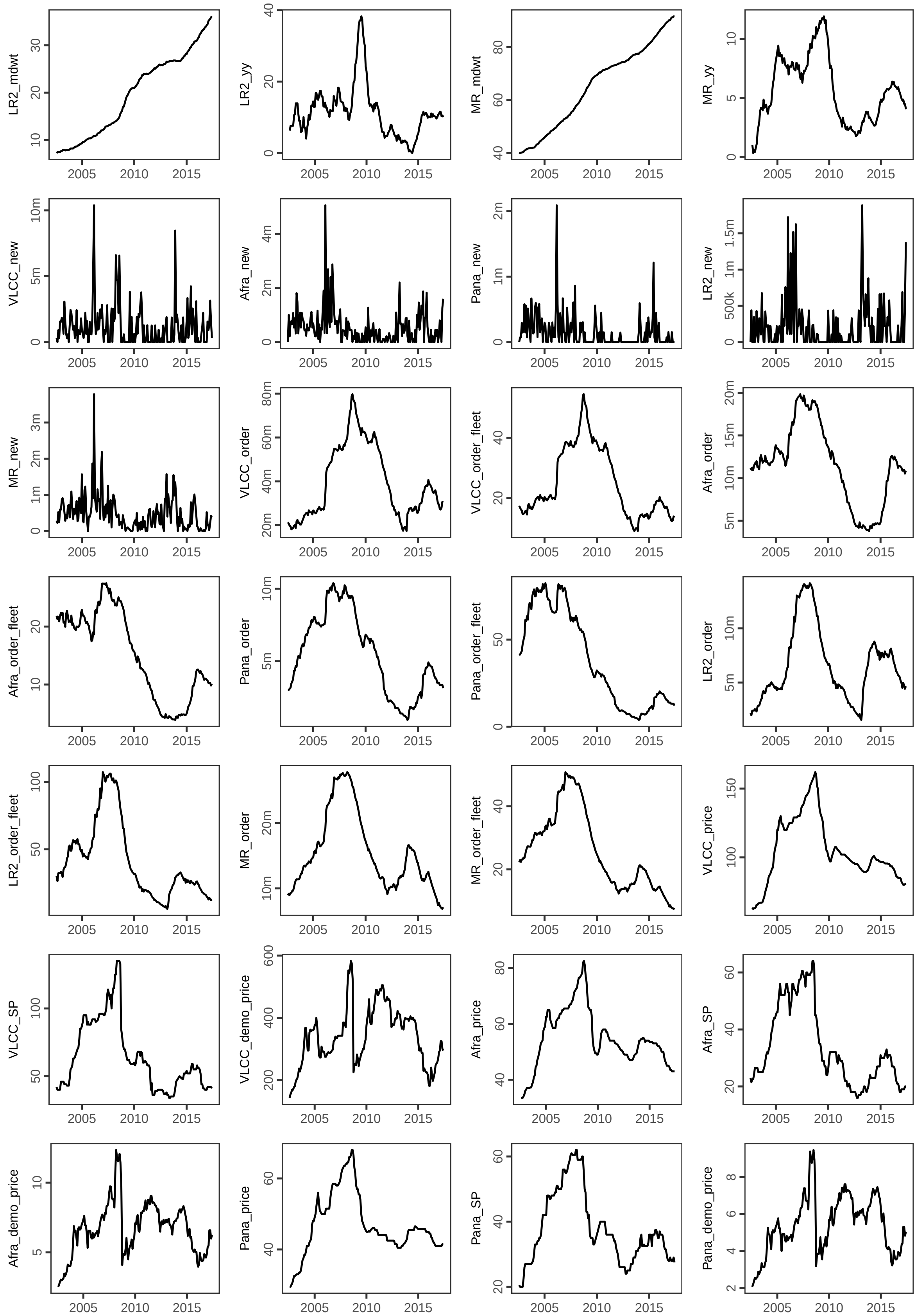


Figure D.8: Displaying all time series considered for this thesis 6/7.

