Cass Business School
CITY UNIVERSITY LONDON

An Examination into the Structure of Freight Rates in the Shipping Freight Markets

Stefan van Dellen

Supervisors:
Professor Nikos K. Nomikos
Dr. Amir H. Alizadeh

A thesis submitted for the degree of Doctor of Philosophy

SEPTEMBER 2011
Cass Business School

This page is intentional left blank.

Table of Contents:

Table of Figures

This page is intentional left blank.

List of Tables

This page is intentional left blank.

List of Abbreviations

$ / Day	US$ per Day
ADF	Augmented Dickey-Fuller Unit Root Test
AFMX	Aframax Vessel
AIC	Akaike Information Criterion
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroscedasticity
ARFIMA	Autoregressive Fractionally Integrated Moving Average Model
ARIMA	Autoregressive Integrated Moving Average Model
ARMA	Autoregressive Moving Average Model
BIFFEX	Baltic International Freight Futures Exchange
CC	Conditional Coverage Test
CDIR	Percentage of Correct Direction Predicted
CF	Cornish-Fisher Expansion
$CI(x;y)$	Cointegrated of Order x and y
CPSZ	Capesize Vessel
D	Level of Demand
d	Order of Fractional Integration
DWT	Dead-Weight Tonnes
EGARCH	Exponential Generalised Autoregressive Conditional Heteroscedasticity Model
ES	Expected Shortfall
ESTAR	Exponentially Smooth Transition Autoregression Model
EWMA	Exponentially Weighted Moving Average
FFA	Forward Freight Agreement
FHS	Filtered Historical Simulation
FHS (200)	Filtered Historical Simulation over a 200 Week Horizon
FHS (400)	Filtered Historical Simulation over a 400 Week Horizon
FIEGARCH	Fractionally Integrated Exponential Generalised Autoregressive Conditional Heteroscedasticity Model
FIGARCH	Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity Model
GARCH	Generalised Autoregressive Conditional Heteroscedasticity Model
GARCHK	Generalised Autoregressive Conditional Heteroscedasticity and Kurtosis Model
GARCHS	Generalised Autoregressive Conditional Heteroscedasticity with Skewness Model
GARCHSK	Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis Model
GARCH-X	Augmented Generalised Autoregressive Conditional Heteroscedasticity Model
GC	Gram-Charlier Expansion
HS	Historical Simulation
HS (200)	Historical Simulation over a 200 Week Horizon
HS (400)	Historical Simulation over a 400 Week Horizon

$I(x)$	Integrated of Order x
IGARCH	Integrated Exponential Generalised Autoregressive Conditional Heteroscedasticity Model
Ind	Independence Test
KPSS	Kwiatkowski, Phillips, Schmidt and Shin Unit Root Test
KSS	Kapetanios, Shin and Snell Unit Root Test
LF	Loss Function
LL	Log-Likelihood
LM	Lagrange-Multiplier Test
LR	Likelihood-Ratio Test
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MDM	Modified Diebold-Mariano Test
MLE	Maximum Likelihood Estimation
NAGARCHSK	Non-Linear Asymmetric Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis Model
OLS	Ordinary Least Squares Estimation
PNMX	Panamax Vessel
QMLE	Quasi-Maximum Likelihood Estimation
RM	RiskMetrics^{TM} Model
RMSE	Root Mean Squared Error
S	Level of Supply
SBIC	Schwartz-Bayesian Information Criterion
SGT	Skewed Generalised *t*-Distribution
SZMX	Suezmax Vessel
TCE	Time-Charter Equivalent Freight Rate
UC	Unconditional Coverage Test
ULCC	Ultra-Large Crude Carrier Vessel
VaR	Value-at-Risk
VAR	Vector Autoregressive Model
VECM	Vector Error Correction Model
VLCC	Very Large Crude Carrier Vessel
WS	Worldscale Freight Rate

Dedication

To my late grandmother Joyce for giving me the love, support and inspiration to complete my PhD.

This page is intentional left blank.

Acknowledgements

Although it is not possible to specifically name everyone who has helped me during the various stages of this thesis and my research studies, I would like to express my greatest thanks and sincere appreciation to my two supervisors Dr. Nikos Nomikos and Dr. Amir Alizadeh for their support, patience and everlasting encouragement throughout these years. Furthermore, I would like to thank my thesis examiners Prof. Roy Batchelor and Dr. Dong-Wook Song for their invaluable feedback and guidance. I am also grateful to Prof. Costas Grammenos for his help throughout my time at Cass and would also like to remember the support of the late Prof. Anthony Lumby. I wish to thank Malla Pratt and Abdul Momin for all their invaluable help, as well as my fellow PhD students at Cass, whom I will not refer to by name for fear of leaving anyone out, for being there through all the good as well as the bad times and providing a listening ear. I would finally like to thank my parents, James and Doria van Dellen, as well as my brothers, Anton, David and Jonathan, and other members of my family for their loving support and encouragement throughout.

Declaration

I grant the powers of discretion to the City University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission covers only single copies made for educational purposes, subject to normal conditions of acknowledgement.

Abstract

This thesis investigates three salient areas of interest in the structure of freight rates in the shipping market, with a particular focus on the tanker and dry-bulk sectors, using recent econometric and time series techniques. The questions asked are: 1) do spot freight rate levels follow a fractionally integrated process, as opposed to being stationary or non-stationary, as had previously been proposed; 2) does spot freight rate volatility also follow a fractionally integrated process; and 3) do freight rates exhibit conditional skewness and kurtosis? It then evaluates the impact that these factors have on the risk exposure of market participants. These concepts are further tested in terms of their respective forecasting performance, relative to other more standard econometric techniques.

An ongoing issue in the shipping literature is whether spot freight rate levels follow a stationary or non-stationary process. This thesis provides another dimension to this discussion by arguing that spot freight rate levels follow a fractionally integrated process. The rationale behind this argument is the fact that the supply and demand dynamics in this market mean that although freight rates are mean-reverting overall, the process of mean-reversion occurs with a delay, which is exactly how one would expect a fractionally integrated process to behave. Although in-sample results were promising in that fractionally integrated models are found to outperform their stationary and non-stationary counterparts across sectors and vessel sizes, out-of-sample forecasts indicate that models that assumed stationarity or non-stationarity outperformed these models, depending on the sector and vessel size.

Additionally, the thesis extends this debate to the volatility of these spot freight rate levels, where it is proposed that volatility also follows a fractionally integrated process. In-sample results from the estimation of Generalised Autoregressive Conditional Heteroscedasticity (GARCH), Integrated Generalised Autoregressive Conditional Heteroscedasticity (IGARCH) and Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) models indicate that FIGARCH models outperformed the other two models across all sectors and vessel sizes, however, when calculating the respective out-of-sample Values-at-Risk for each

vessel type, non-parametric models are found, in most cases, to outperform their parametric counterparts across sectors and vessel sizes.

This thesis finally examines whether freight rates exhibit conditional skewness and kurtosis, where the shape of the supply function in the shipping freight markets indicates that these would not be constant over time, as is assumed by other standard models. Results for the in-sample period indicate that the Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis (GARCHSK) models outperformed GARCH and FIGARCH models. This being said, when calculating the respective out-of-sample Values-at-Risk for each vessel type, non-parametric models are found, in most cases, to outperform their parametric counterparts across sectors and vessel sizes.

1 Review of the Shipping Market and Hypotheses

1.1 General Introduction

For almost as long as people have been trading goods, the primary means of transportation for these goods has been by ship, thus demonstrating the enduring importance of the shipping industry to the world economy as a whole. To illustrate this point, it was estimated that seaborne trade in 2007 was over 7,500 million tonnes, with the demand for shipping services equating to over 32,900 billion tonne-miles.[1] For this reason, from when they were first discussed by Koopmans (1939), freight rates, or the price of transporting goods by sea, have been a constant source of practical and academic interest.

The academic interest has primarily been focused on modelling the freight rates, but an ongoing debate as to the degree of stationarity, and therefore the correct models to use, has continued for almost twenty years since the advent of unit root tests. Within this debate, general and partial equilibrium models, for example those proposed by Beenstock and Vergottis (1989) and Koekebakker, *et al.* (2006), amongst others, argue that the first moment of freight rates, i.e. spot freight rate levels, are stationary. The contrasting view, as outlined in such papers as Berg-Andreassen (1997) and Kavussanos and Visvikis (2004), amongst others, proposes that unit root tests indicate that these spot freight rates are in fact non-stationary. The debate extends to the issue of the structure of volatility, or second moment, of these freight rates where Kavussanos (1997) argues that volatility exhibits persistence, while Kavussanos and Nomikos (2003) are happy to assume that there is no persistence in volatility. This thesis adds a new dimension and middle ground to the debate by suggesting that both the first and second moments of freight rates are fractionally integrated where, to the best of the author's knowledge, this is done for the first time in the shipping literature. In addition, this thesis examines, where once again, the author is unaware of previous research on this topic in the shipping literature, the third and fourth moments of

[1] One should note that a tonne-mile is defined as the transportation of one tonne of cargo over one nautical mile.

freight rates, and in particular, introduces the concept of conditional skewness and kurtosis to the shipping academic literature.

An interesting point to highlight is that the shipping market is, perhaps, one of the few markets in which the underlying good provided is a service, and hence intangible. This being said, it is worth considering that these methodologies are not market specific; already a significant amount of interest in fractionally integrated models exists in other markets, see for example Baillie, *et al.* (1996b) and Kang, *et al.* (2009), amongst many others. In addition, this methodology can be readily applied to other markets, such as the real estate market, in which real assets are traded.

In summary, this thesis aims to give an insight into the structure of freight rates, through the examination of the various moments in these series. A thorough, and correct, understanding of the structure is of great interest as freight rates play a pivotal role and form the basis of almost every function in the shipping industry, from the determination of the price of the transport service through to the valuation of second-hand vessels. For this reason a correct model for freight rates is vital for all participants in the market, from the ship-owners and charterers themselves, right on down through the market to ship-brokers, maritime lawyers, hedge funds and other auxiliary parties involved. To give a structure to this concept, this chapter begins by reviewing the structure of the underlying shipping market, before moving on to outline the four main hypotheses that form the basis of the thesis.

1.2 Review of the Shipping Market

The shipping industry can be divided into several segments, such as, those for tanker, dry-bulk, container, reefer and cruise vessel vessels; however, a more general approach, commonly taken, is to divide the shipping market into two main sub-markets, namely the liner and bulk-shipping markets. Liner shipping is generally characterised by vessels that operate along pre-specified, fixed routes according to a regular, fixed schedule, where the majority of these vessels are now container ships. Essentially, what characterises the liner sector as distinct from the bulk sector is the fact that liner operators essentially provide a complete logistics, i.e. door-to-door service. In contrast, the bulk operator’s responsibility only begins when the goods are loaded onto the vessel, and ends when they are offloaded onto the quayside. The freight rates in the liner sector are consequently generally fixed and there is usually very little negotiation between parties with respect to these. Bulk-shipping, on the other hand, is different in that the vessels in this sector usually operate when and where the charterer demands them. Additionally, freight rates in the bulk sector vary widely and are a matter of negotiation between the ship-owner, and are therefore mostly negotiated privately. Focusing on the bulk-shipping sector, this can be further sub-divided into the dry-bulk and tanker sectors, where this classification depends on the characteristics of the cargo that the vessel will carry. In general, dry-bulk vessels will carry dry cargo, such as coal, grain or iron ore, while tankers generally carry liquid cargo, such as crude oil or oil products, although combined carriers do exist which may operate in both markets, even though these are no longer in vogue.

As a result of the historic stability of freight rates in the liner industry, the consequent lack of volatility renders asset play and freight modelling strategies unnecessary and therefore costly, and provides little scope for interest from auxiliary parties, such as hedge funds. To illustrate this point, Clarksons Research Services Ltd. (2010) show that there were only 123 container vessels sold, totalling 2.6 million DWT and worth US$742 million, during 2009, as opposed to 153 tanker vessels, totalling 14.9 million DWT and worth US$3,674 million, and 584 dry-bulk vessels, totalling 32.2 million DWT and worth US$8,846 million, during the same period.[2] In addition to this, the vast majority of vessels lie within the bulk-shipping sector, where for example, ISL Bremen (2008) indicates that, in 2007, 77% of the world seaborne fleet was comprised of either dry-bulk or tanker vessels.[3] Consequently, this research will focus exclusively on the bulk-shipping sector of the shipping market; and, hence, on the freight rates exhibited in the dry-bulk and tanker sectors of the shipping market. Given this fact, however, it is very interesting to note that the freight rates in the liner sector have become much more volatile recently, mostly as a result of the credit crisis; therefore the methodologies and hypotheses outlined in this thesis may be applied to this sector as a future potential extension of the work considered here.

[2] This disparity is further illustrated in Figures A01 to A03, in Appendix A of the thesis.
[3] A graphical breakdown of the composition of the world seaborne fleet, both in terms of the number of vessels and tonnage, can be found in Figures A04 and A05, in Appendix A of the thesis.

1.2.1 The Dry-Bulk and Tanker Sectors

As mentioned above, the bulk-shipping sector is commonly sub-divided into the dry-bulk and tanker sectors. The two sectors share a number of common characteristics, the first of which is that, in both sectors, the commodities carried tend to be low-value products, hence their suitability for slow, cheap transportation in large consignments, where this is to achieve economies of scale. One reason for this is that the costs of alternative means of transportation, such as transporting these goods by air or container, is prohibitive, in addition to the fact that these goods are generally transported in such large consignments as to make other means of transport unfeasible. The second common characteristic is that these sectors constitute the majority of seaborne trade; where, for example, ISL Bremen (2008) illustrate that of the 7,572 million tonnes of cargo transported by sea, at least 58% of this was either dry- or liquid-bulk cargo.[4] The final common characteristic is that the size of the vessels, in both sectors, is measured in dead-weight tonnes (DWT), as opposed to the container sector, where ship size is generally measured in terms of the number of twenty foot containers they can carry. Stopford (2009) defines the DWT as the maximum amount of cargo, in terms of weight, that a vessel can carry, without being classified as overloaded. One should note that this measure includes the weight of any fuel, stores, water ballast, fresh water, passengers and baggage. Having established these three characteristics, one can now proceed to individually examine characteristics of each market.

Table 1.1 – Classification of Dry-Bulk Vessels

Vessel Class	Size (DWT)
Capesize Dry-Bulk Vessels	Over 80,000
Panamax Dry-Bulk Vessels	60,000 to 79,999
Handymax Dry-Bulk Vessels	40,000 to 59,999
Handysize Dry-Bulk Vessels	10,000 to 39,999

Source: Clarksons Research Services Ltd. (2010)

Beginning with the dry-bulk sector, ISL Bremen (2008) show that, although only 16% of the world fleet in 2007, in terms of vessels, was comprised of dry-bulk vessels, this sector comprised 35% of the cargo carrying capacity, and 26% of world trade, in

[4] A graphical breakdown of the composition of world seaborne trade, by commodity, can be found in Figure A06, in Appendix A of the thesis.

terms of all cargo carried. Of these commodities carried, the predominant cargoes are iron ore, coal and grain, commonly known as the major bulks, where in 2007 trade of these exceeded 1.9 billion tonnes, and where, as mentioned above, this was equal to the 26% of the total world trade by sea.[5] The size of dry-bulk vessels can range from small coastal vessels, of less than 1,000 DWT, to huge Capesize vessels of over 300,000 DWT. The major classifications of these vessels are provided in Table 1.1.

The dry-bulk trade has grown steadily over the years (particularly with the development of China and India as major economic powers) where this has been fuelled by world economic growth and the consequent demand for bulk commodities, such as coal and iron ore.[6] This growth is not immune to current economic conditions, and with the current recessions, levels of trade have fallen, although the markets appear to have picked up somewhat recently. An indication of this is that annual weighted average earnings, i.e. earnings weighted for the size of vessels, as reported by Clarksons Research Services Ltd. (2010), fell from a high of US$43,649 per day in 2007 to US$15,016 per day in 2009, although these have recovered to US$21,677 per day, for the year to date in 2010.[7]

Table 1.2 – Classification of Tanker Vessels

Vessel Class	Size (DWT)
Ultra Large Crude Carrier (ULCC) Tankers	Over 300,000
Very Large Crude Carrier (VLCC) Tankers	200,000 to 299,999
Suezmax Tankers	120,000 to 199,999
Aframax Tankers	80,000 to 119,999
Panamax Tankers	55,000 to 79,999
Product Tankers	10,000 to 54,999

Source: Clarksons Research Services Ltd. (2010)

Changing the focus to the tanker sector, the predominant cargoes here are either crude oil, generally carried by larger tankers, or refined oil products, such as heating oil, jet fuel, liquefied natural gas, to name but a few. As with the dry-bulk sector, when looking at the figures from ISL Bremen (2008), the fact tankers comprise only 25% of

[5] A graphical breakdown of the composition of trade in the three major dry-bulk commodities, by commodity, can be found in Figure A07, in Appendix A of the thesis.
[6] This is illustrated in Figure A08, in Appendix A of thesis, which shows the evolution of dry-bulk trade against world trade for the ten year period between 1988 and 2007.
[7] Figure A09, in Appendix A of the thesis, illustrates the evolution of average weighted earnings for dry-bulk vessels for the ten year period between 2001 and 2010.

the world fleet, in terms of the number of vessels, severely under-represents the importance of the sector as it comprises 41% of the world fleet, in terms of cargo carrying capacity. Moreover, trade in liquid-bulk cargoes comprised 32%, or 2.4 billion tonnes, of the total world trade of 7.6 billion tonnes.[8] The size of tankers ranges from small barges, used to transport bunkers, to the enormous Ultra Large Crude Carriers (ULCCs), which can range up to 550,000 DWT. The classification of tankers is summarised in Table 1.2. For the purposes of this study, the ULCC and Very Large Crude Carrier (VLCC) markets have been combined for ease.

As with the dry-bulk sector, the tanker sector has grown steadily over the years as world economic growth and increased demand for crude oil and oil products has increased the demand for transportation.[9] However, the tanker sector is, as stated above, also not immune to current economic conditions, and, with the advent of the latest recession, tanker rates plummeted. To illustrate this point, average annual weighted earnings for tankers, as reported by Clarksons Research Services Ltd. (2010), fell from US$44,130 per day in 2008 to $15,511 per day in 2009; however, the market has shown signs of recovery in 2010, to date, with average weighted earnings of $21,583 per day.[10]

A further small point to note regarding the dry-bulk and tanker sectors is the manner in which freight rates are quoted differs across sectors. In the tanker sector, freight rates are quoted in terms of the Worldscale (WS), which can then be converted into US$ per tonne, according to the route, using the Worldscale Book, which is revised annually. In contrast, dry-bulk sector freight rates are simple quoted in terms of US$ per tonne, regardless of the route on which the goods are transported.

Having established the characteristics of each of sector, one can now move on to explore the underlying characteristics, in terms of supply and demand, for the shipping market as a whole, and how these play a role in determining freight rates.

[8] A graphical breakdown of the composition of tanker trade, by commodity, can be found in Figure A10, in Appendix A of the thesis.
[9] This is illustrated in Figure A11, in Appendix A of thesis, which shows the evolution of tanker trade against world trade for the ten year period between 1988 and 2007.
[10] Figure A09, in Appendix A of the thesis, illustrates the evolution of average weighted earnings for tanker vessels for the ten year period between 2001 and 2010.

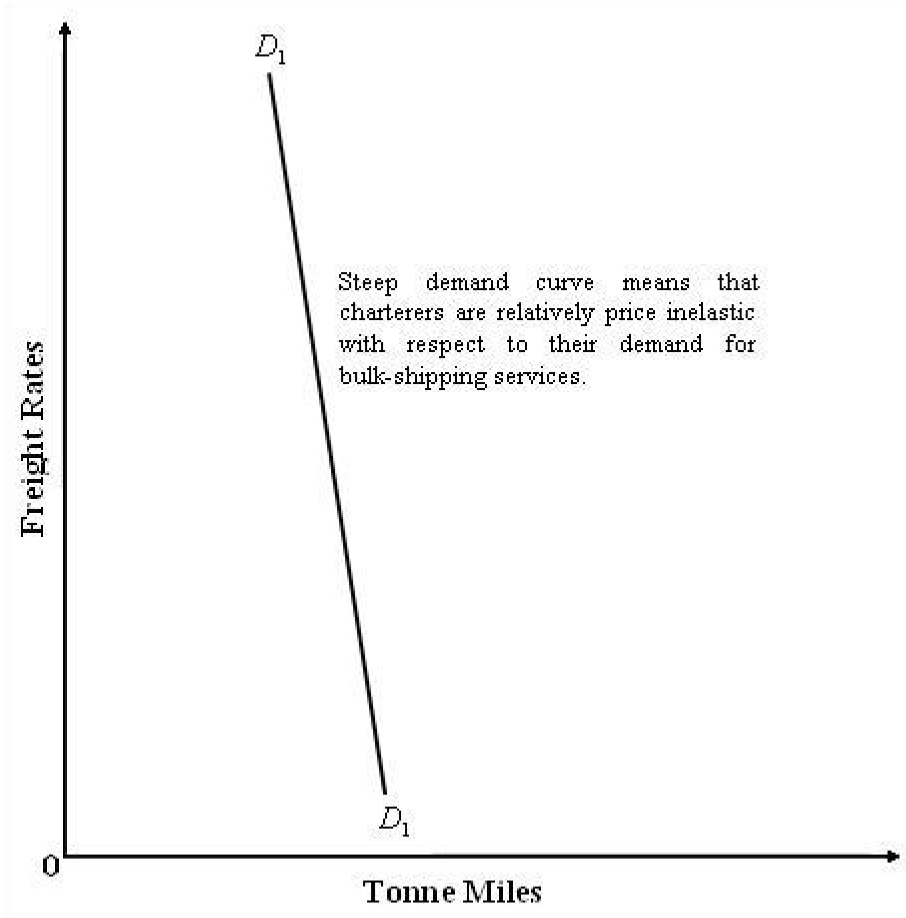
1.2.2 The Demand for and Supply of Bulk-Shipping Services

Having examined the structure of the individual sectors of the shipping market, one can now move on to examine what factors drive the bulk-shipping sector, i.e. the determinants of the demand and supply functions, both of which determine the level of freight rates in the market.

Since the demand for bulk-shipping services is a demand for the transportation of goods, this demand is derived from the demand for the goods being carried and is therefore susceptible to the cyclical nature of world trade. However, although this is a derived demand, Stopford (2009) illustrates that it can be divided into five main determinants, namely: 1) the level of world economic activity; 2) the level of seaborne commodity trade; 3) the average haul, or, in other words, the distance over which the commodities must be carried; 4) current political events; and, finally, 5) the level of transportation costs.