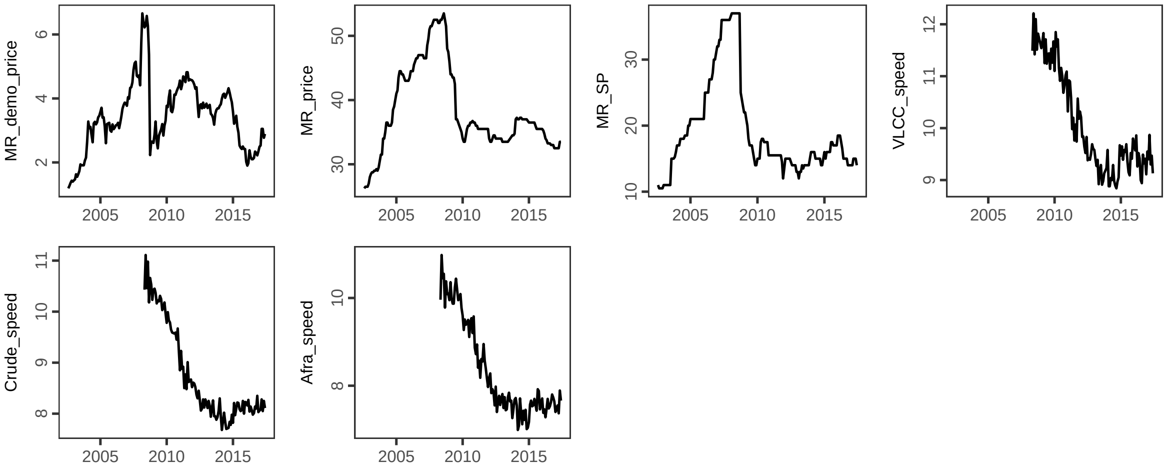


Figure D.9: Displaying all time series considered for this thesis 7/7.

References

Abouarghoub, W. (2013). *Implementing the new science of risk management to tanker freight markets* (PHD, University of the West of England). Retrieved from http://eprints.uwe.ac.uk/20836

Abouarghoub, W. & Mariscal, I. B.-F. (2011). Measuring level of risk exposure in tanker shipping freight markets. *International Journal of Business and Social Research*, *1*(1), 20–44. Retrieved from https://EconPapers.repec.org/RePEc:mir:mirbus:v:1:y:2011:i:1:p:20-44

Abouarghoub, W., Mariscal, I. B.-F., & Howells, P. (2012). Dynamic earnings within tanker markets: An investigation of exogenous and endogenous structure breaks. *American International Journal of Contemporary Research*.

Abouarghoub, W., Mariscal, I. B.-F., & Howells, P. (2014). A two-state markov-switching distinctive conditional variance application for tanker freight returns. *International Journal of Financial Engineering and Risk Management*. Retrieved from http://www.inderscienceonline.com/doi/abs/10.1504/IJFERM.2014.058762

Abouarghoub, W., Nomikos, & Petropoulos. (2017). On reconciling macro and micro energy transport forecasts for strategic decision making in the tanker industry. *Transportation Research Part E: Logistics and Transportation Review*. doi:https://doi.org/10.1016/j.tre.2017.10.012

Adland, R. (2003). *The stochastic behavior of spot freight rates and the risk premium in bulk shipping* (PHD, Ocean Systems Management, MIT).

Adland, R. & Cullinane, K. (2005). A time-varying risk premium in the term structure of bulk shipping freight rates. *Journal of Transport Economics and Policy (JTEP)*, *39*(2), 191–208. Retrieved from https://www.ingentaconnect.com/content/lse/jtep/2005/00000039/00000002/art00004

Adland, R. & Cullinane, K. (2006). The non-linear dynamics of spot freight rates in tanker markets. *Transportation Research Part E: Logistics and Transportation Review*, *42*(3), 211–224.

Adland, R. & Koekebakker, S. (2007). Ship valuation using cross-sectional sales data: A multivariate non-parametric approach. *Maritime Economics & Logistics*, *9*(2), 105–118. doi:10.1057/palgrave.mel.9100174

Alderton, P. M. & Rowlinson, M. (2013). *The economic of shipping freight markets*. Informa Law from Routledge.