If the level of world economic activity is high, there will, of course, be a high demand for commodities, and, consequently, a high demand for the means with which one would transport these, hence, the demand for bulk-shipping services will increase. Moreover, even if the above does not hold, if the level of seaborne commodity trade is high, then the demand for shipping will naturally increase, due to the nature of the service provided. Another factor that would increase the demand for vessels in the bulk sector would be if the average haul were to increase, for example, if new oil fields were discovered in a remote part of the world. The reason for this is that charterers would seek to take advantage of the economies of scale offered by bulk-shipping to drive down their transportation costs, as well as the fact that fewer vessels would be available for hire. Looking at the fourth factor, political events can either have a positive or negative effect on the demand for bulk-shipping, depending on whether the news is perceived as good or bad. An example of how good news for the bulk-shipping industry would impact on demand for their services was when the Suez Canal was closed during the 1970s, where tankers transporting oil from the Middle East to Europe and the US East Coast were forced to round the Cape of Good Hope, thus resulting in increased average hauls. However, on the negative side, the introduction of the Oil Pollution Act of 1990, which made ship-owners and charterers

Figure 1.1 – The Demand Function for Bulk-Shipping Services



liable under certain cases for clean-up costs in the event of a spill, meant that charterers looked for other alternatives, such as overland pipelines, to transport their cargoes, thereby reducing the demand for shipping as a means of transport. The final factor is that if transport costs are high, then bulk-shipping may become attractive to charterers as they seek to take advantage of the economies of scale provided, thereby increasing the demand for their services.

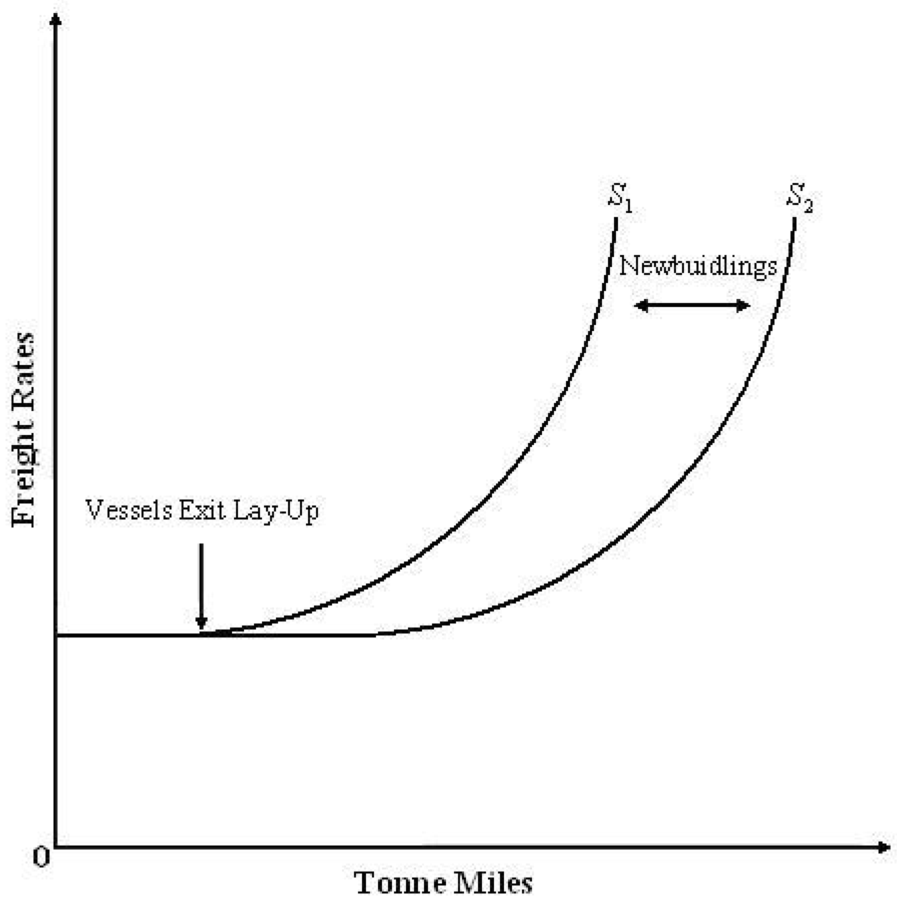
Looking at the demand function in Figure 1.1, one should note that the demand function is relatively price inelastic, i.e. demand for bulk-shipping services will only decrease by a very small amount, for a fixed increase in the freight rates. A reason for this is that, on the whole, the cost of transportation is only a very small part of the overall cost of the product, therefore, although charterers are still sensitive to price changes, they are much less sensitive than they would be, for example, with respect to the cost of refining the crude oil. Another reason for this price inelasticity in demand is that, in some cases, shipping is the only means of transport available to the charterer, and therefore it is a case of necessity to pay whatever is demanded to be able to transport the commodities.

As mentioned above, ISL Bremen (2008) show that, in 2007, over 7.5 billion tonnes of cargo were transported by sea, and these were transported over 4,439 billion nautical miles, corresponding to a demand for shipping of 22,018 billion tonne-miles. Of this demand, 12,440 billion tonne-miles, or 38% of the total, corresponded to the dry-bulk sector, and 10,397 billion tonne-miles, or 32% of the total, to the tanker sector, thus serving to reinforce the importance of both these sectors.

Changing focus to the supply-side of the market, Stopford (2009) argues that the supply of shipping services depends on five major factors, namely: 1) the fleet stock, i.e. the number of vessels operating; 2) ship-building production, i.e. the number of new vessels being built; 3) scrapping, i.e. the number of vessels scrapped for scrap metal, and losses, or vessels lost at sea or damaged beyond repair; 4) fleet productivity; and 5) the current level of freight rates in the market. Furthermore, one should also note that the supply of shipping is fixed in the short-term, for reasons discussed below.

Beginning with the fleet stock, it is fairly obvious that the higher the fleet stock, the greater the supply of bulk-shipping in that there are more vessels available to carry goods. Clarksons Research Services Ltd. (2010) reports that, as of May 2010, there are 7,541 dry-bulk vessels, totalling 481 million DWT, and 5,353 tankers, totalling 442 million DWT, meaning that the bulk-shipping fleet is at its highest levels for at least forty years. Moving on to ship-building production - as new vessels enter the market, so supply increases and Clarksons Research Services Ltd. (2010) report that, during 2009, 544 dry-bulk vessels and 580 tankers were delivered, corresponding to 43 and 48 million DWT, respectively, once again the highest figures for over 40 years. As far as scrapping and losses are concerned, an increase in these would lead to a corresponding decrease in the supply of bulk-shipping. Scrapping levels are relatively low at the moment, but increasing, while the current market conditions, in terms of the low freight rates, mean that a lot of the fleet is slow-steaming, i.e. not operating a the maximum speed possible, thereby further reducing the supply of shipping. Despite these three factors, the size of the fleet stock and the level of

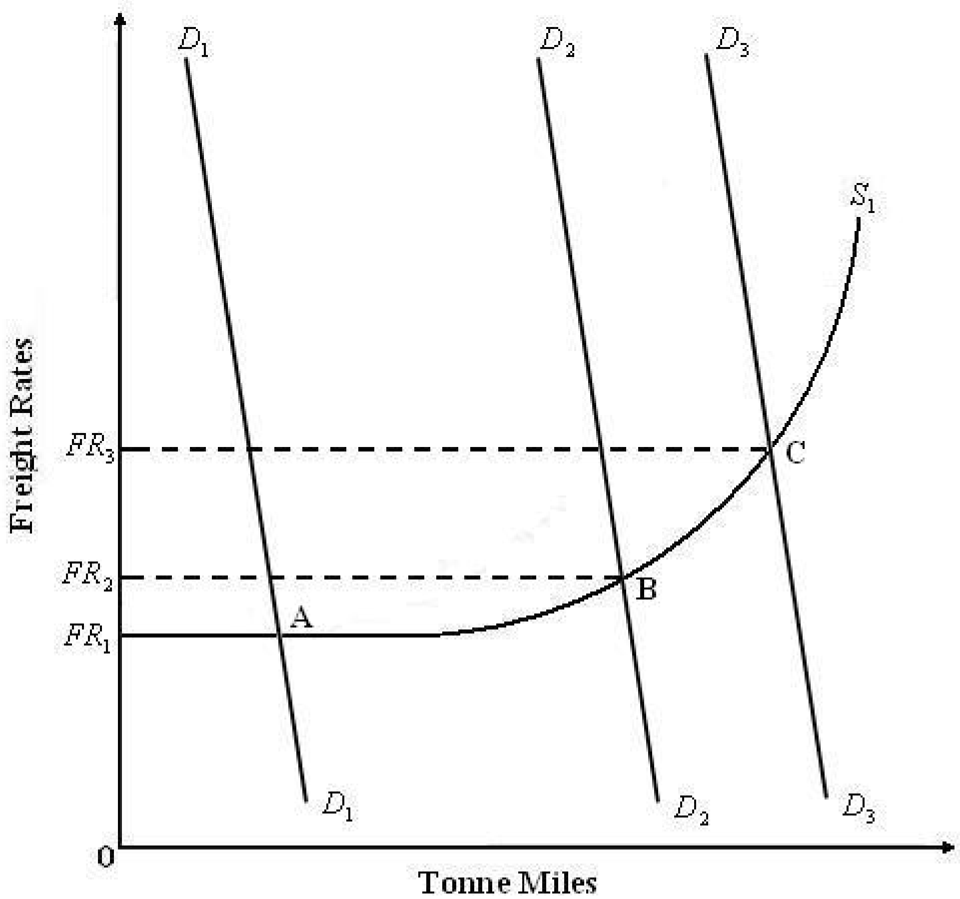
Figure 1.2 – The Supply Function for Bulk-Shipping Services



new-buildings being delivered mean that the supply of bulk-shipping is at record levels, and, given the current state of the order-book for new vessels, is likely to keep growing. This is not an ideal position for the bulk-shipping to be in, considering that the world is currently only just recovering from one of the worst recessions since the Great Depression, and where shipping is merely a means of transport for this world trade.

The shape of the supply function, illustrated in Figure 1.2, can be explained by the fact that supply is fixed in the short-term. The reason for this is that it can take up to three years, sometimes even longer, for a vessel to be delivered, i.e. from the time the vessel is ordered to when the ship-owner takes possession. This means that, with the exception of scrapping vessels or placing vessels in lay-up, when freight rates are so low that operating the vessel becomes unprofitable, the level of supply remains

Figure 1.3 – The Short-Run Market Equilibrium for Bulk Shipping Services



relatively fixed, at S_1 until the new-buildings are delivered, where the supply function in Figure 1.2 will shift outwards from S_1 to S_2. However, one should note that even delivering vessels for scrapping may take a few months; hence this may also delay any adjustments to the supply function.

Having established the characteristics of the supply and demand functions for bulk-shipping, one can now move on to examine the mechanics of the freight rate mechanism and the impact that this has on volatility in the shipping markets.

1.2.3 Freight Rates and Volatility in the Bulk-Shipping Sectors

The prevailing freight rate, which is the name for the prices of the transport service, are determined from the interaction between the supply and demand functions, discussed above. The mechanics of this freight rate mechanism are illustrated in

Figure 1.3. If we remember that the supply function is fixed in the short-run, then, as the demand function shifts along this fixed supply function, the prevailing freight rates will change accordingly. This means that the momentary equilibrium will increase (decrease), if the demand for the bulk-shipping services increases (decreases) accordingly, with the magnitude of the change in price depending on the price elasticity of the supply function at the point of equilibrium. Using the example given in Figure 1.3, if one begins at Point A, a 50% increase in demand, from D_1 to D_2, will result in a small increase in freight rates, from FR_1 to FR_2, corresponding to Point B, while a 10% further increase in demand, from D_2 to D_3, would result in a much bigger increase in freight rates, from FR_2 to FR_3, corresponding to Point C. The reason for this discrepancy is that the price elasticity of the supply function at Point B is much lower than at Point C, as there is still excess capacity in terms of supply, whereas at Point C, supply is pretty much at the maximum short-term level.

In the long-run, the supply function may shift either inwards, if the level of scrapping exceeds the level of new-building deliveries, or outwards, if the reverse applies. This ratio of scrapping vessels to new-building orders will depend on market sentiments about the future direction of freight rates. If market sentiments are good, which would correspond to ship-owners feeling that freight rates are likely to remain high or increase in the future, then the supply function may shift to the right, as the ship-owner delays any scrapping activity and places orders for new vessels to benefit from the boom. The reverse would apply when ship-owners felt that freight rates are depressed, or are likely to fall in the future, as happened last year, in which case ship-owners would seek to scrap unprofitable vessels and delay any orders. Over time, as the long-run supply of ships adjusts from a previous under- or over-supply of tonnage, freight rates will revert to the mean level that existed prior to the observable changes in the supply function.

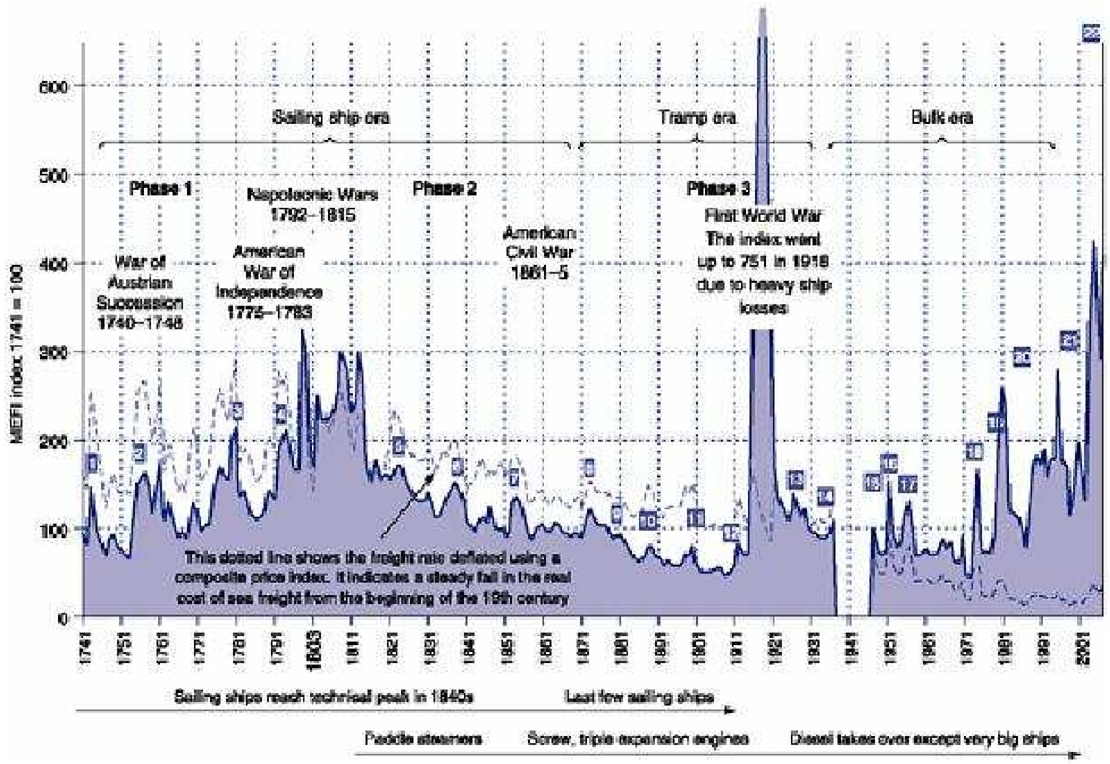
Moving to look at the volatility of shipping freight rates, or the risk inherent in the market, one should be aware that the shipping industry is highly dependent on a number of external factors, over which market agents have no control. This means that, as a result of this lack of direct control, freight rates in the shipping market are exceptionally volatile, with volatility increasing as ship-size increases. The reason for this size-volatility relationship is that smaller ships are more versatile in terms of the cargo they can carry and ports they can visit, hence they can operate in multiple markets. In contrast, a larger vessel is limited in terms of the cargo it can carry, due to the size of the cargo consignments required, and ports it can visit, as a result of draft constraints. Another reason for this volatility is the volatile nature of the demand for the goods being carried. For example, it is common knowledge that the demand for crude oil and refined oil products increases during the Northern Hemisphere winter, due to the increased demand for heating, however, this demand falls dramatically in spring. This means that the fixtures for VLCCs and other types of tankers will increase as the Northern Hemisphere winter approaches, and falls sharply during the spring months. In summary, all these factors cause the shipping industry to be a highly volatile market, hence any assistance that can be given as to predicting the nature of and relationship between factors in the industry, such as freight rates and volatility, will be in high demand.

1.2.4 Shipping Market Cycles

No study on freight rates within the shipping market would be complete without some comment on the market cycles within the industry. Shipping market cycles vary in length and frequency, but it is generally accepted that there have been 22 dry-cargo cycles between 1741 and 2007, where Figure 1.4 illustrates these cycles in the dry-bulk sector. There are generally four stages to the shipping market cycle, namely: the trough, the recovery, the peak, and the collapse.

During the first stage, i.e. the trough, freight rates are low, as a result of low demand. This would mean that ships queue up at loading ports, vessels generally slow-steam to conserve fuel and there are distress sales, i.e. sales as a result of default on loans or due to cash shortages. The second stage, or recovery, is characterised by freight rates beginning to increase, ships being removed from lay-up, and second-hand prices beginning to recover as the freight market improves. Throughout the third stage, or peak, freight rates are high, the fleet operates at full-speed, second-hand prices are above book value, order-books are almost full, and there is no idle tonnage available, so demand tends to outstrip supply. The final stage, known as the collapse, generally

Figure 1.4 – Dry-Bulk Sector Market Cycles from 1741 to 2007



Source: Stopford (2009)

occurs as a result of new orders, ordered either during the recovery or peak, being delivered, which means that the supply of tonnage now once again exceeds demand, hence freight rates plummet and no new orders are placed, with many ship-owners trying to cancel existing orders.

This collapse could occur as a result of a drop in trade, due to a slowdown in the world economy, where these effects tend to be exacerbated by negative market sentiment. Interestingly, the shipping markets are probably coming towards the end of the collapse stage and entering the trough phase of the cycle, where this is as a result of the world only now beginning to recover from one of the worst recessions since the Great Depression.

Being able to predict the future direction of freight rates, and therefore, to an extent, the market cycles, would enable one to predict when each of the stages will occur, thus allowing to profitably enter and exit the market, both in terms of day to day operations and investment timing, at the most opportune times.

1.2.5 A Brief Overview of the Prevailing Market Conditions

Having outlined the main characteristics of the shipping markets in general, this subsection provides a brief overview of the prevailing market conditions in the dry-bulk and tanker sectors in order to give context to the environment in which the research hypotheses were formed.

Examining firstly the evolution of the tanker spot freight rates, as illustrated in Figure 1.5, freight rates for the tanker market were relatively stable for the ten year period up to 2001 when they experienced the first of a series of peaks which followed over the next seven or eight years. The first peak coincides with the initiation and the process of accelerated phasing out of single-hulled tankers in favour of the double-hulled alternates, as a result of the amendment to the Marpol Convention. This led to reduction in the number of vessels in the tanker fleet and a consequent decrease in the supply of tanker services and a resultant dramatic increase in freight rates as a result of the shape of the supply curve. The second peak corresponds to the second Gulf War in 2003 as well as a further amendment to the Marpol Convention which increased the scrapping schedule of single-hulled tankers, with corresponding effects on the price of oil and supply of shipping services, and therefore freight rates. This peak then leads on to the further peaks resulting from an increased demand for oil, increased oil prices and the development boom in China, and therefore an increased demand for transportation services for the oil needed. These series of peaks were followed by an unprecedented collapse in the freight rate market in late 2008, caused by both the world undergoing arguably the most severe economic slowdown since the Great Depression, combined with massive over-ordering during the previous peaks resulting in a huge number of new tankers entering the fleet thus increasing the supply of vessels to record levels and causing an extremely and sudden fall in freight rate levels. One could also argue that these peaks may be attributed to the rescaling of the Worldscale rates in January, however, this is not felt to be a major factor as many of the peaks occur in the middle of the year.

When changing focus to look at the dry bulk market, the picture is somewhat more tranquil, as illustrated in Figure 1.6, in that freight rates remained relatively stable at fairly low levels until 2003. During the period between 2004 and 2005, a first peak is

Figure 1.5 – Evolution of Tanker Spot Freight Rates Between 13 January 1989 and 26 June 2009

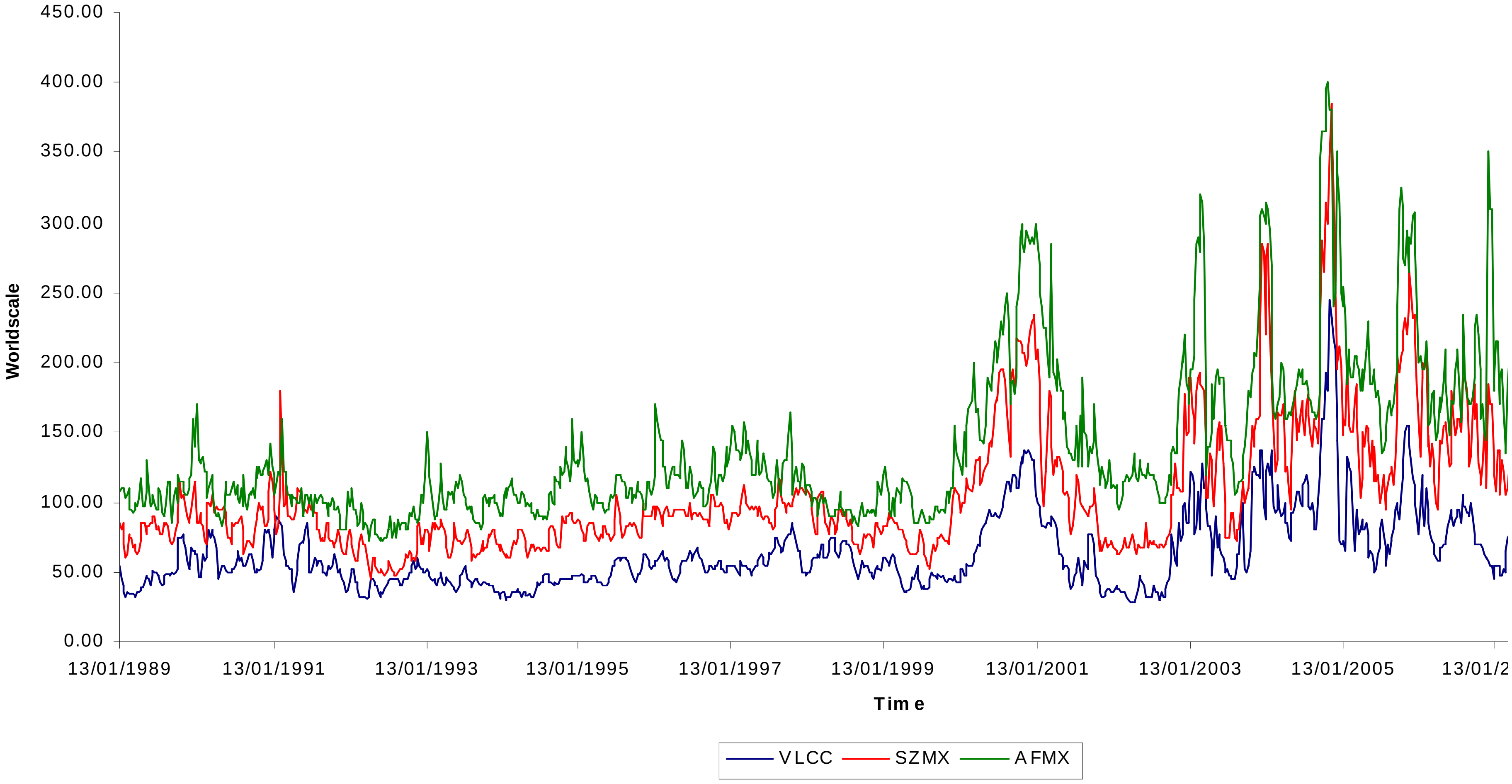
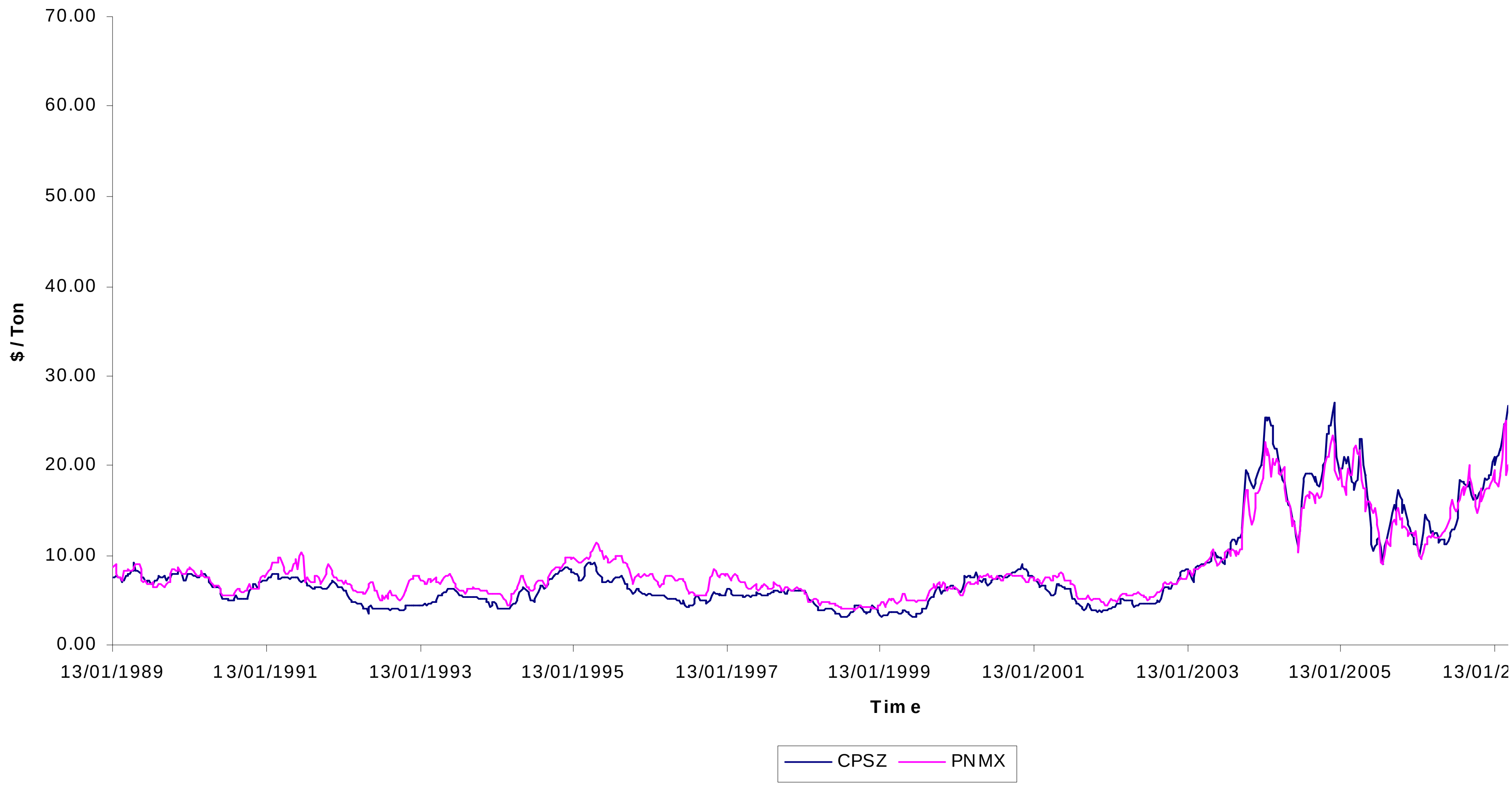


Figure 1.6– Evolution of Dry-Bulk Spot Freight Rates Between 13 January 1989 and 26 June 2009



found, corresponding with increased demand for commodities driven by the growth of the Chinese economy. The correction in the market was then followed by a massive increase in freight around the end of 2006 and beginning of 2007, driven by a rapid increase in the demand for commodities by China, congestion in world ports leading to tonnage being tied up, and a dramatic increase in the price of commodities. However, as was the case in the tanker market, a dramatic slowdown in world economic growth, as well as over-ordering, led to as extreme a fall in freight rates, although in this case, continued demand for commodities such as coal and iron ore, by China led to a much faster and somewhat greater recovery.

The following section outlines the research hypotheses with the ultimate goal of trying to achieve exactly this outcome, by understanding the structure and nature of freight rates themselves.

1.3 Research Hypotheses

The aim of this research is to expand on the traditional models of the structure of freight rates through the use of Autoregressive Fractionally Integrated Moving Average (ARFIMA), Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) and Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis (GARCHSK) models. Once this has been done, these models will be used to forecast spot freight rate levels and freight rate volatility, and then evaluate the underlying risk for market participants through the use of the Value-at-Risk methodology. This will give one an understanding of the behaviour of the different moments of freight rates, thereby enabling participants in the shipping markets to have a better understanding of both the direction of spot freight rate levels, and the underlying risk.

This research is of interest to a number of different parties, both in terms of participants in the shipping market, as well as financial markets as a whole. One reason for this is that it aims to add another dimension to the debate as to the exact structure of freight rates, as well as the degree of stationarity of these, as well as providing insight as to how the higher moments would affect freight rate risk as a whole. As mentioned above, this is of interest to participants in the shipping markets as this kind of understanding is crucial for investment timing and planning decisions, as well as for indirectly linked parties to be able to quantify their exposure to the market. Interestingly, the fact that the shipping freight market is perhaps one of the few markets in which the underlying asset is a service, as well as the fact that it may be used as a proxy for world trade, means that this market can be of interest to participants in financial markets.

1.3.1 Background to Research Questions

As mentioned above, shipping is one of the few markets in which the good being priced is a service, as well as being one of only a few markets in which the underlying asset is a real asset. This research has been planned to provide critical insight into the function of the price series within this very different type of market. Within the shipping market, freight rates play a pivotal role in that they form the basis of almost every function, from determining the cost of transporting goods from point A to point B, to the valuation process for vessels themselves. Therefore, a deep understanding and correct modelling of freight rates is essential for all participants in this market, from the ship-owners and charterers themselves, right on down to ship-brokers, maritime lawyers, institutional investors, such as hedge funds, and other auxiliary parties.

Perhaps one of the most obvious uses of freight rate modelling is for decision-making purposes, where, through the forecast of freight rates and the respective market risk, charterers can determine when it is optimal to transport their cargo, and ship-owners can determine where to position their vessels as well as when to enter and exit the market. This would mean that ship-owners could maximise their earnings by repositioning their vessels prior to the freight rate rising, and make allowances for falling freight rates, such as re-arranging the financing of their vessels. In contrast, while charterers could minimise their transportation costs by planning to transport as much as possible of their goods only when market conditions are favourable.

Another, although perhaps not quite as obvious, reason for their importance is that they allow investment timing decisions to be made. For ship-owners, as freight rates form the basis on which prices of new and second-hand vessels are valued, a thorough understanding of the structure of freight rates would mean that they could determine the ideal time to invest in new or second-hand tonnage, or either sell or scrap their vessels. This type of asset play in the market can be crucial for risk management purposes, due to the notoriously volatile nature of the market. An example of this is highlighted by Clarksons Research Services Ltd. (2010), who illustrate out that year-on-year returns from tanker earnings between 2001 and 2010 ranged from -62.43%, in 2009, to 88.85%, in 2003. Furthermore, in the dry-bulk sector, this ranged from -19.84%, in 2005, to 151.84% in 2010. To give an idea of the size of this market, Clarksons Research Services Ltd. (2010) show that at the peak of the market, in 2007, trade in second-hand vessels was worth US$47.8 billion, consisting of 1,873 individual transactions and 85.6 million DWT. For institutional investors, such as banks providing loans, or hedge funds, this is essential for both evaluating their risk exposure to the market, and identifying the ideal time in which to invest in the market. As ships are predominately financed through loans from banks, any fall in freight rates, and hence the value of the ship, would place their loans in the precarious position of a possible default. For other investors, in derivatives products, such as Forward Freight Agreements (FFAs), where these are based on analysts' assessments of future freight rate direction, an understanding of the structure of these would mean that they could determine the optimal time to invest.

In order to correctly model this, one needs to address the debate regarding the degree of stationarity of freight rates and the persistence of volatility in the market, as well as understand the impact that higher moments have on market risk in this context. This thesis proposes that freight rate levels are neither purely stationary, nor non-stationary, but that they follow a fractionally integrated process. In addition, this thesis proposes that this argument extends to the volatility of freight rates, where volatility also follows a fractionally integrated process. This implies that volatility does not decay rapidly, as implied by the traditional Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models, utilised in most papers on volatility in the shipping markets; however, neither does it persist indefinitely, as would be implied by the Integrated Generalised Autoregressive Conditional Heteroscedasticity (IGARCH) model. The final proposal in this thesis is that higher moments have an impact on the inherent risk in the shipping market, as skewness and kurtosis evolve over time, hence, introducing the concepts of conditional skewness and kurtosis.

1.3.2 Hypothesis 1 – The Dynamics of the First Moment

Traditional general equilibrium models, such as those proposed by Hawdon (1978) and Beenstock and Vergottis (1989), suggest that freight rates are mean reverting. This implies that any imbalance in the supply and demand functions would be corrected, thus causing freight rates to revert to the mean level. In contradistinction, newer research, mostly during the 1990s, including Berg-Andreassen (1996), Glen (1997) and Kavussanos and Nomikos (2003), which coincided with the development of new tests for stationarity, found that freight rates were not mean reverting, but followed a random walk process, thus implying that they are non-stationary. This means that any imbalances in freight rate would persist indefinitely, thus leading to an "explosive" series. However, more recent research, for example papers by Adland and Cullinane (2006) and Koekebakker, *et al.* (2006), using partial equilibrium models, propose that the original assumption of mean reversion is correct, and any conclusions otherwise were as a result of deficiencies in the unit-root tests, a fact outlined by Schwert (1989). Furthermore, this literature stream argues that this mean reversion process will be almost immediate, therefore implying stationarity. Another element that may cast doubt on the validity of any assumptions of non-stationarity is the length and frequency of the data set examined. Schwert (1989) highlights the fact that the longer the data set and the greater the frequency, the greater the number of observations in the sample and the more likely one is to observe mean reversion in the data and the better the understanding of the dynamics of the data. This issue is addressed in this thesis in that the sample of freight rates consists of weekly observations for the period extending from 13 January 1989 to 26 June 2009.

This thesis puts forward, for the first time in the shipping literature, the proposal that the answer may in fact lie somewhere in between these two rival conclusions, i.e. that freight rates follow a fractionally integrated process. The rationale behind this statement are that in the short-term, the supply function for shipping services is fixed, while demand is relatively price inelastic, however, in the longer-term, as new vessels are delivered, the supply function will expand accordingly. This means that in the short-term, freight rates will exhibit non-stationary behaviour in that, due to the fixed nature of supply, as demand increases, so will freight rates, up to the point where freight rates make other, more expensive, alternative means of transportation viable, as illustrated in Figure 1.3 above. However, as high freight rates induce ship-owners to order new vessels, and these vessels are delivered, usually after between 18 and 36 months, but this can be extended to over five years, the supply function will shift to the right, as illustrated in Figure 1.2 above, and freight rates will revert to their mean level. Thus, as one can see, freight rates are mean reverting; however, this mean-reversion process will occur with a lag, which is the definition of a fractionally integrated process.

In order to test this hypothesis, Chapter 5 of the thesis presents the results from applying Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Fractionally Integrated Moving Average (ARFIMA) models to a series of freight rates, and subsequently produces forecasts based on each of these models. Whichever of the ARMA, ARIMA or ARFIMA models provides the best model of the underlying freight rate series, will demonstrate whether freight rates follow a stationary, non-stationary or fractionally integrated process, respectively.

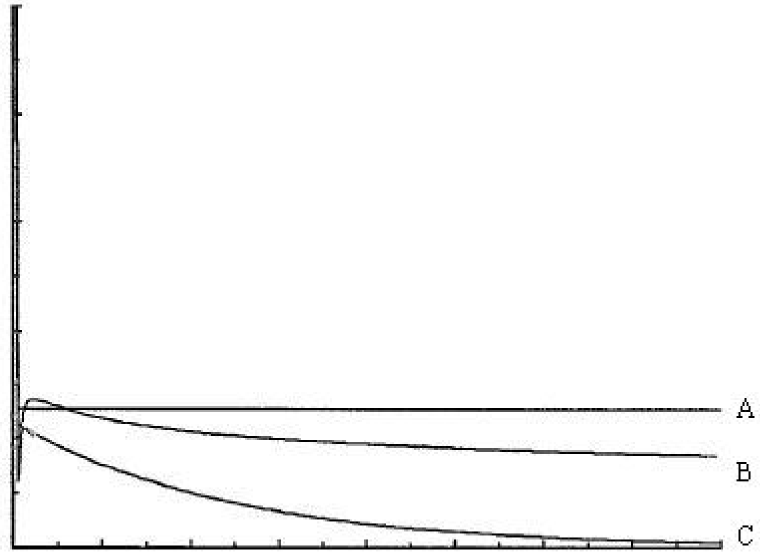
The interesting aspect of this hypothesis is that it would enable market participants to better forecast freight rates. This would enable ship-owners to better plan the positioning of their vessels, to take advantage of higher freight rates, as well as better make decisions as to the optimal time in which to invest or pull out of the market. These factors would, in turn, lead to increased profits for market participants, which could also have a run-on effect on other markets, as most of the commodities traded in the world are transported by sea. Furthermore, a better understanding of the transport costs involved would enable charterers to better forecast their costs, and potentially pass on these cost-savings to clients and participants in other markets. Better forecasts of freight rates, as stated above, would also enable a better understanding of investment timing, where these methods could then be applied to other markets in which real assets are traded. A final benefit is that there are of course the policy and decision making implications, where a better understanding of the structure of freight rates would enable one to make better decisions regarding company policies, investments, and the structure of the market as a whole.

1.3.3 Hypothesis 2 – The Dynamics of the Second Moment

Having established the dynamics of the first moment of the underlying freight rates, the obvious next question is whether a similar structure applies to the second moment, or volatility, of these freight rates. One should note here that volatility is a measure of risk in the market, where, the higher the standard deviation, or variance, the greater the level of associated risk. As has been mentioned in the review of the shipping markets above, the shipping market is exceptionally volatile and therefore any method that could be used to correctly model this underlying volatility would be most welcome. For this reason, there has been a clear interest in the study of the volatility in the shipping context, where the predominant models used have been the Autoregressive Conditional Heteroscedasticity (ARCH) family of models, first introduced by Engle (1982). Examples of this include Kavussanos (1996), who uses an ARCH model; Kavussanos (1997), who uses a Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model; and Kavussanos and Nomikos (2000b) and Kavussanos and Nomikos (2003), who use an augmented GARCH (GARCH-X) model. One should note that Kavussanos (1997) also proposes that an Integrated Generalised Autoregressive Conditional Heteroscedasticity (IGARCH) model could be used, although this was never estimated.

Having established the applicability of ARCH-type models to estimate volatility in the shipping markets, one can now move on to see how the structure of the underlying freight rate series could affect the model selection. In this respect, the concepts of stationarity, non-stationarity and fractionally integration can be extended from the spot freight rate levels to the volatility of freight rates, as illustrated in Baillie, *et al.* (1996a). This research is therefore, to the best of the author's knowledge, the first in the shipping literature to test the hypothesis as to whether the volatility series follows a fractionally integrated process. To give a graphic understanding of what is meant, one should examine the different impulse response functions in Figure 1.7, where this measures the time it takes for a shock to volatility to dissipate. In this sense, should shocks to the volatility decay exponentially, as indicated by impulse response function C in Figure 1.7, where $\lambda_k = \left(\phi_1 - \beta_1\right)\phi_1^{k-1}$; $k > 1$, then the volatility series could be argued to follow a "stationary" process; while if these shocks persist indefinitely, then the volatility

Figure 1.7 – Cumulative Impulse Response Functions



Source: Baillie, *et al.* (1996a)

series would follow a "non-stationary" process, as portrayed by news impact curve A, in Figure 1.7, where $\lambda_k = (1-\beta_1)$; $k > 1$. Once again, this research proposes that a middle ground, where, should shocks to the volatility decay in a hyperbolic manner, as illustrated by news impact curve B in Figure 1.7, where $\lambda_k \left[1-\beta_1-(1-d)k^{-1}\right] \cdot \Gamma(k+d-1)\Gamma(k)^{-1}\Gamma(d)^{-1}$; $k > 1$, then the volatility series could be argued to follow a "fractionally integrated" process. The rationale behind this hypothesis is the same as for the spot freight rate levels. Imbalances in supply and demand in the short-term cause freight rate levels to "explode". Consequently, the volatility, or standard deviation, of these freights will also increase dramatically, until such a time as the level of spot freight rates stabilises. As new vessels are delivered, spot freight rates revert to the mean spot freight level, and volatility stabilises, however, this process of stabilisation occurs with a lag, due to the fixed nature of supply in the short-term.

In order to test this hypothesis, Chapter 6 of the thesis presents the results from GARCH, IGARCH and Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) models. Whichever of the GARCH, IGARCH or FIGARCH models provides the best model of the underlying volatility, will be demonstrate whether freight rate volatility follows a stationary, non-stationary or fractionally integrated process, respectively.

As with the structure of the first moment, the structure of the second moment of freight rates is of particular interest for investment timing decisions. Ship-owners, charterers, hedge funds, and other market participants are all, in one way or another, exposed to volatility in the freight rate markets. By being able to better understand the underlying structure of this volatility, market participants are able to determine the ideal time to enter and exit the market in order to minimise the market risk exposure. In addition, auxiliary parties, such as hedge funds, may seek to take advantage of this volatility in order to trade in freight rate derivatives. Finally, an understanding of volatility and the inherent risk in the market is essential for portfolio optimisation, whether of vessels themselves, or freight rate derivatives, and is vital for determining one's correct Value-at-Risk, although a more correct term in the context of this thesis would be the ship-owner's or charterer's Profit-at-Risk, as one is not trading a portfolio here but minimising the risk exposure of market participants' profits. One should note that these techniques can once again be applied to any market in which real assets are traded.

1.3.4 Hypothesis 3 – The Dynamics of the Higher Moments

Moving on with the analysis of the moments, this research examines the higher, i.e. third and fourth, moments. Incorporating skewness and kurtosis into models of price series is well established, however, a relatively new introduction to the financial markets literature is the concept of conditional skewness and kurtosis. The concept of conditional higher moments was introduced by Harvey and Siddique (1999), who developed the Generalised Autoregressive Conditional Heteroscedasticity with Skewness (GARCHS) model to measure the impact of conditional skewness on the volatility of stock prices. This work was extended by Brooks, *et al.* (2005) who examined the impact of conditional kurtosis, proposing the Generalised Autoregressive Conditional Heteroscedasticity and Kurtosis (GARCHK), on stocks and bonds. To amalgamate both these concepts, i.e. conditional skewness and conditional kurtosis, León, *et al.* (2005) developed and proposed the use of the Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis (GARCHSK) model, to model the impact of these conditional higher moments on the returns from a series of stocks and exchange rates, thereby illustrating the importance of these conditional moments in determining the true structure and risk inherent in these series,.

This thesis adopts these concepts, and tests the hypothesis that conditional skewness and kurtosis also plays a significant role in the structure of the underlying freight rates, where this is, to the best of the author's knowledge, the first time this is examined in the shipping literature as well as the literature for markets in which real assets are traded. The rationale behind this hypothesis is that the shape of the supply function in the freight markets is such that when one is positioned at a relatively price elastic portion of the supply curve, the degree of skewness and excess kurtosis will be relatively low. This being said, as the price elasticity decreases, as short-term supply reaches its maximum level, and freight rates shoot up, so will the degree of skewness and excess kurtosis, resulting in an extremely fat-tailed, positively skewed distribution.

As with the second moment, skewness and kurtosis play a huge role in determining the market risk exposure of participants in the freight rate markets. For this reason, an understanding of these is vital for investment-timing and decision making purposes. In addition, as Christoffersen (2003) highlights, if one does not incorporate these into any Value-at-Risk calculation, one significantly underestimates the risk exposure of market participants. One can conclude by adding that these concepts can, as with the previous two, be readily applied to other markets in which real assets are traded.

1.4 Structure of the Thesis

Having outlined the aims and contributions of the thesis to shipping market literature, this section outlines the organisation of the thesis, which consists of the nine chapters, including this introduction. The general structure of each of the four empirical chapters, i.e. Chapters 5 through 8, is similar in that each begins with the general aim of the discussion, before moving onto a brief discussion of the methodology, a description of the data, a thorough analysis of the empirical findings, and a final brief summary of the findings and conclusion.

Chapter 2 of the thesis is devoted to an in depth discussion and review of the market literature regarding each of the hypotheses discussed. It begins by outlining the general theorems regarding the structure of the shipping markets, before going onto discuss the issue of stationarity, with respect to the spot freight rate levels. Following this, the chapter outlines the literature regarding the impact of volatility in the shipping market, and how FIGARCH models have been used in other markets. The chapter continues with a discussion of the current literature on conditional higher moments, and concludes by clearly outlining and summarising the contribution of this thesis to the existing literature.

Chapter 3 outlines the methodology to be used in the empirical chapters. This begins with a particular focus on the ARMA, ARIMA and ARFIMA models for spot freight rate levels, as well as how one would evaluate forecasts of these. Following this a discussion of the GARCH, IGARCH and FIGARCH models, and the implications therefore, before moving onto look at the higher moments, with a particular focus on the GARCHSK model.

A general description of the data is provided in Chapter 4, which includes a discussion on such matters as the data sourcing, sample period and general characteristics of the data. In addition, standard unit root tests are performed to give a preliminary idea as to the degree of integration of the spot freight rate levels, as well as tests to determine the degree of autoregression in the various moments.

The first hypothesis regarding the degree of integration is addressed in the fifth chapter of the thesis. Tests for fractional integration are also performed on the residuals of the models to determine if the models have been properly specified. Following this, forecasts are performed and evaluated to determine the best model with which to identify the future direction of the spot freight rates themselves.

Chapter 6 introduces the concept of fractional integration, in the volatility series, to the shipping literature and highlights the degree of persistence in the shocks to volatility. Chapter 7 extends the discussion in Chapter 6 by introducing the GARCHSK model to the analysis, and highlights how this contributes to the understanding of the true structure of freight rates.