Alexander, C. (2008a). *Market Risk Analysis, Practical Financial Econometrics (Volume II)*. John Wiley & Sons, Ltd.

Alexander, C. (2008b). *Market Risk Analysis, Quantitative Methods in Finance (Volume I)*. John Wiley & Sons, Ltd.

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

Alizadeh, A. & Nomikos, N. (2011). Dynamics of the term structure and volatility of shipping freight rates. *Journal of Transport Economics and Policy (JTEP)*. Retrieved from http://www.ingentaconnect.com/content/lse/jtep/2011/00000045/00000001/art00006

Alizadeh, A. & Talley, W. K. [Wayne K.]. (2011). Vessel and voyage determinants of tanker freight rates and contract times. *Transport Policy*, *18*(5), 665–675. doi:10.1016/j.tranpol.2011.01.001

Anyanwu, J. O. (2013). Analysis of the interrelationships between the various shipping markets. *Analysis*, *3*(10).

Baumöhl, E. & Lyócsa, S. (2009). Stationarity of time series and the problem of spurious regression. *Munich Personal RePEc Archive*, (27926). Retrieved from http://mpra.ub.uni-muenchen.de/27926/

Beard, D. E. (2011). *Tanker and Bulk Shipping Industry*. Iberia Capital Partners. Retrieved from http://www.tradewindsnews.com/incoming/article268101.ece5/binary/Shipping%20overview

Beenstock, M. & Vergottis, A. (1989). An Econometric Model of the World Tanker Market. *Journal of Transport Economics and Policy*, *23*(3), 263–280.

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

Benth, F. E. & Koekebakker, S. (2016). Stochastic modeling of supramax spot and forward freight rates. *Maritime Economics Logistics*, *18*(4), 391–413.

Benth, F. E., Koekebakker, S., & Taib, C. M. I. C. (2015). Stochastic dynamical modelling of spot freight rates. *IMA Journal of Management Mathematics*, *26*(3), 273–297. Retrieved from http://dx.doi.org/10.1093/imaman/dpu001

BP. (2017). BP Statistical Review of World Energy 2017. Retrieved from https://www.bp.com/content/dam/bp/en/corporate/pdf/energy-economics/statistical-review-2017/bp-statistical-review-of-world-energy-2017-full-report.pdf

Clarksons. (2017a). *Expert commentary.* Retrieved January 5, 2018, from https://www.clarksons.com/

Clarksons. (2017b). Glossary. Retrieved June 5, 2017, from http://www.clarksons.com/glossary/

Clarksons Research Services Limited. (2017). Shipping Intelligence Network and World Fleet Register databases. Retrieved September 12, 2017, from https://clarksons.net/

Cullinane, K., Teng, Y., & Wang, T.-F. (2005). Port competition between shanghai and ningbo. *Maritime Policy & Management*, *32*(4), 331–346.

Dafas, P. A. (2004). *Estimating the parameters of a mean-reverting markov-switching jump-diffusion model for crude oil spot prices* (Master's thesis, Finance Birkbeck College).

Diebold, F. X. (2013). Comparing Predictive Accuracy , Twenty Years Later : A Personal Perspective on the Use and Abuse of Diebold – Mariano Tests. *Journal of Business and Economic Statistics*, (November). doi:10.1080/07350015.2014.983236

Diebold, F. X. (2017). *Forecasting in Economics, Business, Finance and Beyond.* Department of Economics, University of Pennsylvania. Retrieved from http://www.ssc.upenn.edu/%7B~%7Dfdiebold/Textbooks.html

Dikos, G., Marcus, H. S., Papadatos, M. P., & Papakonstantinou, V. (2006). Niver lines: A system-dynamics approach to tanker freight modeling. *Interfaces*, *36*(4), 326–341. doi:10.1287/inte.1060.0218

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78. doi:10.1145/2347736.2347755. arXiv: 9605103 `[cs]`