Chapter 8 highlights the practical applications of freight rate volatility models. It highlights the difference in the performance of different standard Value-at-Risk methodologies and selects the best model, given a set of pre-determined critical levels, for determining the Profit-at-Risk, Costs-at-Risk and Value-at-Risk for ship-owners, charterers and auxiliary parties in the shipping freight market, respectively. It also aids in highlighting the contributions of Chapters 6 and 7 to understanding the concept of risk in the market.

The final chapter, Chapter 9, presents a summary of the thesis findings and highlights the main conclusions. The implications and limitations of the findings of each empirical study are also discussed further. The thesis concludes with suggestions for future research.

2 Literature Review

2.1 Introduction and Key Papers

From the time when they were first discussed by Tinbergen (1931), Tinbergen (1934) and Koopmans (1939), freight rates have been a constant source of academic interest. The reason for this, as outlined in Beenstock and Vergottis (1989), amongst others, is that freight rates form the basis on which all shipping decisions are made, from when and where to operate, to the crucial investment decisions of when and where to buy vessels. This research focuses on three main hypotheses, namely the use of Autoregressive Fractionally Integrated Moving Average (ARFIMA), Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) models, and Generalised Autoregressive Heteroscedasticity with Skewness and Kurtosis (GARCHSK) models, to model the freight rate process in the shipping freight market. This section aims to examine the relevant literature on each of these topics and critically evaluate its potential application to solve the issues raised in the research hypotheses in Chapter 1 of the thesis.

The literature review begins with an analysis of the first econometric models of freight rates and the shipping markets. These provide a basic understanding of the determinants of freight rates, and how these behave under different market conditions. The aim of this is to give an understanding of the functioning of the markets, and how the various factors interact to form the general structure of freight rates in these different shipping markets.

Section three of this literature review examines the stationarity of freight rates, beginning with the structural models, which argue that, due to the supply and demand dynamics of the freight markets, the constant adjustment of supply to a relatively inelastic demand will cause freight rates to follow a mean reversion process. The reason for this is that when freight rates are low, the supply of shipping services will naturally also be low; however, as the demand for shipping services increases, so supply will gradually increase in response until excess supply is exhausted. If the level of demand continues to increase, due to the lag between the ordering and delivery of new vessels, freight rates will increase exponentially; however, as soon as the new vessels are delivered to the market, the increase in supply will lead to freight rates decreasing back to the mean level. The literature then moves on to an argument that this traditional maritime economic theory is incorrect and that modern econometric tests, such as the Augmented Dickey-Fuller test, proposed by Dickey and Fuller (1981), and Phillips-Perron test, introduced by Phillips and Perron (1988), amongst others, in fact prove that freight rates are non-stationary. This literature then moved on from this to use cointegration analysis to forecast future freight rates. Finally, the most recent literature proposes that in fact the structural models were correct, and that freight rates are indeed stationary, the error lying in the relative weakness of the stationarity tests utilised. Therefore, by applying still more modern econometric tests, such as the KSS test proposed by Kapetanios, *et al.* (2003), which is used in the paper by Koekebakker, *et al.* (2006), it is proposed that freight rates are indeed stationary. The aim of the current research is to contribute to this existing literature and maritime economic theory by proposing that in fact neither of these arguments are correct, and that the answer to the question of the stationarity of freight rates lies somewhere in the middle, i.e. that that freight rates are fractionally integrated.

The fourth section of literature looks at risk in the form of the volatility of freight rates; however, unfortunately, there has been very little published on the topic. Volatility in the shipping freight markets was originally believed to be best modelled using the Autoregressive Conditional Heteroscedasticity (ARCH) models, developed by Engle (1982). However, later papers found that the use of the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) framework, proposed by Bollerslev (1986), provided the most appropriate parameterisation for volatility. As the literature regarding freight rates moved towards freight rates following a non-stationary process, so did the literature on volatility, where the use of the Integrated Generalised Autoregressive Conditional Heteroscedasticity (IGARCH) framework, outlined in Engle and Bollerslev (1986), as was proposed by Kavussanos (1997) but never actually implemented. The aim of the current research, in this respect, is to contribute to the existing literature by arguing that, as there is long memory in freight rates, there is long memory in the volatility of freight rates, and therefore the most appropriate model to use to estimate volatility would be a FIGARCH model.

Section five looks at the concepts of fractional integration and long memory and how these have been used in different financial markets, with a particular focus on the stock, exchange rate and interest rate markets. It reviews both the concepts of fractional integration in terms of price levels and volatility to give an understanding of how these work, and how they may be applied in the shipping context. It therefore provides a link between the second and third sections of the literature review.

The sixth section of literature looks at the impact of higher moments on the behaviour of price series. This literature argues that as the behaviour of volatility can vary across time, i.e. the concept of conditional volatility, so can the behaviour of skewness and kurtosis, thus introducing the concepts of conditional skewness and kurtosis. The argument here is that the assumption of constant skewness and kurtosis leads to market participants severely underestimating their risk exposure, hence Harvey and Siddique (1999) and Brooks, *et al.* (2005) developed the Generalised Autoregressive Conditional Heteroscedasticity with Skewness (GARCHS), to enable one to model conditional skewness, and the Generalised Autoregressive Conditional Heteroscedasticity and Kurtosis (GARCHK) model, to examine conditional kurtosis, respectively. The limitations of these models, in that the GARCHS model enabled one to model conditional skewness but not conditional kurtosis, and vice versa for the GARCHK model, led León, *et al.* (2005) to propose the Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis (GARCHSK) model, which enabled one to jointly analyse the impact of conditional skewness and kurtosis. With these concepts in mind, this research aims to contribute to the literature by examining, for the first time in the shipping literature, the impact of these conditional higher moments on the structure of freight rates and then evaluate how this impacts on the risk of market participants using the Value-at-Risk methodology.

The final section of literature examines the different econometric methodologies used to forecast freight rate levels and volatility in the shipping freight markets. This research aims to contribute to the literature by introducing the notions of fractional integration and conditional higher moments to the shipping literature, and then comparing the results to provide a conclusion as to the impact of these in the market.

2.1.1 Key Papers

2.1.1.1 Early Econometric Models of the Shipping Markets

1. Tinbergen (1931)
2. Tinbergen (1934)
3. Koopmans (1939)

2.1.1.2 Key Papers Regarding the Stationarity of Freight Rates

Papers on Structural Models

1. Zannetos (1966)
2. Hawdon (1978)
3. Beenstock and Vergottis (1989)

Papers on Non-Stationarity and Cointegration

1. Berg-Andreassen (1996)
2. Berg-Andreassen (1997)
3. Veenstra and Franses (1997)
4. Kavussanos and Nomikos (1999)
5. Kavussanos and Nomikos (2003)
6. Kavussanos and Visvikis (2004)

Papers on Partial Equilibrium Models

1. Tvedt (1997)
2. Tvedt (1998)
3. Tvedt (2003)
4. Adland and Strandenes (2004)
5. Adland and Cullinane (2006)
6. Koekebakker, *et al.* (2006)

2.1.1.3 Key Papers Regarding the Volatility of Freight Rates

1. Kavussanos (1996)
2. Kavussanos (1997)
3. Kavussanos and Visvikis (2004)

2.1.1.4 Key Papers on Long Memory and Fractional Integration

1. Adelman (1965)
2. Baillie (1996)

2.1.1.5 Key Papers on Conditional Higher Moments

1. Harvey and Siddique (1999)
2. Brooks, *et al.* (2005)
3. León, *et al.* (2005)
4. Bali, *et al.* (2008)

2.1.1.6 Key Papers on Forecasting in the Shipping Markets

1. Batchelor, *et al.* (2007)
2. Angelidis and Skiadopoulos (2008)

2.2 Some Early Models of the Shipping Markets

Despite the rapid growth of international seaborne trade in the early 1900's, as well as the increasing importance of the shipping industry in providing the means with which to connect sources of supply and demand for different types of commodities, it was not until the 1930's, with the pioneering studies by Tinbergen (1931), Tinbergen (1934) and Koopmans (1939), that the foundations of the analysis of this industry were set out. In the first of these, Tinbergen (1931) provided the first quantitative analysis of the dynamics of the shipbuilding market, in which the important variables for this market were empirically identified. In a later study, Tinbergen (1934) investigated, for the first time, the formulation of freight rates using a supply and demand framework, introducing the concept of market equilibrium to the market, with a particular focus of the sensitivity of these freight rates to changes in the factors affecting the supply of and demand for shipping services. The factors included the price of bunkers, the size of the shipping fleet, and the fact that the demand function for shipping services is price inelastic. Koopmans (1939) provided the first attempt to analyse the shipping freight market, in which the behaviour of the supply and demand schedules for shipping services under different market conditions are distinguished. These studies provided the foundation upon which subsequent studies in the literature on shipping and shipbuilding markets are built. For this reason, these studies are discussed in more detail below.

2.2.1 Tinbergen (1931) and the Dynamic Shipbuilding Model

Tinbergen (1931) provided the first empirical analysis of the cyclical nature of the shipbuilding market. The model developed in this paper provided the basis for all subsequent studies in the literature through the connection of shipping freight rates and shipbuilding activities via the size of the world fleet, which Tinbergen denoted K_t. Tinbergen began the analysis by assuming that should the world fleet increase (decrease) in size, i.e. should there be an increase in the supply of ships, this would results in a negative (positive) effect on the prevailing freight rates in that market, denoted FR_t, hence:

$$FR_t = f_1\left(\overset{-}{K_t}\right) \qquad (2.1)$$

In addition to this, Tinbergen argued that any change in the size of the fleet, denoted ΔK_t, where this is adjusted for losses and scrapping activity, will be proportional to the orders placed k periods earlier, denoted OR_{t-k}, where k denotes the lag between an order being placed and the order being delivered. This means that:

$$\Delta K_t = f_2\left(\overset{+}{OR_{t-k}}\right) \qquad (2.2)$$

Tinbergen then assumed that the level of new orders at period $t-k$, i.e. OR_{t-k}, are positively related to the level of freight rates at that period, denoted FR_{t-k}, thus:

$$OR_{t-k} = f_3\left(\overset{+}{FR_{t-k}}\right) \qquad (2.3)$$

Tinbergen (1931) thus derived a model, through substituting Expression (2.3) into Expression (2.2), which relates the expansion, or contraction, of the world fleet size to the levels of freight rates at time $t-k$. One should note that, in Expressions (2.1) to (2.3), the signs above variables correspond to the signs of the partial derivatives of these variables. Following this, Tinbergen continued the analysis by estimating the model, using data for the period between 1870 and 1913, reaching the conclusion that the shipbuilding industry follows a cyclical pattern, where each of these cycles has a duration of approximately eight years from peak to peak.

2.2.2 Tinbergen (1934) and the Shipping Freight Rate Model

In a subsequent study Tinbergen (1934) provided the first study of shipping freight rates within a supply-demand framework. Within this framework, Tinbergen evaluated the sensitivity of shipping freight rates to the determinants of supply and demand, where there are denoted Q^S and Q^D, respectively. Tinbergen proposed that the important determinants of the level of supply in the shipping are the bunker, or fuel, prices, denoted BP, the size of the world fleet, denoted K, and the prevailing freight rates in the market, denoted FR. On the demand-side, Tinbergen argued that the demand for shipping services is perfectly price inelastic, as changes in freight rates do not appear to influence the level of demand. Based on these assumptions, the following equations for supply and demand were proposed:

$$Q^S = f\left(\overset{+}{K};\overset{-}{BP};\overset{+}{FR}\right) \qquad (2.4)$$

$$Q^D = \text{inelastic demand} \qquad (2.5)$$

One should note that, in Expression (2.4), the signs above variables correspond to the signs of the partial derivatives of these variables. Tinbergen proposed that the level of supply, which is measured in tonne-miles, is negatively related to the price of bunkers as an increase in this will force ship-owners to reduce the speed of their vessels in order to reduce and optimise fuel costs. In contrast, the level of supply is positively related to the size of the world fleet as it is obvious that an increase in the size of the fleet will directly increase the level of supply in the market. Additionally, freight rates are positively related to supply as an increase in freight rates will incentivise ship-owners to increase the speed of their vessels as well as take idle vessels out of lay-up as shipping operations become more profitable.

Tinbergen (1934) continued by proposing that, under the market clearing assumption, freight rates move instantaneously to bring the market into equilibrium, where $Q_t^S = Q_t^D$. This implies that one can therefore derive the freight rate equation by solving the system of equations in Expressions (2.4) and (2.5) as follows:

$$FR_t = f\left(\overset{+}{Q_t^D}; \overset{-}{K}; \overset{+}{BP} \right) \qquad (2.6)$$

Following this, Tinbergen estimated the model in Expression (2.6), using annul data for the period between 1870 and 1913. This is done using the following log-linear form in order to determine the significance and elasticity's of the respective variables, where:

$$\ln FR_t = \alpha \ln Q_t^D + \beta \ln K_t + \gamma \ln BP_t \qquad (2.7)$$

Tinbergen reported that all the estimated parameters from Expression (2.7) have the correct sign, i.e. $\alpha > 0$, $\beta < 0$ and $\gamma > 0$. Tinbergen this established the important influences that demand, supply and bunker prices have on the determination of the prevailing freight rates.

2.2.3 Koopmans (1939) and Tanker Freight Rates and Shipbuilding

Koopmans (1939) provided the first study within the shipping literature to distinguish between the dry cargo and tanker sector. In this study, Koopmans examined the tanker freight market using a supply-demand framework, where most of the theory proposed by Tinbergen (1931) and Tinbergen (1934) was examined using tanker market data for the period between 1920 and the mid 1930's. Koopmans' most interesting contribution, apart from examining the tanker market as a separate entity, was that the study distinguishes between periods of prosperity and depression in the tanker markets. This meant that Koopmans could explain how the price elasticity of the supply schedule would change from being relatively elastic when freight rates are low, i.e. the market is depressed, to being inelastic when freight rates are high and almost all the fleet is employed, i.e. the market is experiencing a period of prosperity.

2.2.4 Criticisms of Early Econometric Models

Although these models form the basis of further empirical analyses of the shipping markets, they are still open to criticism. The first obvious criticism is that Tinbergen (1931) and Tinbergen (1934) completely ignored the fact that the bulk shipping market is comprised of two separate sectors, namely the dry-bulk and tanker sectors, each with their own particular characteristics of supply and demand. Although this issue was somewhat addressed by Koopmans (1939), in that the study focuses on the tanker sector, this still means that the dry-bulk sector was largely ignored, where Chang and Chang (1996) argued that the dry-bulk sector is exposed to the most risk as it is the most perfectly competitive sector. A further criticism is that, with the exception of Koopmans (1939), no attempt was made to distinguish between the behaviour of supply and demand at various stages in the shipping cycle. A final criticism was that no attempt is made in any of these papers to determine whether the shipping markets are efficient or provide any form of forecast of the future direction of shipping freight rates. These criticisms are partially addressed in the following section, which investigates the question as to whether freight rates follow a stationary or non-stationary process, a question to which this thesis provides an alternative answer by proposing that these are in fact fractionally integrated.

2.3 Are Spot Freight Rate Levels Stationary?

As mentioned above, since the pioneering work of Tinbergen (1931), Tinbergen (1934) and Koopmans (1939), there has been tremendous academic interest in the area of maritime economics, particularly on the structure of freight markets and the modelling of the spot freight rate for bulk shipping. Having an understanding of the structure of these freight rates, ship-owners charterers, and other participants in the shipping markets may be more secure as to the accuracy of models and forecasts made with regard to the future direction of freight rates, and, accordingly, therefore more able to make the correct operational and investment decisions. This section focuses on the structure of freight rate levels, and in particular on the question as to the degree of stationarity of freight rate levels in the shipping freight market?

Before going any further it is worth noting what is meant by a stationary time series. Brooks (2002) defines a series as weakly stationary, what is commonly meant as a stationary series, if it has a constant mean, a constant variance and a constant autocovariance. Engle and Granger (1987) extend this argument by stating that a series, given that the it has no deterministic component, is said to be integrated of order *d*, i.e. $I(d)$, if this series is a stationary, invertible, Autoregressive Moving Average representation after it has been differenced *d* times. For example, a series is said to follow an $I(1)$ process if the first-difference of the series is stationary.

2.3.1 Structural Models of Shipping Freight Rates

The early econometric models of the shipping industry, discussed above, provided the first ideas regarding the formation of freight rates and the structure of the shipping freight markets. These ideas were then expanded by Zannetos (1966) to provide the first complete structural model, which was in turn expanded on by Hawdon (1978) and Beenstock and Vergottis (1989). The basis of these models was to use supply and demand fundamentals to develop a definitive model in order to forecast the future direction of spot freight rates. This section reviews these papers and provides a brief critical commentary on their value.

2.3.1.1 Zannetos (1966) and the Structure of Tanker Freight Rates

Zannetos (1966) extended the earlier econometric model to provide one of the first structural models, and the first study to distinguish between the determination of spot and time charter freight rates. Although this extensive study primarily dealt with the relationship between spot and period freight rates, Zannetos argued that prevailing spot freight rates, i.e. short-term freight rates, in the tanker market are solely a function of the number of ships in lay-up at any point in time. In order to test this hypothesis, Zannetos tested the relationship that:

$$S_t = a + b\left(\frac{1}{LU}\right) \qquad (2.8)$$

One should note that, in Expression (2.8) above, S_t denotes the tanker spot freight rate level, while $1/LU$ denotes the proportion of the tanker fleet in lay-up. It is worth mentioning that $1/LU$ is included to provide for the empirical observation that tanker freight rates never fall below certain rates, known as the so-called lay-up points.

Hawdon (1978) highlighted the fact that the validity of this simple model depends on the assumption that autonomous shifts in demand occur, and that the supply schedule in the tanker market is fixed. Although this assumption would hold in the short-term, the supply function can change when looking at a longer horizon. Hawdon proposed that Zannetos (1966)'s assumption can be tested by examining the correlation between the spot freight rate levels and $1/LU$ for sub-periods of approximately equal length, the results of which indicate that the supply curve has indeed shifted significantly between sub-periods.

Hawdon (1978) suggested several possible technical reasons for the failure of this hypothesis. The first of these is that the development of combined carriers led to a reduction in the specificity of the tanker fleet, enabling these vessels to switch to the dry-bulk market, rather than going into lay-up. Another reason is that vessels have significantly increased in size, hence the subsequent economies of scale would mean that vessels are able to operate a lower freight rates, thus leading to a reduction in the number of lay-ups. A further reason proposed is that Zannetos (1966) fails to allow for the fact that the costs of inputs into the tanker market may change over time, thus affecting the number of lay-ups.

This being said, Hawdon (1978) argued that there are more fundamental flaws in Zannetos (1966)'s model. The first of these is that, due to a lack of data, the data and methods used to construct $1/LU$ are open to question. The other is that $1/LU$ does not account for other factors, such as sailing speeds, which may affect the level of supply in the market. This leads one nicely onto Hawdon (1978)'s model, which is discussed in more detail in the following sub-section.

2.3.1.2 Hawdon (1978) and Tanker Freight Rates Across Time Horizons

Hawdon (1978) extended the previous literature by arguing that shipping freight rates in the tanker sector can be viewed as a series of interactions between the market for tanker services and the market for tankers. Hawdon proposed that both the supply and demand functions are price inelastic in the short-run. However, this picture changes in the long-run, where, although demand remains inelastic, the level of supply is likely to be affected by the current and expected values of freight rates in the market, where the market for tankers will initially alter as freight rates in the tanker market change.

Hawdon (1978) continued by proposing that as tankers predominantly carry oil, the demand for tanker services is likely to a function of the world demand for oil. In addition, Hawdon argued argues that supply of tankers in the market, or the size of the current fleet, is derived as follows:

$$F_t = F_{t-1} + D_t - Sc_t \qquad (2.9)$$

One should note that, in Expression (2.9) above, F_t and F_{t-1} denotes the size of the fleet in the current and previous periods, respectively, while D_t and Sc_t denote the number of vessels delivered and scrapped during the current period, respectively. In this function, Hawdon proposed that D_t and Sc_t are likely to be dependent on the expectations in the tanker market with respect to future rates, as well as some other market specific factors.

Hawdon (1978) hypothesised that the demand for tanker services is a simple function of total world trade in oil, where at any moment in time a certain proportion of the existing tanker fleet will be employed, or active, while the remainder will either be in lay-up or lying idle. One can measure the size of this active fleet as follows:

$$AF = a + bT \qquad (2.10)$$

In Expression (2.10) above, AF denotes the size of the active fleet, while T denotes the level of demand for tanker services. Consequently, this would imply that freight rates will remain relatively low, until the point at which the proportion of the total that fleet that is active reaches a critical value, at which point freight rates will subsequently increase rapidly as this proportion increases. This being said, Hawdon argued that freight rates in the tanker market are likely to vary with the proportion of the total fleet that is active, rather than with the size of the active fleet, hence:

$$R_t = f\left(a/F_t + bT/F_t\right) \qquad (2.11)$$

In Expression (2.11), R_t denotes the spot freight rate, a and b are the coefficients from Expression (2.10), F_t is as calculated in Expression (2.9) and T is given. Hawdon believed that if this analysis is correct, then a direct exponential relationship may exist between R_t, a/F_t and bT/F_t, since the proportion of the total fleet that is active is equivalent to the total fleet minus the number of lay-ups. Hawdon noted, however, that this model of freight rates is not sufficient as, in order to complete this specification one must take into account such supply-side factors. These include bunker costs (BC), the price of new tankers (PS), the cost of labour (Lab), the prevailing freight rate in the dry-bulk sector (DR), which acts as a substitute market, as well as the average size of a tanker (AS). Hawdon thus derived the following reduced form freight function:

$$\log R_t = \alpha + a/F_t + bT/F_t + cBC + dPS + eLab + fDR + gAS \qquad (2.12)$$

Hawdon went on to describe other characteristics of the tanker market, namely how orders for new tankers, second-hand ship prices, the level of scrapping and the number of new deliveries are described. However, as these are not strictly relevant for the current research topic, it would not be necessary for these to be considered further in this thesis.

2.3.1.3 Beenstock and Vergottis (1989) and the Dry-Bulk Market

Beenstock and Vergottis (1989) extended the previous empirical analyses by, amongst others Koopmans (1939), Hawdon (1978) and Wergeland (1981). This was done by developing and estimating, using annual data for the period between 1960 and 1985, an aggregate econometric model of the dry-bulk sector, in which freight rates, the level of lay-up, new and second-hand prices of vessels, and the size of the fleet were jointly determined. One should note that in this paper, Beenstock and Vergottis assumed rational expectations, i.e. that the best forecast of future values is provided by a random walk, and that time-charter rates are hypothesised to reflect rational expectations of current freight rates in the spot or voyage market. It should be further noted that Beenstock and Vergottis did not test this hypothesis, although later papers, such as Berg-Andreassen (1997), amongst others, did test this and found that it does not hold in the shipping context.

Beenstock and Vergottis (1989) assumed that ship-owners maximise profits under conditions of perfect competition, in which they are seen a price-takers, both in terms of voyage income, as reflected by the prevailing freight rates, denoted F, and voyage costs, denoted PB, such as bunkers, wages and port charges. Beenstock and Vergottis highlighted the fact that the number of voyages made by a vessels is directly proportionate to the average speed of the vessel, denoted S, however, voyage costs vary disproportionately with speed, in fact they increase exponentially. One can therefore determine the profit from a vessel, during any time period, as follows:

$$\Pi = SF - S^{\alpha}PB - OC \qquad (2.13)$$

In Expression (2.13) above, S^{α} reflects the hypothesis exponential relationship between the speed and voyage costs of a vessel, while OC denotes the operating, or fixed, costs for the respective vessel.

Beenstock and Vergottis (1989) highlighted the fact that freight rates are determined by the interaction between the demand and supply schedules in the market, where the supply of vessels is proportional to the size of the trading fleet multiplied by the average speed of this fleet. In addition, the size of the world fleet is inelastic in the short-run as a result of the lead-time on shipbuilding activities. One should also note that the number of vessels in lay-up will depend on the prevailing freight rates, voyage costs and running costs relative to the costs of lay-up. In the longer term, the fleet size varies as a result of shipbuilding and scrapping activity, where the level of shipbuilding varies directly with the price of new vessels, and scrapping varies inversely with second-hand prices of vessels. Beenstock and Vergottis therefore proposed that as the size of the fleet affects freight rates, while freight rates affect the stock demand vessels, the level of freight rates, ship prices and fleet sizes are dynamically interdependent.

Following this, Beenstock and Vergottis (1989) hypothesised that the short-term level of supply, which is measured in tonne-miles, can be determined as follows:

$$M^{S} = f_1\left(K^{*}; F/PB; Z_1\right) \qquad (2.14)$$

In Expression (2.14) above, M^S denotes the level of supply, K^* denotes the size of the fleet trading in the dry-cargo market, i.e. what Hawdon (1978) defined as the active fleet, F/PB denotes the size of the world fleet divided by the price of bunkers and Z_1 denotes a vector of exogenous variables that affect the level of supply.

Beenstock and Vergottis (1989) continued by defining the active fleet as follows:

$$K^* = (1-\mu)K + COM \qquad (2.15)$$

In Expression (2.15) above, μ denotes the proportion of the dry-bulk fleet in lay-up, COM denotes the number of combination carriers operating in the dry-bulk market, and K denotes the total dry-bulk fleet.

Beenstock and Vergottis (1989) proposed that the market behaves slightly differently in the longer-term as new vessels are delivered to the market and ship-owners are able to scrap their vessels. This would imply that the size of the dry-bulk fleet at the end of period *t* is calculated as follows:

$$K_t = K_{t-1} + D_t - S_t - L_t \qquad (2.16)$$

In Expression (2.16) above, D, S and L denote the number of new vessels delivered, the number of vessels scrapped, and the number of vessels lost, respectively. Beenstock and Vergottis proposed that the number of vessels sent for scrapping will vary inversely with the second-price of vessels relative to the price of scrap, and directly with the age profile of the fleet. In addition, drawing on previous models proposed, such as those in Witte (1963), Beenstock and Vergottis argued that the supply of new vessels, measured by the size of the order-book, varies directly with the new-building price and inversely with the price of other types of vessels, where the latter provide alternative means of income for the shipyard. Beenstock and Vergottis also highlighted the decision to invest or disinvest in vessels is also a function of the expected operating profit, which, as defined in Expression (2.13) above, is at least in part a function of the expected freight rates.

On the other side of the market equilibrium equation, the level of demand, denoted M^D, which is again expressed in tonne-miles, inevitably reflects the volume of seaborne trade. The volume of seaborne trade, in turn, reflects the level and structure of world economic activity, both in terms of geographic location and the type of commodities traded. Beenstock and Vergottis therefore simply defined the demand function as follows:

$$M^{D} = \tilde{M} \qquad (2.17)$$

One should that in Expression (2.17) above, $\tilde{M}$ denotes the exogenously determined volume of seaborne trade in dry-bulk commodities. Beenstock and Vergottis argued that, although in theory the level of demand should vary inversely with the level of freight rates, the fact that there is limited scope for substitution means that demand is relatively price inelastic. It should be further noted that Beenstock and Vergottis assumed that the market was in equilibrium, i.e. $M^{S} = M^{D}$, hence freight rates would move to clear the market. This would imply that periods of high (low) freight rates would induce ship-owners to invest (disinvest) in vessels, thereby leading to an increase (decrease) in the level of supply of shipping services, and a subsequent fall (rise) in freight rates, and hence follow a mean reversion process.

A flaw with Beenstock and Vergottis (1989)'s argument that freight rates follow a mean reversion process is that they do not examine the speed at which these adjustments occur. Although this research does not argue with the premise that freight rates are bounded in the long-run, it does feel that this will not necessarily be the case in the short-run due to the lag between the ordering and delivering of new vessels. Therefore, it may be more appropriate to talk about freight rates being fractionally integrated rather than completely stationary.

2.3.1.4 Concluding Comments on Structural Models

The structural models discussed above assume that freight rates are bounded by supply-side factors. One could therefore, using more recent time series theory, conclude that freight rates are mean reverting in the short-run and are therefore stationary. The reasons for this is that any dramatic increase in freight rates will be accompanied by an increase in the supply of shipping services as ship-owners take their tonnage out of lay-up and increase the speed at which their vessels are operating. This is because ship-owners do not want to miss the opportunity to earn high returns. The reverse will apply if freight rates fall, i.e. ship-owners will begin to lay-up their vessels and operate their vessels at a greatly reduced freight rate in order to minimise costs such as bunkers and the cost of wages. On the demand side, if freight rates increase dramatically, charterers will begin to look at other substitute methods of transport, such as air, therefore causing a decrease in the demand for shipping services. Therefore these factors all combine to create a mean reverting process, as outlined in later papers, such as Tvedt (1997) and Tvedt (1998).

2.3.2 Critique of General Equilibrium Models

While this research does not disagree with the argument that freight rates follow a mean reversion process, presented above, it proposes that the mean reversion process is 'delayed' in that freight rates display a long-memory and therefore this process will not happen as quickly as authors such as Zannetos (1966) and Hawdon (1978) may have implied. This would mean that the level of freight rates may exceed the two standard deviation unit band that a stationary process suggests. Hence, this research surmises that the answer to the question of the degree of integration lies somewhere in the middle of stationary and non-stationary process, i.e. freight rates are fractionally integrated $\left(0 < d < 1\right)$.

Another criticism of structural models is that they require one to forecast both the demand and supply sides of the freight rate process. While the supply side of the process is endogenous to the shipping market, and therefore may be able to be accurately forecasted using variables within the shipping market, the demand side factors are exogenous, in that the demand for shipping services is derived from world trade. As a result of this, one cannot forecast the demand side of the market using endogenous variables, and, even if one were to able to obtain all the relevant exogenous variables, many have tried to model world trade and failed. This research avoids this problem by using a partial equilibrium approach where the only item being used to model freight rates are freight rates themselves.

A final criticism of these models is that these all assume investors have 'rational assumptions'. This means that future freight rates will be solely be a function of the level of freight rates today. This hypothesis has, however, been queried in a number of papers, where, for instance, Berg-Andreassen (1997) found that this does not hold in the shipping context.

2.3.3 Arguments Regarding Non-Stationarity and Cointegration

This section reviews the literature on the cointegration between freight rates and other factors in the shipping market. The aim of this is to use the concept of cointegration, which requires that the underlying series are both integrated of the same order, where the order is greater than or equal to one, to determine whether freight rates follow a non-stationary process. To put this simply, in order to be cointegrated with another factor from the shipping markets, freight rates must follow at least an $I(1)$ process.

2.3.3.1 Berg-Andreassen (1996) and the Structure of Freight Rates

In the first paper reviewed here, Berg-Andreassen (1996) tested two fundamental hypotheses regarding the structure of freight rates in the shipping market. First, Berg-Andreassen tested if freight rates are stationary or not; and, second, then tested whether the assumption that freight rates are normally distributed holds true. The motivation behind the first hypothesis was that standard econometric techniques have serious shortcomings when performed on non-stationary variables, hence one needs to ensure that all variables, in this case freight rates, are stationary before continuing. As far as the second hypothesis is concerned, Berg-Andreassen argued that one cannot implement any risk-reducing diversification strategies if the characteristics of the distribution of the underlying series are not known.

Regarding the first hypothesis, Berg-Andreassen (1996) noted that Engle and Granger (1987) defined a series as integrated of order d if the d^{th}-difference of the series is stationary. To this end, Berg-Andreassen performed Augmented Dickey-Fuller tests, developed by Dickey and Fuller (1981), on a sample of daily freight rates, extending from 4 April 1985 and 23 December 1988 and across 13 routes. The results indicated that one could not reject the null hypothesis for the spot freight rate levels; however, one could for the first differences. Berg-Andreassen thus concluded that freight rates followed an $I(1)$ process, and were thus non-stationary in levels. Berg-Andreassen argued further that these results indicated that freight rates followed a random walk

process, hence forecasts of future freight rates based on historical statistics became nigh on impossible and therefore the conventional “last done” wisdom would provide just as good a forecast as any other more sophisticated method.
Looking at the second hypothesis regarding the distribution of freight rates, Mandelbrot (1963), in response to the common occurrence of extreme leptokurtosis in financial time series, developed a new family of distributions named the Paretian distribution. Berg-Andreassen (1996) adopted this as a starting assumption, which was then tested using the Jarque-Bera test, developed by Jarque and Bera (1980), the results for which indicated that one could reject the null hypothesis of normality, where further descriptive statistics indicated the presence of extreme leptokurtosis, as suggested by the Paretian assumption, as well high coefficients of skewness.

Berg-Andreassen (1996) thus reached the overall conclusion that freight rates over the sample period were non-stationary, following an $I(1)$ process. In addition, it was argued that the hypothesis of a Paretian distribution of these freight rates was met. This latter conclusion is of definite interest to the discussion of the structure of the higher moments later in this chapter.

This being said, there are doubts regarding the efficiency of the tests used to determine whether the data series are stationary or not. If the data series are not non-stationary as proposed by Berg-Andreassen (1996), then the results of the cointegration analysis will be invalid. In fact, traditional maritime economic theory, as outlined by Zannetos (1966), amongst others, argues that due to the supply and demand fundamentals, freight rates are in mean reverting, while more recent studies, such as Adland and Strandenes (2004), propose that freight rates are indeed stationary. However, this research argues that due to the lag between the ordering and delivery of new vessels, freight rates are in fact fractionally integrated.

2.3.3.2 Berg-Andreassen (1997) and Freight Rate Generation

Berg-Andreassen (1997) extended the earlier work on the structure of freight rates by empirically evaluating five different prevailing theories on the freight rate generation process in the time charter markets in shipping. Berg-Andreassen sought to test the market notion that changes in spot freight rates formed the basis of the market expectation regarding future period freight rates against those hypotheses laid out regarding the structure of period freight rates in papers by Zannetos (1966), Beenstock and Vergottis (1989) and Hale and Vanags (1989). In brief, Zannetos argued that the short-run rate is an important determinant in the long-run expectation process, while Hale and Vanags based their reasoning on the expectations hypothesis, i.e. that the spot-period freight rate relationships should be similar in nature to those in other markets. Beenstock and Vergottis argued that the current period rate should be a function of both the expected short-term rates and the voyage cost. Using the Koyck-lag estimation procedure, developed by Koyck (1954), it was shown that most of the explanatory power for the model was gleaned from the constant and lagged dependent variable, where this was the period rate.

Berg-Andreassen (1997) then formalised these concepts into five separate hypotheses. The first of these was the Zannetos Hypothesis, derived from the Zannetos (1966) model, which argued that the period was a function of both the spot rate levels and the changes in the spot rate, consequently:

$$Y_{i;t} = f\left(S_{i;t}; \Delta S_{i;t}\right) \qquad (2.18)$$

In contrast, the second, i.e. the Lagged Zannetos Hypothesis, extended the above hypothesis by postulating that it would hold if the explanatory variables were lagged one period, hence:

$$Y_{i;t} = f\left(S_{i;t-1}; \Delta S_{i;t-1}\right) \qquad (2.19)$$

The third hypothesis, known as the Koyck-Lag Hypothesis, which was derived from Beenstock and Vergottis (1989), stated that the period rate was a function of all previous period's voyage costs. In simpler terms it merely stated that the time charter rate was a function of the lagged time charter rate and the contemporary voyage costs, thus it can be expressed as follows:

$$R_{i;t} = f\left(R_{i;t-1}; V_{i;t}\right) \qquad (2.20)$$

Fourth, the Rational Expectation Hypothesis, as outlined in Hale and Vanags (1989), proposed that period freight rates were a function of the difference between the long- and short-term freight rates, the rate spread, and the lagged level of the short-term rate, therefore one could state that:

$$R_{i;t} = f\left(R_{i;t-1}; -Z_{i;t-1}; Z_{i;t-1}\right) \qquad (2.21)$$

The fifth, and final, hypothesis, referred to by Berg-Andreassen as the Conventional Wisdom Hypothesis, stated the market notion that the period freight rate was a function of only the changes in the short term freight rate:, i.e.:

$$Y_{i;t} = f\left(\Delta S_{i;t}\right) \quad (2.22)$$

One should note that in Expressions (2.18) to (2.22), $R_{i;t}$ and $Y_{i;t}$ denote the time charter freight rate in the market *i* at period *t*, where these were measured in $ / day and $ / dwt / month, respectively. Additionally, $S_{i;t}$ and $Z_{i;t}$ denote the spot freight rate in the market *i* at period t, where these were measured in $ / ton and $ / day, respectively; while, $V_{i;t}$ denotes the voyage costs in the market *i* at period *t*, where these were measured in $ / day.

Having outlined the hypotheses, Berg-Andreassen (1997) then moved on to examine the methodology to be used in the empirical evaluation of these. The first step in the process was to determine the degree of integration of the underlying freight rate series, using Dickey and Fuller (1981)'s Augmented Dickey-Fuller test for stationarity. Having done that, Berg-Andreassen then tested for cointegration between the underlying series using the test developed by Johansen (1988). Using a data set comprised of 10 years of quarterly data over three different routes, Berg-Andreassen reached the conclusion that all variables in the data set were first-difference stationary. Berg-Andreassen found that the residuals for all series were $I(1)$, therefore the Zannetos and Lagged Zannetos hypotheses could be rejected outright. Moving on, the results for the Johansen test indicated no cointegrating relationship between the time charter rates and voyage costs, therefore the Koyck Lag hypothesis could be rejected, while mixed results for the Rational Expectations Hypothesis led Berg-Andreassen to conclude that the Johansen test was more reliable and therefore reject this hypothesis. Of all the hypotheses, the only set of series for which there was a cointegrating relationship was between the time charter rates and changes in the spot freight rates. Consequently Berg-Andreassen concluded that only the Conventional Wisdom Hypothesis is valid in these cases.

One should note, however, that a series of major problems with this paper exist. The first of these is that the whole analysis is only based on three routes, which places serious doubts on whether or not these results are generalisable across the entire freight market. Furthermore, as argued previously in this chapter, there are doubts regarding the efficiency of the tests used to determine whether the data series are stationary or not. If Berg-Andreassen (1997)'s conclusion that the freight rates are non-stationary fails, then the results of the cointegration analysis will be invalid. In fact, traditional maritime economic theory, as outlined by Zannetos (1966), amongst others, argues that due to the supply and demand fundamentals, freight rates are in fact mean reverting, a fact supported by more recent literature, such as Tvedt (1998), where these suggest that freight rates follow a stationary process. This being said, this research argues that due to the lag between the ordering and delivery of new vessels, freight rates are in fact fractionally integrated.

2.3.3.3 Veenstra and Franses and the Efficiency of Shipping Markets

Veenstra and Franses (1997) took a slightly different approach to previous papers in that they looked at cointegration between data series, however these data series were solely comprised of freight rates. Veenstra and Franses developed a Vector Autoregressive (VAR) model to model these freight rates and then assessed forecasts derived from this model to determine whether this improved the accuracy of short- and long-term forecasts. The main justification provided for this process was that shipping freight markets were assumed to be approximately efficient, hence these freight rates would contain all publicly available information and no extra variables beyond these would be required for model building. Furthermore, Veenstra and Franses argued that one might have expected that freight rates for different parts of the shipping industry were correlated; therefore one could try to indentify the underlying structure that could have been summarised in a multivariate time series model, such as that provided. Veenstra and Franses therefore identified three main aims for the research, i.e. to generate forecasts; to identify the long-term trend behind the freight rates; and, to investigate whether the VAR model outperformed the no-change forecasting model. Veenstra and Franses suggested that should the above aims not be the case, then one could feel confident regarding the validity of the Efficient Markets Hypothesis in the case of the shipping markets.

In the analysis, Veenstra and Franses (1997) used a data set comprised of the natural logarithms of monthly freight rates across six separate routes and two vessel classes, three per class, for the period between September 1983 and February 1995, where the period between September 1993 and February 1995 was used for ex-ante evaluation of the forecasting performance. Veenstra and Franses then performed Augmented Dickey-Fuller tests, first proposed by Dickey and Fuller (1981), to test for the stationarity of the data series, the results for which indicated that all data series were non-stationary at the 1% level of significance. Having calculated the respective correlations for all data series, Veenstra and Franses argued that the correlation and unit root results all provided evidence that the series could have common properties. Veenstra and Franses proposed the use of the VAR model in order to describe these properties, where this would have explained the characteristics of the underlying six freight rate series through their own lagged values, where this was somehow restricted to reflect the common features. In order to test the assumption that the series were cointegrated, Veenstra and Franses performed unit root tests on the un-weighted differences between all possible combinations of two out of the six series. As an alternative to this pair-wise approach, Veenstra and Franses investigated cointegration amongst all six variables in one step using use the Johansen (1991) test for this purpose. Veenstra and Franses noted that Granger Representation Theorem, outlined in Engle and Granger (1987), suggests that should a cointegration relation between a set of series exist, then the VAR can be written as a Vector Error Correction (VEC) model, and vice versa. Based on the results of these tests, Veenstra and Franses thus concluded that there were five cointegrating relationships within the sample set and then obtained the estimated cointegration parameters within the VEC model using the Ordinary Least Squares methodology.

The results from this analysis indicated that only four adjustment coefficients were significantly different from zero, where, in the subsequent forecasting exercise, all other parameter were set equal to zero. Veenstra and Franses (1997) highlighted the fact that a common phenomenon when modelling ocean freight rates is that different freight rates exhibit quite similar patterns, where this similarity may indicate the existence of a common stochastic trend within the data set. To this end, Johansen (1991) established that a condition for one common trend to exist for six freight rates would be that there are five cointegration relations in the set of freight rates. Veenstra and Franses therefore once again considered the VEC specification of their VAR (1) model in order to find an expression for the stochastic trend, and to discover a way of calculating that trend. Veenstra and Franses noted that the stochastic trend could not be forecasted as it was, by definition, a random walk process, and therefore simply portrayed the unexplained part of the data once any deterministic or common relationships within the provided data set had been removed.

Veenstra and Franses (1997) then moved on to look at the forecasting performance of the specified VAR model, where the forecasting equation was derived on the basis of the general VAR (1) model used above. Veenstra and Franses confined their forecast analysis to dynamic forecasts from 1 up to 18 periods ahead and forecasts for 1 period ahead for the each of the 18 periods and then compared these forecasts with the naive forecasting method, where the observation in the previous period was used to provide the forecast for this method. Veenstra and Franses found that the forecasts did not appear to pick up the actual freight rate movements; however, they observed that the realized observations were usually well within the 95% forecasting intervals. Veenstra and Franses noted that these results may be as a result of the fact that longer term forecasts for VAR models tend toward the estimated average of the series in the model. Another reason why the model may not have performed well in the long term was illustrated by the stochastic trend described above. Veenstra and Franses therefore concluded that their proposed multivariate VEC model with 5 cointegration relations was defeated by a naive forecasting strategy for both the short- and long-term forecasts of these series.

Veenstra and Franses (1997) thus reached the conclusion that an economically meaningful structure exists in a set of ocean dry bulk freight rates. Further, the results did not seem to be in conflict with the efficient market hypothesis as it applies to ocean freight rates as, even though there appear to be long-run relationships between freight rates, they found that such relationships do not result in improved forecasts.

As with the previous papers, there are doubts regarding the efficiency of the tests used to determine whether the data series are stationary. If the data series are not non-stationary, then the results of the cointegration analysis will be spurious. In fact traditional maritime economic theory, such as Zannetos (1966), amongst others, argues that due to supply and demand fundamentals, freight rates are in fact mean reverting, with more recent partial equilibrium studies, such as Adland and Cullinane (2006), adding that freight rates are stationary. This being said, this research argues that due to the lag between the ordering and delivery of new vessels, freight rates are in fact fractionally integrated.

2.3.3.4 Kavussanos and Nomikos (1999) and the Issue of Unbiasedness

The focus changes somewhat now to examine the interaction between futures and spot prices, where Kavussanos and Nomikos (1999) investigated the Unbiased Expectations Hypothesis of futures prices in the freight future markets. Cointegration techniques, which were employed in order to examine this hypothesis, illustrated that futures prices provided unbiased forecasts of realised spot prices superior to those generated from Error Correction, Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing and Random Walk models. Hence, Kavussanos and Nomikos argued that it appeared that users of the freight futures market, i.e. the Baltic International Freight Futures Exchange (BIFFEX) market, received accurate signals from these markets as to the future direction of spot freight rates and could therefore utilise the information generated by these freight futures prices to guide their decisions in the physical market.

Kavussanos and Nomikos (1999) argued that price discovery is one of the main reasons for the extent to which future contract prices reflect unbiased expectations of the spot price on the date of delivery is important to market participants. One should note that by price discovery, the authors meant the process through which market participants are able to use futures prices to ‘discover’ future equilibrium prices in spot markets. If futures prices are not unbiased forecasts of these equilibrium prices, then they may not perform this price discovery role efficiently as they do not represent accurate predictors of expected spot freight rates. Several studies, including Lai and Lai (1991) in various FOREX markets, Chowdhury (1991) in commodity markets, and Crowder and Hamed (1993) in the oil futures market, found that the unbiased expectations hypothesis held in that futures prices were unbiased forecasts of the realised spot prices. On the other hand, Krehbiel and Adkins (1994) found that the Unbiased Expectations Hypothesis failed in the quarterly Treasury bill, Eurodollar and Treasury bond futures markets. One should note that one of the common features of these studies was their use of cointegration techniques due to the fact that spot and futures price series followed a non-stationary process. According to Kavussanos and Nomikos, the Unbiased Expectations Hypothesis is comprised of two suppositions, namely, that the price of a freight futures contract before the date of maturity is equal to the expected spot freight rate on the date of maturity; and, that the expectation of the spot freight rate is formed rationally. The authors examined this notion empirically by testing the parameter restrictions that $(\beta_1;\beta_2)=(0;1)$ in the following expression:

$$S_t=\beta_1+\beta_2F_{t;t-n}+\varepsilon_t \;\; ; \;\; \varepsilon_t \sim iid\left(0;\sigma^2\right) \qquad (2.23)$$

If futures were an unbiased predictor of the future spot price, then the current futures price would contain all the relevant information required to forecast the next period's spot price, where these were denoted $F_{t;t-n}$ and S_t, respectively, in Expression (2.23) above. In addition, investigated the short run dynamic properties of spot and futures prices with the aim of identifying the speed with which spot and futures prices responded to deviations from their long-run relationship.