Doreswamy & Vastrad, C. M. (2013). Performance Analysis of Regularized Linear Regression Models for Oxazolines and Oxazoles Derivatives Descriptor Dataset. *International Journal of Computational Science and Information Technology*, *1*(4), 111–123. Retrieved from https://arxiv.org/pdf/1312.2789.pdf

EIA. (2012). PADD regions enable regional analysis of petroleum product supply and movements. Retrieved from https://www.eia.gov/todayinenergy/detail.php?id=4890

EIA. (2017a). EIA: Japan Oil Market Overview. Retrieved from https://www.hellenicshippingnews.com/eia-japan-oil-market-overview/

EIA. (2017b). The Strait of Malacca, a key oil trade chokepoint, links the Indian and Pacific Oceans. Retrieved from https://www.eia.gov/todayinenergy/detail.php?id=32452

EIA. (2017c). Three important oil trade chokepoints are located around the Arabian Peninsula. Retrieved from https://www.eia.gov/todayinenergy/detail.php?id=32352

EIA. (2017d). What Drives Crude Oil Prices? Retrieved from https://www.eia.gov/finance/markets/crudeoil/demand-oecd.php

EIA. (2018). U.S. Energy Information Administration (EIA) homepage. Retrieved from https://www.eia.gov/

EnQuest. (2017). Acquisition of Magnus Oil Field & Sullom Voe Oil Terminal. Retrieved from http://www.enquest.com/%7B~%7D/media/Files/E/Enquest/2017%20-%20magill/MagillEnQInv%7B%5C_%7Dlowres.pdf

Escribano, Á. (2002). Modeling electricity prices: international evidence. *EFMA 2002 London ...* (June), 2–34. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract%5C_id=299360

ExxonMobil. (2017). 2017 Outlook for Energy. Retrieved from http://cdn.exxonmobil.com/%7B~%7D/media/global/files/outlook-for-energy/2017/2017-outlook-for-energy.pdf

Fan, S., Ji, T., Wilmsmeier, G., & Bergqvist, R. (2013). Forecasting baltic dirty tanker index by applying wavelet neural networks. *3*, 68–87.

Fearnley. (2018). *Expert commentary.* Fearnley Securities. Retrieved February 11, 2018, from http://www.fearnleysecurities.no/

Find a Port. (2017). Chiba. Retrieved November 5, 2017, from http://www.findaport.com/port-of-chiba

FRED. (2017). Federal Reserve Economic Data. Retrieved September 17, 2017, from https://fred.stlouisfed.org/

Grotenhuis, M. & Thijs, P. (2015). Dummy variables and their interactions in regression analysis: examples from research on body mass index, 1–22. arXiv: 1511.05728. Retrieved from http://arxiv.org/abs/1511.05728

Gujarati, D. N. (2011). *Econometrics.* McGraw-Hill.

Halls-Moore, M. (2016). Maximum Likelihood Estimation for Linear Regression. Retrieved May 1, 2017, from https://www.quantstart.com/articles/Maximum-Likelihood-Estimation-for-Linear-Regression

Hampton, M. J. (1991). *Long and short shipping cycles.* Cambridge: Cambridge Academy of Transport.

Harrell, F. E. (2001). *Regression Modeling Strategies.* doi:10.1007/978-1-4757-3462-1. arXiv: arXiv:1011.1669v3

Hasinur, M., Khan, R., Bhadra, A., & Howlader, T. (2016). Stability Selection for Lasso, Ridge and Elastic Net Implemented with AFT Models Running Title: Stability Selection for AFT Models. *Cornell University Library ArXiv e-prints*, 1–24. arXiv: arXiv:1604.07311v1. Retrieved from https://arxiv.org/pdf/1604.07311.pdf

Hawdon, D. (1978). Tanker freight rates in the short and long run. *Applied Economics*, *10*(3), 203–218. doi:10.1080/758527274

Hudson, R. S. & Gregoriou, A. (2014). Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns. *International Review of Financial Analysis*, *38*(0), 151–162. doi:10.1016/j.irfa.2014.10.008