Kavussanos and Nomikos (1999) then tested for cointegration using Johansen's estimation procedure, outlined in Johansen (1988), where under this specification, the joint distribution of spot and futures prices can be described as a Vector Error Correction Model (VECM). One should note that when spot and futures prices follow a non-stationary process, cointegration is a necessary condition for the Unbiased Expectations Hypothesis to hold. If this is not the case, then spot and futures prices will tend to drift apart over time and therefore futures prices cannot be unbiased predictors of the realised spot prices. It is important to realise that although cointegration is a necessary condition for the Unbiased Expectations Hypothesis, it is not a sufficient condition, i.e. it does not necessarily mean that the hypothesis holds, as demonstrated by Hakkio and Rush (1989). Before one can test for cointegration, one needs to establish that the component data series follow a non-stationary process both in terms of seasonal and ordinary unit roots. In order to test this, Kavussanos and Nomikos employed the methodology outlined by Hylleberg, *et al.* (1990) to test for seasonal unit roots, and the Augmented Dickey-Fuller and Phillips-Perron tests to test for ordinary unit roots, where the these were developed by Dickey and Fuller (1981) and Phillips and Perron (1988), respectively. The results of these tests indicated that the underlying series exhibit no seasonal unit roots, however, all series were found to be first-difference stationary. Once this was identified, cointegration techniques were used by Kavussanos and Nomikos to test the Unbiased Expectations Hypothesis, where three steps may be distinguished in this process. First, the authors arrived at a well-specified VECM; second, the existence of a cointegrating vector was investigated within this VECM using the maximum and trace tests proposed by Johansen (1988); and finally, once the existence of the cointegrating relationship had been established, the Unbiased Expectations Hypothesis was investigated by testing the parameter restrictions that $\beta_1 = 0$ and $\beta_2 = 1$ in the cointegrating relationship using the Likelihood Ratio statistic outlined in Johansen and Juselius (1990). If these restrictions hold, then the price of the futures contract is an unbiased predictor of the realised spot price.

The results obtained by Kavussanos and Nomikos (1999) indicated a VECM model with lags of 1, 2 and 3 for the one-, two- and three-months data, respectively, was well-specified and that a cointegrating relationship did exist between spot and futures prices. In the case of the one- and two-months futures prices, Kavussanos and Nomikos found that the null hypothesis that the Unbiased Expectations Hypothesis held could not be rejected, hence futures prices one and two months prior to maturity were unbiased predictors of the realised spot prices. However, in the case of the quarterly futures prices, the restriction was rejected and therefore futures prices three months prior to maturity provided biased forecasts of the realised spot prices. Possible reasons for this included thin trading, as proposed by Gilbert (1986), or that this bias reflected imbalances between long and short hedging demand in the market, particularly for non-storable commodities, as proposed by Kolb (1992) and Deaves and Krinsky (1995).

Kavussanos and Nomikos (1999) then tested the forecasting performance of futures prices to test whether the findings of Ma (1989), Kumar (1991) and Hafer, *et al.* (1992), who found that, broadly speaking, futures prices provide superior forecasts of the realised spot prices than do forecasts from alternative models, hold in the shipping market. Kavussanos and Nomikos (1999) compared futures price forecasts with those generated by bi-variate VECM, univariate ARIMA and Holt-Winters Exponential Smoothing, as outlined in Holt (1957) and Winters (1960), models. The results obtained indicated that for the two- and three-months horizons, futures prices outperformed the other models considered; however, for the one-month horizon, the VECM provided marginally better forecasts than the futures prices. Kavussanos and Nomikos (1999) concluded that the results obtained had two major implications, the first of which was that participants in the futures market received accurate signals from futures prices, which could then be used to guide their physical market decisions. The second implication was that for the one- and two-month periods, the "average" hedger could use this market to efficiently forecast realised spot rates without having to paying any form of risk premium.

Once again there are doubts regarding the efficiency of the tests used to determine whether the data series are stationary or not. If the data series are not non-stationary as proposed by Kavussanos and Nomikos, then the results of the cointegration analysis will be spurious. In fact, traditional maritime economic theory, such as was outlined in Zannetos (1966), amongst others, argues that due to supply and demand fundamentals, freight rates follow a mean reversion process, while more recent literature, such as Koekebakker, *et al.* (2006), suggests that freight rates follow a stationary process. This being said, this research argues that due to the lag between the ordering and delivery of new vessels, freight rates instead follow a fractionally integrated process.

2.3.3.5 Kavussanos and Nomikos (2003) and Granger Causality

Kavussanos and Nomikos (2003) extended their earlier work by, in addition to investigating the price discovery relationship in the freight futures market, as in the earlier work, investigating the causal relationship between spot and futures prices in the Baltic International Freight Futures Exchange (BIFFEX) market. The authors argued that futures prices must lead the underlying spot prices in order to fulfil their price discovery role. This was illustrated in papers by Stoll and Whaley (1990), Wahab and Lashgari (1993), Hung and Zhang (1995) and Tse (1995), amongst others, where the overall conclusion was that causality between spot and futures prices can run in one or both directions, where, in all cases, futures prices contribute to the discovery of new information regarding the future level of spot prices in the underlying market.

In order to establish these relationships, Kavussanos and Nomikos (2003) used the VECM, first proposed by Johansen (1988), and the Johansen statistics, developed by Johansen (1991), using the methodology outlined in Granger (1986). Before being able to establish cointegration between variables, one must first establish the degree of integration for each of the underlying variables, hence the authors used Augmented Dickey-Fuller, Phillips-Perron and KPSS tests, attributed to Dickey and Fuller (1981), Phillips and Perron (1988) and Kwiatkowski, *et al.* (1992), respectively, to test whether spot and futures prices were non-stationary. The results of these tests indicated that both spot and futures prices followed an $I(1)$ process, that spot and futures prices were cointegrated and that futures prices tended to discover new information more rapidly than spot prices. Following this, Kavussanos and Nomikos compared the forecasting performance of the resultant Vector Error Correction Model (VECM) with that of Vector Autoregressive (VAR), Autoregressive Integrated Moving Average (ARIMA) and Random Walk models, where the accuracy of these forecasts was assessed using the Diebold-Mariano test, first outlined by Diebold and Mariano (1995). The results of these assessments indicated that, as long as futures prices were formulated as a VECM, these provided a more accurate forecast of the realised spot prices than the other models, thereby confirming the results of their earlier paper, i.e. Kavussanos and Nomikos (1999).

As with the previous papers, there are doubts regarding the efficiency of the Augmented Dickey-Fuller test, where Harris (1995) and Maddala and Kim (1998), amongst others have criticised this test on the basis that they are not powerful enough in rejecting the null hypothesis of a unit root, particularly in cases where there is mean reversion which is long relative to the sample length. Traditional maritime economic theory, such as outlined in Beenstock and Vergottis (1989), amongst others, argues that this is indeed the case in the shipping markets. This being said, this research addresses this issue by arguing that freight rates are fractionally integrated.

2.3.3.6 Kavussanos and Visvikis (2004) and the FFA Market

Another paper that looked into the forecasting power of futures prices with respect to spot prices was that by Kavussanos and Visvikis (2004), however, in this case, unlike previous studies, this was in the Forward Freight Agreement (FFA) market. A further difference is that Kavussanos and Visvikis extended the examination by investigating the volatility of the series, where the volatility component shall be examined later in this literature review. The ultimate aim of this research was to investigate the lead-lag relationship, in terms of both returns and volatility, between the spot and futures markets in the shipping industry.

In terms of the returns, Kavussanos and Visvikis (2004) began by determining the order of integration of the underlying spot and futures prices using Augmented Dickey-Fuller, Phillips-Perron and KPSS tests, where these were developed by Dickey and Fuller (1981), Phillips and Perron (1988) and Kwiatkowski, *et al.* (1992), respectively. The results of these tests suggested that both the daily spot and FFA series were first-difference stationary. Having established this, Kavussanos and Visvikis then tested for cointegration between the data series using the Johansen procedure, developed by Johansen (1988), the results for which indicated that spot and FFA prices were integrated across all routes. According to Granger (1988), should series be cointegrated, causality between the series should exist in at least one direction, therefore the authors implemented the Granger Causality test in order to establish the direction of this causality finding be-directional relationships between spot and FFA prices, however, FFAs played a leading role in incorporating new information. Kavussanos and Visvikis therefore concluded that FFA prices played a crucial price discovery role in the shipping markets.

Once again the lack of power with regard to the tests for stationarity causes concern here. If the data series are not non-stationary, then the results of the cointegration analysis will be spurious. In fact traditional maritime economic theory, such as Zannetos (1966), amongst others, argues that due to supply and demand fundamentals, freight rates are in fact mean reverting, where more recent literature, such as Adland and Cullinane (2006), suggests that freight rates follow a stationary process. In contrast, this research argues that due to the lag between the ordering and delivery of new vessels, freight rates are in fact fractionally integrated.

2.3.3.7 Other Papers on Non-Stationarity and Cointegration

In addition to the research discussed above, other research on this topic, amongst others, included significant papers by Kavussanos (1996), Glen and Rogers (1997), Haigh (2000), Kavussanos and Nomikos (2000b) and Haigh and Holt (2002). Kavussanos (1996), while investigating volatility in the spot and time-charter markets for dry-bulk vessels, also tested for stationarity in the freight rate series examined. Results from the Dickey-Fuller and Augmented Dickey-Fuller tests, attributed to Dickey and Fuller (1979) and Dickey and Fuller (1981), respectively, indicated that the logarithms of the spot and time-charter freight rates followed an $I(1)$ process. Furthermore, Glen and Rogers (1997), when testing for cointegration between the component freight rates of the SSY Capesize Index, and Kavussanos and Nomikos (2000b), when investigating the relationship between spot and futures prices, used Augmented Dickey-Fuller and Phillips-Perron tests, where the latter was first developed by Phillips and Perron (1988), to test the stationarity of the underlying time series. The results of these tests suggested that all component freight rate series were first-difference stationary. Providing yet more support of this phenomenon, Haigh (2000) and Haigh and Holt (2002) both, when also testing for cointegration between spot and futures prices in the freight market, found that the results of the Augmented Dickey-Fuller test showed that both the underlying spot and futures freight rates series were integrated of order one.

Once again there are doubts regarding the efficiency of the tests used to determine whether the data series are stationary or not. If the data series are not non-stationary as proposed by these papers, then the results of the cointegration analysis will be spurious. In support of this criticism, traditional maritime economic theory, such as outlined in Beenstock and Vergottis (1989), amongst others, argues that due to supply and demand fundamentals, freight rates are in fact mean reverting; where this research extends this argument by proposing that due to the lag between the ordering and delivery of new vessels, freight rates are in fact fractionally integrated.

2.3.3.8 Concluding Comments

These papers all had one thing in common, i.e. that they found that freight rates were non-stationary, which therefore enabled the respective authors to use cointegration to try and determine the relationship between freight rates and some other variable. This enabled the authors to better understand the behaviour of freight rates, and improve on their ability to model this stochastic variable. The major assumption here, as stated previously, was that freight rates followed an $I(1)$ process. This means that freight rates are no longer considered to be mean reverting, a result completely contrasting traditional maritime theory, as outlined in, for example Zannetos (1966), and freight rates would no longer be assumed to be constrained within a upper and lower bound. As a result of this, interesting new models, such as ARIMA and VECM, were introduced to try and increase the accuracy of freight rate forecasting, although the accuracy was still surprisingly low and many papers argued that the most accurate method of forecasting future spot freight rates was to look at the futures market.

2.3.4 Critique of Non-Stationarity and Cointegration

This research does not agree with this assumption of non-stationarity. As stated above, traditional maritime economic theory, such as outlined in Zannetos (1966) and Beenstock and Vergottis (1989), amongst others, states that supply and demand constraints are such that freight rates will illustrate mean reverting properties. The reason for this is that any dramatic increase in freight rates will be accompanied by an increase in the supply of shipping services as ship-owners take their tonnage out of lay-up and increase the speed at which their vessels are operating. This is because ship-owners do not want to miss out on the chance to earn high returns. The reverse will apply if freight rates fall, i.e. ship-owners will begin to lay-up their vessels and operate their vessels at a greatly reduced freight rate in order to minimise costs such and bunkers and the cost of wages. On the demand side of things, if freight rates increase dramatically, charterers will begin to look at other substitute methods of transport, such as air, therefore causing a decrease in the demand for shipping services. Therefore these factors all combine to create a mean reverting process, as discussed in Tvedt (1997) and Tvedt (1998).

In addition to this, Schwert (1989) illustrated the low power of the Augmented Dickey-Fuller and Phillips-Perron tests, developed by Dickey and Fuller (1981) and Phillips and Perron (1988), respectively, particularly in the case of stationary processes where there is a large negative unit root in the moving average term. Furthermore, more recent papers using partial equilibrium models, such as those by Adland and Cullinane (2006) and Koekebakker, *et al.* (2006), have proposed that spot freight rates are stationary, at least in the tail of their distribution, however, these results are not conclusive.

This research aims to settle this dispute by proving that freight rates follow a mean reverting process but that the mean reversion process is 'delayed' in that freight rates display a long-memory and therefore this process will not happen as quickly as authors such as Tvedt (1998) and Koekebakker, *et al.* (2006) may have implied. Therefore, the level of freight rates may exceed the upper and lower constraints that a stationary process suggests. Hence, this research feels that the answer to the question of whether freight rates are stationary or not lies somewhere in the middle of stationary and non-stationary process, i.e. freight rates are fractionally integrated in that they are integrated of a fractional order between zero and one, i.e. $0 < d < 1$.

2.3.5 Stationarity and Partial Equilibrium Models

As discussed above, Schwert (1989)'s argument that the most commonly used tests for non-stationarity, i.e. the Augmented Dickey-Fuller and Phillips-Perron tests, attributed to Dickey and Fuller (1981) and Phillips and Perron (1988), respectively, lack accuracy casts doubt on many of the statements regarding the non-stationarity of freight rates. Additionally, traditional maritime economic theory, such as outlined in Zannetos (1966) and Beenstock and Vergottis (1989), suggests that freight rates are constrained by an upper and lower limit due to the dynamics of the supply and demand functions in the shipping markets. Recently, new literature has re-opened the debate regarding the structure of the freight rate process. This literature is examined in the discussion below.

2.3.5.1 Tvedt (1997) and the Geometric Mean Reversion Process

Tvedt (1997) examined the structure of freight rates as a geometric mean reversion process, which was then used to address the question of how to value a VLCC tanker. Tvedt argued that as the value of an asset is influenced by the future discretion of its owner to adjust its use or properties in reply to stochastic events, in this case changes in the underlying freight rates, freight rate uncertainty would have been of paramount importance to the market for VLCC vessels. Therefore, Tvedt believed that the improved valuation of these vessels was a sufficient motivation for a search to correctly describe the stochastic nature of the underlying freight rate. One should note that the demand function for shipping services is inelastic due to the high relative costs of substitute transportation. Furthermore, the supply function, in the short-run, will also be relatively inelastic when there are no idle vessels available, with the only increase in supply here coming from the increasing speed and efficiency of the vessels concerned. This inelasticity of supply is primarily due to the lag experienced, where this is generally longer than a year, between the ordering of new tonnage and its delivery. On the reverse side, if freight rates are exceptionally low, vessels may enter lay-up; resulting in a decrease in supply, however, should the underlying freight rates pick up, this short-run supply can be reintroduced to the market.

Tvedt (1997) argued that as a result of this inelasticity of demand and the short-term upper limit to supply, freight rates may increase to very high levels. This being said, these higher freight rates will act as a trigger for ship-owners to start ordering new vessels, however, as a result of the lag discussed above, there would be a delayed effect to the supply function. Consequently, Tvedt proposes that although these very high freight rates would only be a temporary occurrence, they may persist for some time. On the reverse side, if freight rates fell too low, then ship-owners may be forced to scrap their vessels, due to liquidity shortages, resulting in a decrease in the level of supply and a resultant increase in freight rates. Tvedt therefore argued that freight rates are bounded by an upper limit beyond which charterers may seek alternative transportation, and a lower limit beyond which ship-owners will cease trading.

Tvedt (1997) noted that, for the lack of a better model, some practitioners in the shipping industry used the Black-Scholes model, first proposed and used by Black and Scholes (1973), to approximate the value of simple options written on the underlying freight rates. Tvedt (1997) highlighted the fact that one of the fundamental assumptions for this Black and Scholes model is that the underlying prices follow a stochastic process described by a Geometric Brownian motion. This being said, Tvedt postulated that there was no reason to assume that freight rates in the shipping market should follow this Geometric Brownian motion. This argument was supported by Bjerksund and Ekern (1995) who suggested that spot freight rates followed an Ornstein-Uhlenbeck process and that cost were constant, where the reason for this assumption was that this process has mean reversion properties. Tvedt (1997) noted, however, that, as discussed above, should freight rates fall to levels where operational costs are not covered, ship-owners would either lay-up or scrap their vessels; another constraint here is that freight rates may not be negative. Tvedt therefore rejects the Ornstein-Uhlenbeck process as it fails to take account of this since the process is not downward restricted; as well as the fact that because the process is normally distributed around the given mean, it often gives negative values if volatility is high.

Following the limitations of the above processes, Tvedt (1997) therefore proposed that freight rates may have been best described by a Geometric Mean Reversion process. This process, like the Ornstein-Uhlenbeck process suggested by Bjerksund and Ekern (1995), has mean reversion properties, however, unlike the Ornstein-Uhlenbeck process, it also fulfils the requirement of being downward restricted in that zero is an absorbing level. Furthermore, the Geometric Mean Reversion process may have proved to be a reasonable approximation as to the fact that freight rates in the spot market often stayed at a moderate level, with relatively low volatility for long periods, followed by periods of high freight rates and volatility. This process secured that the mean reversion would be strong when freight rates were low, and vice versa; and, due to the geometric nature of the last term, the process also related high rates to high volatility, and vice versa.

Bearing the above in mind, Tvedt (1997) noted that although the Ornstein-Uhlenbeck and Geometric Mean Reversion processes may not have provided the best Markov specification of freight rates, these processes belong to a very small class of stochastic differential equations that are analytically solvable. Tvedt therefore argued that choosing the most appropriate specification of a series for valuation purposes will always involve a trade-off between analytical tractability and goodness of fit to the observations of the series in question.

The most important thing to note, with respect to this thesis, is that Tvedt (1997) assumed that freight rates followed a mean reversion pattern. When freight rates are high, even though demand is relatively inelastic, charterers will utilise substitute forms of transportation, while, when the reverse applies, ship-owners will scrap, or at the very least lay-up, their vessels.

Some important concerns regarding Tvedt (1997) are that this paper was purely theoretical and did not give any consideration to the empirical characteristics of the data concerned. Additionally, this paper used highly technical stochastic models, which are not easily implemented by laymen in the field. Finally, this paper only considered the valuation of one specific type of vessel, which raises queries as to whether results from these models would be generalisable or not. Therefore, this research feels that further research into this topic is required.

2.3.5.2 Tvedt (1998) and Valuation Assuming Stationary Freight Rates

In a later paper, Tvedt (1998) examined the structure of the underlying freight rates in the Baltic International Freight Futures Index (BIFFEX) derivatives market. Tvedt argued that while some practitioners used either the Black or Black-Scholes formulae developed by Black (1976) and Black (1976), respectively, to price these options. This being said, Tvedt (1998) argued that the behaviour of the underlying freight index, i.e. the Baltic Freight Index (BFI), most probably did not follow the Geometric Brownian motion assumed by these models, hence Tvedt proposed that practitioners should have considered other pricing models for the shipping models, where this was set out as the aim of Tvedt's paper.

In Chapter 3.3 of Gray (1990), Gray argues that the low of 553.5 experienced by the BFI in 1986 represented the lowest income level at which a ship-owner would have continued to operate in the market, before laying up their vessel. On the reverse side, Gray believed that, historically, there appeared to be an upper resistance level in the BFI of 1,650, beyond which increased supply capacity, due to more efficient vessel use and the entry of combination carriers from other sectors, has a dampening effect on freight rates. Therefore, taking into account these apparent upper and lower limits, it appeared that the underlying freight rates followed a mean reversion process. These market dynamics would have obviously affected the value of any derivatives in the market, hence Bjerksund and Ekern (1995) assumed that the underlying freight rates for the optioned being valued followed an Ornstein-Uhlenbeck process, i.e. that the freight rate process was normally distributed and that freight rates gradually reverted to a constant mean level after any shock.

Tvedt (1998) thus felt that the appropriate process for discussing the underlying freight rates of the BFI should have had mean reversion properties and been restricted downwards by the laying-up of vessels, as had been discussed in Mossin (1968) as well as Dixit and Pindyck (1994). Hence, Tvedt (1998) argued that the increment of the BFI was given by a mean reversion process with an absorbing level, i.e.:

$$dX_t = \kappa \times [\alpha - \ln(X_t - \lambda)] \times [X - \lambda] dt + \sigma [X_t - \lambda] dZ_t \qquad (2.24)$$

One should note that in Expression (2.24), X_t is the index value at time t; dZ_t is the increment of a standard Brownian motion; and κ, α and σ are constants. This process had mean reversion properties, thus, if $\ln\left(X_t - \lambda\right)$ was above α, the trend of the process would have been negative, and vice versa. Additionally, the mean reversion was stronger for high values of the BFI, as opposed to low values, for the same absolute deviation of $\ln\left(X_t - \lambda\right)$ from α, and the process exhibited increasing volatility as the index level rose.

Tvedt (1998) concluded that freight rates reverted downward if they were above average, and vice versa, where the mean reversion property was due to frictional capacity adjustments to changes in the demand for shipping services. This would thus have influenced the value of any derivatives on the underlying index via the variance of the futures price process.

Some important things to note about this paper are that this paper is purely theoretical and does not give any consideration to the empirical characteristics of the data concerned. In addition, this paper uses highly technical stochastic models, which are not easily implemented by laymen in the field. Finally, this paper considers the valuation a futures index, which although linked to the underlying freight rate process, leaves room for further research into the structure of that process itself.

2.3.5.3 Tvedt (2003) and the Dollar-Yen Effect on Stationarity

In yet another later paper, Tvedt (2003) proposes that freight rates and second-hand ship prices in the dry-bulk shipping market follow a stationary process when transformed from US dollars into Japanese yen. Tvedt proposes that the random walk hypothesis, i.e. that freight rates follow an $I(1)$ process, can be rejected in most cases. The results obtained by Tvedt confirm the classical maritime economic theory, which argues in favour of stationarity in freight rates.

Tinbergen (1931) introduced the concept that high freight rates trigger the ordering of new vessels, which in turn causes downward pressure on the prevailing freight rates. This being said, as a result of the lag in the delivery of these vessels, this fall in freight rates will not be instantaneous and high freight rates may prevail for a short while. In the reverse instance, excess capacity may cause freight rates to fall to lower levels, where the rate of reversion to the mean level will depend either on the speed at which demand catches up with the excess capacity and equilibrium is restored, or whether excess tonnage is destroyed through scrapping. This concept of the ‘mean reverting’ nature of freight rates suggests that freight rates, as a part of their stochastic nature, have a downward trend when freight rates are high, i.e. have an upper bound, and vice versa, as was discussed in Beenstock and Vergottis (1989) and Hawdon (1978). This assumption of mean reversion in freight rates has been used in theoretical asset valuation papers by Bjerksund and Ekern (1995) and Tvedt (1998), amongst others.

With very few exceptions, the evidence from the shipping literature, as presented by Berg-Andreassen (1996), Veenstra and Franses (1997) and Glen (1997), amongst others, tends towards the fact that freight rates in the dry-bulk shipping markets follow a random walk process. In contradistinction, as stated previously, the traditional maritime economic theory, as well as the earlier studies by Tvedt, i.e. Tvedt (1997) and Tvedt (1998), does support the mean reversion hypothesis.

Tvedt (2003) uses the Augmented Dickey-Fuller test, see Dickey and Fuller (1981), to test for unit roots in the data, and found that, when denominated in US dollars, the data series follow a random walk process, thus confirming the random walk hypothesis outlined in the earlier literature. However, when all observations are converted from US dollars to Japanese yen, the results change and dry-bulk freight rates appear to be stationary around a trend. Additionally, the random walk hypothesis was also rejected, at the 5% level of significance, in the case of the Baltic Freight Index, although in the cases of new-building and second-hand vessel prices, one cannot reject this hypothesis. The fact that one could not reject the random walk hypothesis in the case of the vessel prices may as a result of the fact that prices in the second-hand market may have been less influenced by the market fundamentals, where Tvedt (2003) argued that these were largely governed by the development of the Asian economies.

Tvedt (2003) therefore concludes that by changing the shipping market perspective from US dollars to Japanese yen, which Tvedt argues may be a more realistic approach, given the fact that the market is dominated by Far East players, freight rates in the dry-bulk market appear to be stationary. Tvedt also adds that although non-Asian ship-owners generally consider international shipping as a US dollar industry, this perception may be somewhat misleading, since yen dominated prices probably better reflect fundamental changes in the industry, thus giving better feedback to market agents. Furthermore, investing in a shipping asset probably means that the ship-owner will have to take a long-position on a yen-related asset, thus implying that there would be substantial exchange rate risk for a ship-owner wishing to maximise their US dollar fortune.

The first of the main concerns with this paper is that the unit root test implements has been shown to be deficient, as outlined in Schwert (1989). In addition, the fact that this paper converts freight rates from a US dollar to a Japanese yen denomination is highly irregular. The convention in international shipping is that all international freight rates are denominated in US dollars, where, for this denomination, the results do not vary from the previous literature supporting the random walk hypothesis.

2.3.5.4 Adland and Strandenes (2004) and Stochastic Freight Rates

Extending the earlier literature, Adland and Strandenes (2004) presented a stochastic extension to the classical partial equilibrium models of the spot freight rate market. In order to do this, supply was based on a microeconomic analysis of the supply characteristics inherent in a given fleet and order-book as well as the stochastic ordering and demolition behaviour. These supply characteristics were then combined with stochastic demand to form a model, which was in turn used simulate scenarios for the future VLCC spot rate.

Adland and Strandenes (2004) argued that two schools of thought had arisen regarding the behaviour of freight rates in the spot shipping market. The first of these focused, in line with the classical literature, on modelling the demand and supply functions in the shipping market. This was done using either static supply / demand models, as was the case in Zannetos (1966), Norman and Wergeland (1981) and Evans (1994), or dynamic econometric models, such as those models outlined in Eriksen and Norman (1976), Strandenes (1986), Beenstock and Vergottis (1989) and Lensberg and Rasmussen (1992). The other school focused on modelling the freight rate directly as a stochastic model, which includes the Ornstein-Uhlenbeck process proposed by Bjerksund and Ekern (1995), Tvedt (1997) and Martinussen (1993); the Geometric Mean Reversion process proposed by Tvedt (1997); and the non-parametric model proposed by Adland (2003). Adland and Strandenes (2004) noted that both supply/demand and stochastic models have limitations in that a supply / demand model relies on a large number of variables, a large set of simultaneous equations, and weak econometric models, where a general discussion of these limitations can be found in Birkeland (1998). In the case of the stochastic models, however, these disregard all information not embedded in the current spot freight rate level and past freight rate process.

Adland and Strandenes (2004) bridged these two schools of thought by modelling the interaction of the supply and demand curves in a stochastic partial equilibrium framework, in combination with microeconomic modelling of the time-varying shape of the supply curve. This model incorporated stochastic ordering and scrapping dynamics into the supply curve as well as tracked corresponding changes in the fleet.

As the term partial equilibrium implies, Adland and Strandenes only looked at freight rate equilibrium within a sector of the bulk ship sector, i.e. the VLCC sector. In addition to this, the potential for short-term differences in freight rates across different geographical regions is ignored, i.e. it is assumed that the spot freight market within a particular sub sector behaves as a single market in the short-term.

The supply function, derived by Adland and Strandenes (2004), conformed to the classical shape proposed by Koopmans (1939) and Zannetos (1966) amongst others. Having created a starting point for the supply function, the next step in modelling the supply function was to model the dynamic response of this function to changes in the freight market in which the vessel operates. The first item examined in terms of the response of the supply function was the level of scrapping, where Dixit (1992) provided a general discussion of the optimal point at which to scrap one's vessel. With this in mind, Adland and Strandenes (2004) argued that scrapping volume follows a stochastic Poisson process, in which the expected number of ships scrapped in the time interval, denoted Δ, was a function of the prevailing freight market conditions on date t_j. The authors assumed that λ_j was the average scrapping rate and S_j was the number of ships scrapped, such that $\mathrm{E}(S_\Delta) = \lambda_j \cdot \Delta$ was the expected number of scrapped vessels in the next time interval, Δ, conditional on the information set available on date t_j. Therefore, according to the Poisson distribution with the parameter $\lambda_j \cdot \Delta$, the probability that k vessels will scrapped during the next time interval Δ was:

$$P(S_\Delta = k) = \frac{e^{-\lambda_j \Delta}(\lambda_j \Delta)^k}{k!} \text{ for } k = 0;1;2;... \qquad (2.25)$$

The next item relevant to the supply function was the level of deliveries, where if it is assumed that new-building projects cannot be accelerated, postponed or cancelled, then the number of new-buildings that will be delivered in the next period is known with certainty. However, this is not the case in practice, where, if freight markets are poor, then ship-owners are able to negotiate for projects to be delayed or cancelled, whereas, if freight markets are in a good position, then these projects may also be accelerated. Therefore Adland and Strandenes proposed that the number of new orders follows a stochastic Poisson process, where the average contracting rate during the time interval Δ is a function of freight market conditions on date t_j.

Adland and Strandenes (2004) assumed that the average contracting rate would be γ_j, where O_j was the number of new VLCC orders placed, such that $\mathrm{E}(O_\Delta) = \gamma_j \cdot \Delta$. Therefore, according to the Poisson distribution, with the parameter $\gamma_j \cdot \Delta$, the probability that k new orders would be placed during the next time interval Δ was:

$$P(O_\Delta = k) = \frac{e^{-\gamma_j \Delta}(\gamma_j \Delta)^k}{k!} \text{ for } k = 0\,;1\,;2\,;... \qquad (2.26)$$

When looking at the demand side of the model, Adland and Strandenes (2004) highlighted the fact that the exact shape of the demand function could not be determined empirically, as demand is exogenous to the market, hence the demand function will always have to be based on supposition alone. However, the authors argued that it did not seem fair to assume that the shape of the demand function would be dependent to some extent on the prevailing level of freight rates. Therefore, Adland and Strandenes assumed that the demand function was a simple linear function with respect to freight rates, where the slope of the function was calibrated so as to replicate historical volatility in the market. The authors assumed that demand for VLCC services followed a simple discrete process as follows:

$$D_j = \eta(\alpha + \beta X_{j-1} + \varepsilon_j) \qquad (2.27)$$

One should note that in Expression (2.27) above, X_t denoted the stochastic demand for shipping, where $\varepsilon \sim N(0;s)$. Adland and Strandenes noted that a potential limitation of this model was that this demand process did not allow for seasonality in the demand for oil transportation, even though this is a well known feature of tanker markets, as outlined in Kavussanos and Alizadeh (2002).

Adland and Strandenes (2004) concluded that their paper developed and estimates empirically a stochastic equilibrium model of the VLCC market, where this model could be applied to any other bulk shipping sub sector. However, the authors noted that simulations revealed that the Poisson process could not fully account for occasional large jumps in level of orders or number of scrappings, therefore the addition of a jump process may be necessary to capture this behaviour.

While this current research does not disagree with Adland and Strandenes (2004) regarding the fact that freight rates following a mean reverting process, it proposes that this mean reversion occurs with a lag due to the delay between the ordering and delivery of new vessels. Therefore, this research feels that freight rates do not follow a stationary process, as suggested by Adland and Strandenes, but instead follows a fractionally integrated process.

2.3.5.5 Adland and Cullinane (2006) and the Partial Equilibrium Model

The discussion the stationarity of the underlying freight rate process in the shipping markets was continued by Adland and Cullinane (2006), which examined the dynamics of the freight rates in the tanker market using a general non-parametric Markov diffusion model. The results indicated that the spot freight rate dynamics in the market could be best described by a non-linear stochastic model. The authors also illustrated that the spot freight rate in the tanker market did illustrate mean reversion; however, this was only mean-reverting in the tails of the distribution, and that the volatility of freight rate changes increase with the level of the freight rate. The result regarding mean reversion in the tails of the distribution, which implied that the spot freight rates process behaves like a Martingale over most of its empirical range, may explain why non-stationarity is difficult to reject over short samples, yet spot freight rates are globally mean reverting as implied by maritime economic theory, as discussed by Koopmans (1939) and Beenstock and Vergottis (1989), amongst others.

Adland and Cullinane (2006) argued that, apart from the work done in Tvedt (1996) and Tvedt (2003), where the spot freight rate was modelling in a stochastic partial equilibrium framework, spot freight rate models had been restricted to simple parametric models adopted from financial econometrics. Examples of these included the Geometric Brownian motion, outlined in Dixit and Pindyck (1994), embedded in Black and Scholes (1973)'s Black-Scholes models used by some practitioners; the Ornstein Uhlenbeck process, outlined in Vasicek (1977), which was used by Bjerksund and Ekern (1995) and Tvedt (1997); and the lognormal process, described by Brennan and Schwartz (1979), which was used in Tvedt (1997). In their work, Adland and Cullinane (2006) proposed to extend the recent literature on non-parametric modelling of economic variables to the maritime market, where this was examined for the first time.

Adland and Cullinane (2006) contributed to the existing literature in at least two important ways. First, the use of a fully functional methodology enabled them to investigate the spot freight rate dynamics in the shipping market in a generalised framework, i.e. it enabled the data to speak from themselves rather than imposing some ‘arbitrary’ parametric restrictions on the data. Second, previous evidence regarding the stationarity of spot freight rates process was ambiguous at best, where preliminary unit root tests either failed to reject the null of non-stationarity, as was the case for Berg-Andreassen (1996) and Glen (1997), amongst others, or provided results where the statistics were very close to the rejection threshold. Koekebakker, *et al.* (2006) proposed that this failure to reject non-stationarity was partially due to the short-coming of the empirical tests, as discussed in Schwert (1989), and that such findings are contrary to the classical maritime economics concept that prices in a freight market must revert towards the long-term costs, as outlined by Zannetos (1966). Adland and Cullinane (2006)’s findings suggested that spot freight rates were locally non-stationary over the range of the process, nevertheless, the existence of a non-linear mean reverting trend in the tails of the distribution was sufficient to pull the series back into the middle region and determine global stationarity. This result was consistent with other empirical findings for short-term interest rates in the non-parametric literature, as discussed in Ait-Sahalia (1996) and Jiang (1998), amongst others.

Adland and Cullinane (2006) modelled the dynamic evolution of the spot freight rate process as a general Markov stochastic differential equation, hence:

$$dX_t = \mu\left(X_t\right)dt + \sigma\left(X_t\right)dZ_t \qquad (2.28)$$

In Expression (2.28) above, Z_t was a one-dimensional standard Brownian motion, where μ and σ were the drift and diffusion functions, or the instantaneous conditional mean and standard deviation, respectively. One should note that μ and σ

were known values of the contemporaneous value of the spot freight rate, denoted X_t. As the Brownian components of this expression were Gaussian, the distributions of the process, i.e. the marginal and transitional densities, were entirely characterised by the drift and distribution functions, i.e. μ and σ, and, by the Markov property, the properties of long transitions could be derived by iterating the short transition. The most general approach to estimating stochastic differential equations is to avoid any functional form specifications of the drift and diffusion terms; hence Adland and Cullinane noted that the recent financial had turned to non-parametric estimation of the model in Expression (2.28).

The spot freight rates used by Adland and Cullinane (2006) were defined as the arithmetic average of the time-charter equivalent (TCE) spot freight rates for selected round voyages for a hypothetical generic vessel. The authors outlined three reasons for choosing a TCE spot freight rate, which is measured in US$ per day, rather than the actual spot freight rate, which is measured either in Worldscale or US$ per tonne, itself. The first of these was that the TCE measure was used for the valuation of contingent claims, as discussed by Tvedt (1997), and time-charters with embedded options. Second, the TCE spot freight rate was directly comparable to the time-charter rate; therefore any empirical results could be used to model the term-structure of freight rates. Finally, while over-the-counter freight derivatives in the tanker market are settled against the Worldscale rate, by using the TCE rate the authors avoided difficulties associated with the annual change in the Worldscale schedule.

Adland and Cullinane (2006) noted that given that the estimators for the drift function outlined in Expression (2.28) were based on the assumption of stationarity, it was essential to establish that the time series adhered to this assumption. In order to ensure this, the authors performed an Augmented Dickey-Fuller (ADF) test, first developed by Dickey and Fuller (1981), where the lag length was chosen based on a minimisation of the Scwhartz Information Criterion, outlined in Schwarz (1978). In the empirical literature, even a slight rejection of the null hypothesis of non-stationarity, which was not the case here, was interpreted as strong evidence of stationarity, as demonstrated by Ait-Sahalia (1996), amongst others, due to the low power of the ADF test. For this reason, Adland and Cullinane (2006) also reported the results for the KPSS and Phillips-Perron unit root tests, first discussed by Kwiatkowski, *et al.* (1992) and Phillips and Perron (1988), respectively. The results of these latter tests supported the authors argument that tanker spot freight rates were stationary, where these results were consistent with the empirical results in Kwiatkowski, *et al.* (1992).

Despite the fact that there was little support in maritime economic theory for any particular functional form of the drift function the Markov stochastic differential equation in Expression (2.28), Adland and Cullinane (2006) proposed that the potential for supply adjustment, through new-building and demolition, guaranteed that extremely high or low freight rates were not sustainable in the long-run. Therefore, the spot freight rate could not exhibit the asymptotically explosive behaviour that would be implied by a non-stationary process, and, as pointed out by Zannetos (1966) and Strandenes (1984), freight rates should revert towards some long-run equilibrium related to cost. However, a flaw with this argument of stationarity is that the long production time for new vessels means that supply can only very slowly adjust to unexpected changes in demand, as highlighted by Koekebakker, *et al.* (2006). Moreover, due to the high volatility of spot freight rates, it is relatively difficult to detect such slow-speed mean reversion in high frequency data, a fact highlighted by Dixit and Pindyck (1994).

Adland and Cullinane (2006) argued that the shape of the diffusion function was more certain, with the characteristic hockey-stick shape of the short-run supply function, described in Koopmans (1939), being well-established, in that it is once again described in Zannetos (1966) and Devanney (1973), amongst others, where the option to lay-up and the upper limit to capacity in the short-run leads to a short-run supply function that is near perfectly elastic at low freight rate levels and close to perfectly inelastic at full capacity. Furthermore, Zannetos (1966) points out that demand is assumed to be highly inelastic with respect to freight rate levels due to the relatively high cost of substitutes.

Despite the fact that Adland and Cullinane (2006) primarily provided a descriptive analysis, the authors considered two specific null hypotheses, viz., that the spot freight rate was a Martingale; and, that the conditional standard deviation was constant. One should note that both of these hypotheses were rejected. All the drift unction appeared to be non-linear, with the speed of mean reversion increasing in conjunction with the freight rate level, in a manner similar to the lognormal process, first described by Brennan and Schwartz (1979), which had been applied to the VLCC market by Tvedt (1997). This being said, while there was statistically significant mean reversion in the tails of the distribution, it was not possible to reject Martingale behaviour over the majority of the range of the spot freight rates. As stated previously, this non-linear behaviour may explain why non-stationarity could have been difficult to reject over short samples, yet the spot freight rate process was globally mean reverting. The results in the tanker sectors confirmed Adland and Cullinane (2006)'s *a priori* expectation of increasing conditional standard deviation in the spot freight rate level. On this basis, the constant volatility specification, such as the Ornstein-Uhelnbeck process, attributed to Vasicek (1977), used in the earlier research by Tvedt (1997) and Bjerksund and Ekern (1995), was rejected.

Adland and Cullinane (2006) concluded that the paper provided empirical evidence that the spot freight rate was locally non-stationary over the range of the process with a drift very close to zero; however, the existence of a non-linear mean reverting drift in the tails of the distribution was sufficient to pull the series back to its middle region and determine global stationarity. The authors also emphasised the importance of Martingale behaviour of the spot freight rate series over most of its range, highlighting that this disputed linear mean reverting models and explained the difficulty in rejecting non-stationarity in short samples. Furthermore, the authors also found a statistically significant level effect in the conditional volatility of spot freight rate changes, which suggested that the diffusion functions of some parametric models were incorrectly specified.

Criticism regarding this research include the fact that, as highlighted by Adland and Cullinane (2006), there are doubts regarding the efficiency of the ADF and Phillips-Perron tests used to determine whether the data series were stationary or not. This raises the question as to whether the result of non-stationarity for a portion of the distribution was a spurious result. In addition, while this current thesis does not disagree with the traditional maritime economic theory regarding freight rates following a mean reverting process, it proposes that this mean reversion occurs with a lag due to the delay between the ordering and delivery of new vessels. For these reasons, this research proposes that spot freight rates in the shipping sectors follow a fractionally integrated process.

2.3.5.6 Koekebakker et al. (2006) and Mean Reversion

In some of the recent maritime economic literature it was argued that empirical tests of stationarity often conclude that spot freight rates follow a non-stationary process. This argument is in contrast with traditional maritime economic theory, which argues that freight rates cannot exhibit asymptotically explosive behaviour, as would be implied by non-stationarity.

Koekebakker, *et al.* (2006) argued that when freight rates are high, the increase in the supply of transportation will inevitably bring freight rates down to a level that yields a normal economic profit and vice versa. This would imply that, in the long-term, one would expect freight rates to be highly correlated with long-term costs, hence, in an economic equilibrium setting, one would expect the freight rate to be a mean reverting variable. Thus, the authors argued that the freight cannot exhibit the asymptotically explosive behaviour implied by a non-stationary process. In support of this argument, Koekebakker, *et al.* noted that there were studies that confirmed this argument, for example Zannetos (1966) and Strandenes (1984), amongst others, despite the fact that most empirical studies on the behaviour of freight rates in the maritime economics literature, such as Berg-Andreassen (1996) and Glen (1997), concluded that the spot freight rate, or its time-charter equivalent (TCE), is non-stationary.

Koekebakker, *et al.* (2006) argued that these findings that freight rates have a unit root is, perhaps, not surprising for at least three reasons. The first of these is that most time series of freight rates are found to be highly persistent, a finding outlined by Adland and Cullinane (2006), which, as Dixit and Pindyck (1994) highlighted, makes the hypothesis of a unit root difficult to reject. The second reason lies in the choice of the model, where the most commonly used test, i.e. the Augmented Dickey-Fuller test, developed by Dickey and Fuller (1981), is based on a linear additive model displaying symmetric adjustment. Adland and Cullinane (2006) applied non-parametric estimation techniques to illustrate that the drift term of the spot freight rate process is mean reverting; however, this was found to only be the case in the tails of the freight rate distribution, with the process exhibiting unit root behaviour over the majority of its empirical range. Adland and Cullinane argued that non-linear behaviour in freight rates could explain why non-stationarity may be difficult to reject over short samples and yet the spot freight process was proved to be globally mean reverting and overall stationary. Koekebakker*, et al.* (2006) thus argued that in a non-linear environment, traditional unit root tests are inherently unsuitable for testing for non-stationarity. The final reason for the lack of surprise in the findings was that even if the assumption of non-stationarity does not hold in the strictest sense, it may be convenient from a technical standpoint as stationary processes are typically more complicated to deal with as far as investment decisions in the shipping industry are concerned. For these reasons, the authors set out the main objective of the paper, the assessment of whether the non-stationarity property put forth in the empirical literature was robust.

To this end, Koekebakker*, et al.* (2006) argued in favour of non-linear stationarity being the natural null hypothesis for testing the stochastic property of shipping freight rates. The most popular tests at the time were linear unit root tests, which take non-stationarity as their null hypothesis and test this against a linear stationary alternative; however, the authors also considered unit root tests against a non-linear stationary alternative. Koekebakker, *et al.* proposed that a standard way to have proceeded with this empirical work was to first apply the Augmented Dickey-Fuller and / or Phillips-Perron tests, and then confirm these results using the KPSS test, which has a null hypothesis of stationarity, where the latter two tests were developed by Phillips and Perron (1988) and Kwiatkowski*, et al.* (1992), respectively. This being said, Schwert (1989) illustrated that both the Augmented Dickey-Fuller and Phillips-Perron tests lack power, while Caner and Kilian (2001) in addition to Kuo and Tsong (2005), amongst others, showed that the KPSS test also has undesirable properties. Kapetanios*, et al.* (2003) suggested a unit root test, named the KSS test, against a non-linear globally stationary exponentially smooth transition autoregression (ESTAR) process, which was found to find to have good size and power properties relative to the traditional Augmented Dickey-Fuller test when the process is stationary and highly persistent. Kapetanios and Shin (2008) investigated efficiency improvements using GLS-detrending and found that GLS-detrending could improve the power performance of the KSS test. Koekebakker, *et al.* (2006) used both the KSS test and the point-optimal version in the empirical part of the paper.

Using the standard Augmented Dickey-Fuller test, Koekebakker, *et al.* (2006) found that only the Suezmax data series showed evidence of stationarity around a constant mean, however, if one added a trend to the alternative hypothesis, the test statistics could not reject the null of non-stationarity for this market. For the KSS test, the results were very different from the linear unit root tests, as the test statistics suggested that all freight rates were non-linear stationary for both the constant and constant with drift alternatives, where the authors arrived at this same result when using both the original and the point-optimal versions of the test. Koekebakker, *et al.* concluded that although traditional linear unit root tests still suggested that freight rates are non-stationary, when the authors employed a non-linear unit root test, the results suggested that freight rates across all bulk shipping sectors are stationary, in line with classical maritime economic theory.

While this current research does not disagree with Koekebakker, *et al.* (2006) regarding freight rates following a mean reverting process, it proposes that this mean reversion occurs with a lag due to the delay between the ordering and delivery of new vessels. Therefore, this research feels that freight rates do not follow a stationary process, as suggested by Koekebakker, *et al.*, but instead follows a fractionally integrated process.