Hyndman, R. J. (2006). Another Look At Forecast-Accuracy Metrics for Intermittent Demand. *Foresight: The International Journal of Applied Forecasting*, (4), 43–46. doi:10.1016/j.ijforecast.2006.03.001. arXiv: arXiv:1011.1669v3

Hyndman, R. J. (2014). Fitting models to short time series. Retrieved from https://robjhyndman.com/hyndsight/short-time-series/

Hyndman, R. J. & Athanasopoulos, G. (2014). *Forecasting: Principles and Practice.* doi:10.1017/CBO9781107415324.004

IAME. (2014). Freight Dynamics within the Tanker Market: A Conditional multi- factor freight return model with Markov regime switching indicator functions as threshold parameters. *IAME 2014* (Paper ID 215). Retrieved from https://orca.cf.ac.uk/73377/1/freightdynamicsv1.pdf

ICE. (2013). Ice crude & refined oil products. *IntercontinentalExchange (ICE).* Retrieved from https://www.theice.com/publicdocs/ICE_Crude_Refined_Oil_Products.pdf

IMO. (2017). The International Maritime Organization (IMO) homepage. Retrieved from http://www.imo.org

Investopedia. (2018a). Expectations Theory. Retrieved January 5, 2018, from https://www.investopedia.com/terms/e/expectationstheory.asp

Investopedia. (2018b). Perfect Competition. Retrieved January 5, 2018, from https://www.investopedia.com/terms/p/perfectcompetition.asp

Jeffries. (2017). *Tanker presentation.* Marine Money (Jeffries report). Retrieved from https://www.marinemoney.com/sites/all/themes/marinemoney/forums/MMWeek13/presentations/Wednesday/2%7B%5C_%7D30%20PM%20Jeff%20Pribor%20-%20Crude,%20Product%20%7B%5C&%7D%20Speciality%20Tankers.pdf

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. doi:10.1080/03088830802079987

JODI. (2017). Joint Organisations Data Initiative. Retrieved October 10, 2017, from https://www.jodidata.org/oil/

Kagkarakis, N. D., Merikas, A. G., & Merika, A. (2016). Modelling and forecasting the demolition market in shipping. *Maritime Policy & Management*, *43*(8), 1021–1035. doi:10.1080/03088839.2016.1185181

Kavussanos, M. G. (2003). Time Varying Risks Among Segments of the Tanker Freight Markets. *Maritime Economics & Logistics*, *5*(3), 227–250. Retrieved from https://ideas.repec.org/a/pal/marecl/v5y2003i3p227-250.html

Kavussanos, M. G. & Alizadeh, A. (2002). Seasonality patterns in tanker spot freight rate markets. *Economic Modelling*, *19*(5), 747–782. doi:https://doi.org/10.1016/S0264-9993(01)00078-5

Kavussanos, M. G. & Visvikis, I. (2016). *The international handbook of shipping finance: Theory and practice.* Palgrave Macmillan UK. Retrieved from https://books.google.no/books?id=JOx6DQAAQBAJ

Klovland, J. T. (2002). Business cycles, commodity prices and shipping freight rates: Some evidence from the pre-wwi period.

Koekebakker, S., Adland, R., & Sødal, S. (2006). Are spot freight rates stationary? *Journal of Transport Economics and Policy.* Retrieved from http://www.jstor.org/stable/20053995

Köhn, S. (2008). *Generalized Additive Models in the Context of Shipping Economics* (Ph.D. University of Leicester). Retrieved from https://lra.le.ac.uk/bitstream/2381/4172/1/2008koehnsphd.pdf

Koopmans, T. C. (1939). Tanker freight rates and tankship building : an analysis of cyclical fluctuations. *Nederlandsch Economisch Instituut. Publication ; Nr. 27.*

Kosater, P. & Mosler, K. (2006). Can markov regime-switching models improve power-price forecasts? evidence from german daily power prices. *Applied Energy.* Retrieved from http://EconPapers.repec.org/RePEc:eee:appene:v:83:y:2006:i:9:p:943-958

Lyridis, D., Manos, N., Zacharioudakis, P., Pappas, A., & Mavris, A. (2017). Measuring tanker market future risk with the use of foresim. *SPOUDAI - Journal of Economics and Business*, *67*(1). Retrieved from http://spoudai.unipi.gr/index.php/spoudai/article/view/2571

Lyridis, D., Zacharioudakis, P., & Ousantzopoulos, G. (2004). A multi-regression approach to forecasting freight rates in the dry bulk shipping market using neural networks. *Annual Conference of the International Association of Maritime Economists, IAME 2004.*

McKinsey. (2015). Refining capacity outlook to 2020: 2015 developments. *Energy insights McKinsey.* Retrieved from https://www.mckinseyenergyinsights.com/insights/

Meinshausen, N. & Bühlmann, P. (2009). Stability selection. *Journal of the Royal Statistical Society*, 1–30. doi:10.1111/j.1467-9868.2010.00740.x. arXiv: arXiv:0809.2932v2

Nau, R. (2017). Regression diagnostics. Retrieved January 9, 2018, from http://people.duke.edu/%7B~%7Drnau/testing.htm

Niedersachsen. (2017). Wilhelmshaven. Retrieved from https://www.nports.de/en/ports/wilhelmshaven/

Norman, V. & Wergeland, T. (1981). Stortankmarkedet frem til 1985: En analyse av markedet for tankskip over 200.000 dwt. Skipsfartsekonomisk markedsforskning. doi:Rapportnr.5/1981

Pagonis, T. (2009). *The chartering practice handbook*. Thomas Pagonis. Retrieved from https://books.google.no/books?id=-RSip8ylfMQC

Pelagatti, M. (2013). *Stationary processes*. University of Milano-Bicocca. Milano. Retrieved from http://www.statistica.unimib.it/utenti/p%7B%5C_%7Dmatteo/lessons/SSE/stationarity.pdf

Port Houston. (2017). Port Houston homepage. Retrieved December 2, 2017, from http://porthouston.com/

Port of Antwerp. (2017). Port of Antwerp homepage. Retrieved December 2, 2017, from http://www.portofantwerp.com/en

Port of NY/NJ. (2017). Port of New York and New Jersey. Retrieved December 2, 2017, from http://www.panynj.gov/port/

Port of Rotterdam. (2017). Port of Rotterdam homepage. Retrieved December 2, 2017, from https://www.portofrotterdam.com/en

Poulakidas, A. & Joutz, F. (2009). Exploring the link between oil prices and tanker rates. *Maritime Policy Management*, *36*(3), 215–233. Retrieved from https://EconPapers.repec.org/RePEc:taf:marpmg:v:36:y:2009:i:3:p:215-233

Pruyn, J. F. J., van de Voorde, E., & Meersman, H. (2011). Second hand vessel value estimation in maritime economics: A review of the past 20 years and the proposal of an elementary method. *Maritime Economics & Logistics*, *13*(2), 213–236. doi:10.1057/mel.2011.6

Randers, J. & Göluke, U. (2007). Forecasting turning points in shipping freight rates: Lessons from 30 years of practical effort. *System Dynamics Review*, *23*(2-3), 253–284. doi:10.1002/sdr.376

Reuters. (2017). U.S. gasoline surplus eliminated by trade flows: Kemp. Retrieved from https://www.reuters.com/article/us-usa-gasoline-kemp/u-s-gasoline-surplus-eliminated-by-trade-flows-kemp-idUSKBN1A52CJ

Reve, T. & Sasson, A. (2012). *A Knowledge-based Norway*. Universitetsforlaget.

Ringheim, J. K. & Stenslet, L. O. D. (2017). *Predicting monthly bulk shipping freight rates* (Master's thesis, Norwegian University of Science and Technology).

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

Strandenes, S. P. (1984). *Price determination in the time charter and second hand markets*. Norwegian School of Economics and Business Administration.

Strandenes, S. P. (1999). Is there potential for a two-tier tanker market? *Maritime Policy & Management*, *26*(3), 249–264.

Strandenes, S. P. (2002). Economics of the markets for ships. *The Handbook of Maritime Economics and Business*, 186–202.

Talley, W. K. [Wayne K]. (2011). *The blackwell companion to maritime economics*. John Wiley & Sons.

Tamvakis, M. (1995). An investigation into the existence of a two-tier spot freight market for crude oil carriers. *22*, 81–90.

Tamvakis, M. & Thanopoulou, H. (2000). Does quality pay? the case of the dry bulk market. *36*, 297–307.

Teekay. (2017). Tanker insights - Strongest Atlantic storm season in years. Retrieved October 18, 2017, from http://teekay.com/blog/2017/09/12/tanker-insights-strongest-atlantic-storm-season-years/

The Baltic Exchange. (2018). The Baltic Exchange homepage. Retrieved February 2, 2018, from https://www.balticexchange.com/

The Maritime Executive. (2017). EU Leads the Way for Ship Recycling. Retrieved from https://maritime-executive.com/editorials/eu-leads-the-way-for-ship-recycling%7B%5C#%7Dgs.I4SvLbo

Tinbergen, J. (1934). Annual survey of significant developments in general economic theory. *Econometrica: Journal of the Econometric Society*, 13–36.

Tønne, E. (2016). Clarksons Platou. Retrieved September 16, 2016, from http://sysla.no/maritim/rederienes-skip-er-mindre-verdt-enn-anslatt/

Tradewinds. (2017a). Euronav wins sustainable storage work for ULCC. Retrieved from http://www.tradewindsnews.com/tankers/1358148/euronav-wins-sustainable-storage-work-for-ulcc

Tradewinds. (2017b). Key products tanker routes. Retrieved December 2, 2017, from http://www.tradewindsnews.com/incoming/article363515.ece5

Tsioumas, V., Papadimitriou, S., Smirlis, Y., & Zahran, S. Z. (2017). A novel approach to forecasting the bulk freight market. *The Asian Journal of Shipping and Logistics*, *33*(1), 33–41. doi:https://doi.org/10.1016/j.ajsl.2017.03.005

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

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

Tvedt, J. (2003). A new perspective on price dynamics of the dry bulk market. *Maritime Policy Management*, *30*(3), 221–230. Retrieved from https://EconPapers.repec.org/RePEc:taf:marpmg:v:30:y:2003:i:3:p:221-230

UNCTAD. (2016). Review of maritime transport 2016. *United Nations Conference on Trade and Development.*

van Dellen, Nomikos, N., & Alizadeh, A. (2011). *An Examination into the Structure of Freight Rates in the Shipping Freight Markets* (Doctor of Philosophy, Cass Business School). Retrieved from http://openaccess.city.ac.uk/7789/

Veenstra, A. W. & Franses, P. H. (1997). A co-integration approach to forecasting freight rates in the dry bulk shipping sector. *Transportation Research Part A: Policy and Practice*, *31*(6), 447–458. doi:https://doi.org/10.1016/S0965-8564(97)00002-5

Velonias, M. (1995). *Forecasting tanker freight rates* (Master's thesis, MIT). Retrieved from https://dspace.mit.edu/bitstream/handle/1721.1/36016/32836540-MIT.pdf?sequence=2

Wärtsilä. (2018). Tankers. Retrieved from https://www.wartsila.com/encyclopedia/term/tankers

Weron, R., Bierbrauer, M., & Trück, S. (2004). Modeling electricity prices: Jump diffusion and regime switching. *Physica A: Statistical Mechanics and its Applications.*

Worldscale Association. (2018). Worldscale homepage. Retrieved March 5, 2018, from https://www.worldscale.co.uk/

Zacharioudakis, P. & Lyridis, D. (2011). Exploring tanker market elasticity with respect to oil production using foresim, 297–315. Retrieved from https://www.worldscientific.com/doi/abs/10.1142/9789814329866_0013

Zannetos, Z. S. (1964). *The theory of oil tankship rates: an economic analysis of tankship operations.* MIT Libraries.

Zhong, M. & Shi, Q. (2007). Proceedings of the ieee international conference on automation and logistics, ical 2007, 709–713.