2.3.5.7 Concluding Comments

Given the findings discussed in this section, one can thus conclude that the general consensus among these six papers is that freight rates should follow a stationary process. Any disagreement between these papers is merely regarding the form of stochastic differential model which the freight rate process follows. In their research, Adland and Cullinane (2006) and Koekebakker, *et al.* (2006) went further by arguing that the reason for the empirical literature which proved that freight rates followed a non-stationary process was that the traditional linear unit root tests are flawed, hence Koekebakker, *et al.* proposed the use of the non-linear KSS unit root, first illustrated by Kapetanios, *et al.* (2003).

2.3.6 Critique of Partial Equilibrium Models

Although these papers conformed to classical maritime economic theory, none of these papers allowed for the fact that the mean reversion process is delayed due to the lag between the ordering and delivery of new vessels. In addition to this, stochastic models, as was pointed out by Adland and Strandenes (2004), disregard all information that is not embedded in the current spot freight rate level and past freight rate process.

Koekebakker, *et al.* (2006) highlighted the fact that it is possible to model freight rates as process with long memory, although non-linearity and long memory have rarely been jointly analysed in the general econometric literature. Kapetanios (2006) illustrated that stationary threshold and Markov switching models may exhibit long memory properties, i.e. slowly decaying auto-covariance, which means that non-linear models may be incorrectly taken to be long memory non-stationary processes. Furthermore, Adland and Cullinane (2006) proposed that freight rates only followed a stationary process in the tails of the respective distribution.

This research aims to address these issues by looking at the possibility that long memory may be responsible for the confusion regarding the stationarity of freight rate process. The research then goes on to look at whether the same can be said regarding the volatility of freight rates, before examining the concepts of whether these may be as a result of the presence of conditional skewness and kurtosis, as is discussed in the following sections of this literature review.

2.4 The Issue of Volatility in the Shipping Freight Markets

An understanding of the concept of uncertainty is crucial for the decision-making process in any industry, it is therefore crucial for players in a market to be able to calculate risks, as measured by the volatility of the price series being examined, and if possible, minimise these through the use of some diversified portfolio of assets. The research by Engle (1982) provided the first insight into the modelling of volatility in the financial markets with the development of the Autoregressive Conditional Heteroscedasticity (ARCH) family of models, thereby enabling market participants to gain a better understanding of and enabling them to reduce their risk exposure. The shipping industry in particular is highly dependent on a number of external factors, over which it has no control. As a result of this lack of direct control, freight rates in the shipping markets are exceptionally volatile, with volatility increasing as ship-size increases as smaller ships are more versatile, and can operate in multiple markets, whereas, for example, a VLCC tanker is limited in the ports it can visit and the cargo it can carry. In addition to this, the demand for shipping services is a derived demand and therefore is exogenous to the market, adding to the volatility of freight rates.

Despite this fact, there has been very little research regarding the modelling of the volatility of freight rates. There are, however, three seminal papers on modelling volatility in the shipping markets, which shall be examined in more detail below. Beginning with Kavussanos (1996), which provided the first empirical examination of the nature of volatility in the shipping markets, this review moves on to examine the later work on the topic in Kavussanos and Visvikis (2004), as well as giving a brief overview of other work on the area, and then concludes by providing a critical analysis of the contribution of this thesis.

2.4.1 The ARCH Family of Models and the Shipping Markets

As stated above, Engle (1982) provided the first insight into the modelling of the structure of volatility in the financial markets; however, there is very little in the maritime economic literature regarding the modelling of the volatility of freight rates. The following subsections provide a detailed account of the two seminal papers on modelling the volatility of freight rates, i.e. the papers by Kavussanos (1996) and Kavussanos and Visvikis (2004), as well as giving a more brief account on other literature on the topic. One should note that the lack of literature on this topic does indicate that this is an ideal area for further research.

2.4.1.1 Kavussanos (1996) and the Introduction of Volatility Models

When introducing the concept of modelling the volatility of freight rates to the shipping literature, Kavussanos (1996) stated that addition to allowing one to estimate the volatility of prices over time, the use of Autoregressive Moving Average (ARCH) models over a simple assumption of constant variance throughout the estimation period, as would be used by Ordinary Least Squares (OLS), allowed for improved estimation results. In this paper, the author extended the use of ARCH class models to investigate the volatility of freight rates in the spot and time charter markets for dry bulk vessels, thereby investigating two keys issues. The first of these issues was whether a market effect existed in terms of whether the volatility of freight rates was higher in the spot or time charter market? The second issue was whether a size effect existed for volatility, i.e. was there a difference between the volatility of spot freight rates for smaller and larger vessels?

Kavussanos (1996) explained that the process of building an ARCH model requires three steps. The first of these is specifying the conditional mean of the variable being studied, in this case freight rates. The second step is specifying the conditional variance of the variable, once again in this case freight rates. Having completed the two prior steps, the final step requires that the conditional density of the error term in the regression equation is specified. Kavussanos noted that freight markets are perfectly competitive, where the market-clearing spot freight rate is a function of:

$$FR_t = f^{FR}\left(\overset{+}{IP_t}; \overset{+}{Pb_t}; \overset{-}{K_t} \right) \qquad (2.29)$$

One should note that in Expression (2.29), above, FR_t, IP_t, Pb_t and K_t denote the spot freight rate level, the level of industrial production, the price of bunkers and the size of the world fleet at time t, respectively, while the signs of the partial derivatives of each variable are given above the respective variable. In contrast, Kavussanos determined the market clearing time charter freight rate, where a one-period time-horizon was assumed, as a function of: the following:

$$TC_t = f^{TC}\left[E_t\left(\overset{+}{FR_{t+1}}\right); E_t\left(\overset{-}{Pb_{t+1}}\right) \right] \qquad (2.30)$$

In Expression (2.30), above, TC_t, $E_t\left(FR_{t+1}\right)$ and $E_t\left(Pb_{t+1}\right)$ denote the time-charter freight rate and the expected value of the spot freight rates and bunker prices in the following period, respectively, where this is all evaluated at time t.

Kavussanos noted that these two expressions could be rewritten as follows:

$$y_t = x_t^{'}b + \varepsilon_t \ ; \ \varepsilon_t \sim N(0;h)$$
$$LL = -\left(\frac{T}{2}\right)\ln h - \left(\frac{1}{2}h\right)\sum_t \varepsilon_t^2 \qquad (2.31)$$

Kavussanos highlighted that in Expression (2.31), above, ε_t denoted a white noise error term, with the usual classical properties, while LL denoted the corresponding log-likelihood function, after omitting the irrelevant constant.

Given the arguments above, Kavussanos (1996) noted that econometric theory argues that, although the Ordinary Least Squares (OLS) and Maximum Likelihood Estimation (MLE) procedures may provide the Best Linear Unbiased Estimators of b and h in Expression **Error! Reference source not found.**, above, which are the parameters of interest, if the assumption that the variance of ε_t, or equivalently y_t, which is given by h, is constant fails, this creates problems with these estimation procedures. The reason for this is that, despite the fact that the estimated parameters would remain unbiased and consistent, these estimated parameters will no longer be efficient. Furthermore, any of the estimated variances of these estimated parameters will be biased estimators of the true variances of these parameters, as outlined in Pindyck and Rubinfeld (1991). Kavussanos (1996) argued that these difficulties may be overcome should one use ARCH models, where the variance of the data series is conditioned on the available information set, in conjunction with the conditional mean.

In order to address this issue, Kavussanos (1996) proposed that the most appropriate parameterisation for h_t was the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model developed by Bollerslev (1986), Bollerslev (1986), where h_t is expressed as a linear function of p values of past squared errors and q past values of the conditional variances, h_t, i.e.:

$$y_t = x_t^{'}b + \varepsilon_t \ \ ; \ \ \varepsilon_t \sim N(0\,;h_t)$$
$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i} \qquad (2.32)$$
$$LL = -\left(\frac{T}{2}\right)\ln h_t - \left(\frac{1}{2}h_t\right)\sum_t \varepsilon_t^2$$

One should note that in Expression (2.32), above, $\alpha_0 > 0$ and $\alpha_i ; \beta_i \geq 0$ so as to ensure the non-negativity of the variance, while $\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i < 1$ so as to ensure the stationarity of h_t. Kavussanos (1996) noted that the MLE estimates of the mean and variance equation parameters, i.e. b, α_i and β_i, are asymptotically superior, in terms of efficiency, to those obtained by using the OLS procedure on the GARCH model, as the parameters achieved through the OLS estimation do not achieve the Cramer-Rao lower bound, a fact outlined in Engle (1982). Kavussanos (1996) went further to note that the MLE estimation procedure is non-linear and is achieved by solving the first-order conditions with respect to the GARCH parameters using a numerical optimisation method. Additionally, Kavussanos argued that the GARCH formulations would capture the tendency for large (small) swings in the freight rate to be followed by large (small) swings of random direction, i.e. volatility. This is a phenomenon that was widely covered in the maritime economic literature, for example in the research by Adland and Cullinane (2006), amongst others. Another important consequence of these models, highlighted by Kavussanos (1996), is that the parameters in these models may be estimated using historical data and then be used to model and thereby forecast future volatility patterns in the respective data series being examined, in this case spot and time-charter freight rates.

Prior to estimating the models, Kavussanos (1996) performed Dickey-Fuller and Augmented Dickey-Fuller tests, attributed to Dickey and Fuller (1979) and Dickey and Fuller (1981), respectively, to test for stationarity in the data series, and found that all variables followed an $I(1)$ non-stationary process, apart from the logarithm of the vessel fleet, which was stationary. Following this, Kavussanos used the Johansen procedure, outlined in Johansen (1991), to test for cointegration, where the results indicated that the variables within each equation turned out to be cointegrated. Kavussanos thus concluded that this had two main implications. First, the regressions were related through a long-run economic relationship; and, second, inferences regarding linear restrictions on the parameters could be made through the classical distribution theory.

Following this, Kavussanos (1996) specified the conditional means of the dependent criteria, using the OLS procedure, which was done according to standard statistical criteria; and then, subsequently, modelled the time-varying error volatilities, utilising the ARCH / GARCH estimation procedure. The results indicated that the demand-side variable, i.e. world industrial production, had a significant and positive effect on freight rates; while the supply-side variables, i.e. bunker prices and the size of the fleet, had a significant positive and negative impacts on freight rates, respectively, Kavussanos noted that the use of monthly data required that the dynamic specification of the equations in order to account for the short-run dynamics present in the conditional means, where this was achieved by including one or two lags of the dependent variables. Kavussanos found that a dynamic specification of the expectations of freight rates, the current and lagged values of freight rates, and one period lag of time-charter rates were sufficient to capture the driving forces in the conditional mean of the time-charters. In addition to this, tests of exogeneity in relation to the logarithm of freight rates could not reject the null hypothesis, which thereby legitimised the use of the OLS and ARCH estimation procedures.

Kavussanos (1996) also found that standard statistical diagnostic tests indicated that the fit of all the equations was good; that there was no serial correlation, with the exception of the Capesize category; and that, in general, the data series illustrated no skewness. This being said, however, Kavussanos did find that there was significant kurtosis in the equations and that heteroscedasticity appeared to be a common finding across all equations. Kavussanos argued that the finding of ARCH effects and leptokurtosis in the OLS results justified the use of ARCH models for these regressions, where the appropriate specification for these was determined using likelihood ratio tests. Kavussanos gave the appropriate specification for the models as follows: the aggregate freight rate was modelled using an $\mathrm{ARCH}(1)$ model; time-charters were modelled using a $\mathrm{GARCH}(2;1)$ model; and the three weight categories were modelled using a $\mathrm{GARCH}(1;1)$ model. The author found that re-run diagnostic tests indicated that the use of the models had solved the problem of heteroscedasticity; that serial correlation statistics improved when compared to those from the OLS estimation; while the same applied regarding the levels of skewness and kurtosis in the models. Having modelled the conditional variances for each market in question, Kavussanos extracted these time-varying measures of risk, as well as their behaviour examined over time, and compared these across the relevant markets.

Kavussanos (1996) concluded that risk in the freight rate and time-charter markets was not constant over time, which was manifested by the need for ARCH modelling of the conditional variances, where such time-varying risk were considered to be a combination of industry-specific risk and idiosyncratic risk. However, the author argued that so long as the market participant was faced with more than one option, this risk could have been diversified. Kavussanos also noted that there was a clear tendency for volatility clustering, where volatility was high during, and just after, periods of large external shocks; and, that volatility appeared to be higher in the time-charter market than in the spot freight market. The reason that the author gave for this disparity in volatility between the two types of market was that as time-charter freight rates reflected future expectations, these were more sensitive to changing perceptions of the future market. This would have meant that a risk-averse ship-owner would have, if presented with a choice between the two types of market, exploited this information by preferring the spot market over time chartering, although this would have meant a lower return. When comparing volatilities between different sizes of vessels, Kavussanos found that risk premiums were generally higher for larger vessels, where this was due to the limitations in the trades in which larger vessels could participate; therefore, a risk-averse ship-owner would most probably have invested in smaller vessels rather than larger ones.

Once again there are doubts regarding the efficiency of the tests used to determine whether the data series are stationary or not. If the data series are not non-stationary as proposed by Berg-Andreassen (1996), amongst others, then the results of the cointegration analysis will be spurious. In fact traditional maritime economic theory, such as outlined in Zannetos (1966), amongst others, argues that due to supply and demand fundamentals, freight rates are in fact mean reverting, while more recent research, such as Adland and Cullinane (2006) argue that they are stationary. This being said, this research adds another dimension to the debate by arguing that, due to the lag between the ordering and delivery of new vessels, freight rates are in fact fractionally integrated. If this is the case, then the correct model to use would not be the GARCH model proposed by Kavussanos (1996), but the Fractionally Integrated

Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) model, developed by Baillie, *et al.* (1996a), where this is a notion that forms one of the hypotheses explored in this research.

2.4.1.2 Kavussanos and Visvikis (2004) and Volatility Revisited

A later study on the volatility of freight rates was the comprehensive study by Kavussanos and Visvikis (2004), which examined the lead-lag relationship between spot and futures markets, in this case the over-the-counter Forward Freight Agreement (FFA) market, for both the returns and volatilities of the price series. The section of this paper that dealt with the relationship between returns is discussed above in the section on non-stationary models, with the focus now being on the modelling of volatility. Kavussanos and Visvikis argued that, at the practical level, the better one understood the mean and variance dynamics of a process, the better one could improve risk management and budgeting decisions.

In an earlier study, Working (1970) argued that price discovery refers to the use of prices from one data series in order to predict the prices in another data series. Working continued by stating that this lead-lag relationship between price-movements in the derivatives market and the underlying spot market illustrates how quickly the one market reflects new information, such as shocks, relative to the other, in addition to illustrating the degree to which the two markets are linked. Bollerslev, *et al.* (1992) moved this argument from the first to the higher moments, by arguing that volatility spill-overs from the one market to the next arose primarily as a result of the realisation that speculative price changes were being interwoven with higher moment dependencies. Kavussanos and Visvikis (2004) gave three reasons for this lead-lag relationship being of interest to academics. The first of these was that the issue was related to market efficiency, where if the futures market is efficient, the volatility of the futures prices will give unbiased estimates of volatility of future spot prices. The second reason was that the derivatives market could have been potentially used as a tool for price discovery. The final reason was that if volatility spill-overs did exist from the one market to the other, then the volatility transmitting market could have been used by market participants as a vehicle of price discovery in order to cover their risk exposure. Kavussanos and Visvikis noted that other papers interested in this said relationship between markets in terms of higher moments, i.e. with respect to the time-varying volatility spill-overs, included Ng and Pirrong (1996) and Koutmos and Tucker (1996), amongst others.

Kavussanos and Visvikis (2004) investigated this lead-lag using a multivariate Vector Error-Correction-Generalised Autoregressive Conditional Heteroscedasticity (VECM-GARCH) model, where the variances and covariances of the underlying series were allowed to vary over time, where this enabled the authors to allow for volatility spill-overs. This procedure also ensured that there was an efficient econometric specification, in addition to improving market analysis and any forecasts that were made. This model used had the following form with augmented positive definite parameterisation, as outlined by Baba, *et al.* (1990):

$$\begin{aligned} H_t = {} & A'A + B'H_{t-1}B + C'\varepsilon_{t-1}\varepsilon'_{t-1}C + S1'u_{1,t-1}u'_{1,t-1}S1 + S2'u_{2;t-1}u'_{2,t-1}S2 \\ & + E'\left(z_{t-1}\right)^2 E \end{aligned} \tag{2.33}$$

In Expression (2.33), above, A is a (2×2) lower triangular matrix of coefficients; B and C are (2×2) diagonal coefficient matrices; β_{kk}^2 and $\gamma_{\text{kk}}^2 < 1$, where $k = 1$; 2 for stationarity; $S1$ and $S2$ are matrices, which contain parameters of spill-over effects; $u_{1,t-1}$ and $u_{2,t-1}$ are matrices whose elements are lagged squared error terms, where $u_{1,t-1}$ represents the volatility spill-over effect from the spot to the derivatives market and $u_{2,t-1}$ represents the volatility spill-over effect from the derivatives to the spot market; $\left(z_{t-1}\right)^2$ is the lagged squared basis; and E is a (2×2) vector of coefficients of the lagged squared basis. The conditional variances were considered to be a function of their own lagged values, i.e. the effect of old ‘news’; their own lagged error terms, i.e. the effect of new ‘news’; and a lagged squared basis parameter; while the conditional covariance was considered to be a function of the lagged covariances and the lagged cross-products of the residuals. In this model, the volatility spill-over effects between the spot and derivatives market volatilities could be tested through the coefficients of the two matrices $S1$ and $S2$.

Prior to estimating the model, Kavussanos and Visvikis (2004) found that the diagnostic tests results indicated that there was excess skewness and kurtosis in all prices series; however, the excess kurtosis seems to be of a greater magnitude in the spot price series. The results also indicated that there was significant serial correlation and heteroscedasticity; while Augmented Dickey-Fuller and Phillips-Perron tests, attributed to Dickey and Fuller (1981) and Phillips and Perron (1988), respectively, indicated that the price series followed an $I(1)$ process, where these test results were confirmed using the KPSS test, developed by Kwiatkowski, *et al.* (1992).

When looking at the lead-lag relationship between spot and derivative volatilities, Kavussanos and Visvikis (2004) noted that, overall, the coefficients of the lagged error terms in spot variance equation were higher than those in the variance equation for the derivative for all routes, thus implying that that past shocks, or new 'news' had a greater impact on the spot rather than the derivative volatility. This being said, the coefficient of the lagged variance in the spot variance equation was lower than that for the derivative variance equation across all routes, thus implying that informed agents in the market would have used past volatility, or old 'news', more in the derivative market than in the spot market. The coefficients of the volatility spill-over effects, i.e. $s1_{21}$ and $s2_{1,2}$, would have picked up the effect of the lagged squared forecast errors, or the residuals, of the spot equation in explaining the volatility of the derivative rates, and vice versa, respectively. A volatility spill-over from market to another, in general terms, would mean that any piece of information that is released by the volatility transmitting market will play a superior information role and therefore, has an effect on the market that receives the volatility spill-over. Kavussanos and Visvikis found that there were no volatility spill-overs on Routes 1 and 1A, a finding that was consistent with Kawaller, *et al.* (1990), amongst others. This being said Kavussanos and Visvikis (2004) found that there were bi-directional volatility spill-overs on Routes 2 and 2A, a result that was found to be consistent with the empirical work of Chan, *et al.* (1991), amongst others. In their empirical analysis, the authors noted that the persistence of the volatility in the spot and derivatives markets, following a shock in the respective market, as measured by $b_{kk}^2 + c_{kk}^2$, showed that the unconditional variances were stationary, i.e. that the persistence factors were less than one. Kavussanos and Visvikis (2004) therefore concluded that derivatives markets discovered information more rapidly, when compared to the spot market, and that for practical purposes, information coming from price discovery vehicles could be used by market participants in the decision making process.

As with the previous research, there are, once again, doubts regarding the efficiency of the tests used to determine whether the data series are stationary or not. If the data series are not non-stationary as proposed by Berg-Andreassen (1996), amongst others, then the results of the cointegration analysis will be spurious. In fact traditional maritime economic theory, such as outlined in Zannetos (1966), amongst others, argues that due to supply and demand fundamentals, freight rates are in fact mean reverting, while more recent research, such as Koekebakker, *et al.* (2006) argue that they are stationary. This being said, this research adds another dimension to the debate by arguing that, due to the lag between the ordering and delivery of new vessels, freight rates are in fact fractionally integrated. If this is the case, then the correct model to use would not be the GARCH model proposed by Kavussanos (1996), but the Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) model, which this research will explore.

2.4.1.3 A Few Other Papers on ARCH Models and Volatility

Other papers in the shipping literature to have used the Autoregressive Conditional Heteroscedasticity (ARCH) family of models, first proposed by Engle (1982), include the papers by Kavussanos (1997), Kavussanos and Nomikos (2000b) and Kavussanos and Nomikos (2000a). In the first paper, Kavussanos (1997) used different forms of the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model, first proposed by Bollerslev (1986), to examine the dynamics of conditional volatilities in the market for second-hand vessels. In this analysis, Kavussanos introduced two different types of volatility modelling to the shipping literature. In the first type, a time series model was fitted to each price series, where the respective time series was described in terms of its own past values and past error terms; however, in this case, no attempt was made to understand the underlying structural economic variables that help determine ship prices. The second type of model used was where prices were explained in terms of other underlying structural variables. Kavussanos concluded that there was some support for the use of GARCH models with respect to modelling second-hand vessel prices, where structural variables, such as interest rates and time-charters, did appear to have a role in explaining the conditional variance of the series. The author noted that there appeared to volatility clustering in the data and that there were differences in the nature of the volatilities for different size vessels. Kavussanos finally proposed that a comparison of these different risk levels between different sizes of vessels could have been used as a tool by ship-owners to guide their holdings of different sizes of vessels in their dynamic varying portfolios.

In the second paper examined here, Kavussanos and Nomikos (2000b) used GARCH and augmented GARCH models to investigate the hedging effectiveness of freight futures on the Baltic International Freight Futures Exchange (BIFFEX) across different shipping routes. The authors found that time-varying hedge ratios out-performed alternate market specifications in reducing the risk inherent in a market participant's spot position; however, this reduction in risk was not of the same magnitude as experienced in other markets examined in the literature. The reason given for this was that these freight futures contracts were employed as a cross-hedge against the fluctuations of individual routes on an aggregate index. In addition to this, there was large inherent basis risk and freight rate fluctuations may have not been accurately tracked by futures prices.

In a later paper, Kavussanos and Nomikos (2000a) once again examined the time-varying hedge ratios generated using GARCH and augmented GARCH-X models, and then compared these with constant hedge ratios. Kavussanos and Nomikos found that in- and out-of-sample testing revealed that the GARCH-X specification provided a greater reduction in the risk levels, compared to the other two specifications. However, once again, these models failed to reduce the level of risk in the spot position to the extent experienced in other markets in the literature, and, once again, the reason given for this was the heterogeneous composition of the underlying index. The authors suggested that the index may have needed to be restructured in order to improve the hedging effectiveness of the freight futures contract.

As with the seminal papers, the papers raise some concerns regarding the regarding the efficiency of the tests used to determine whether the data series are stationary or not. If the data series are not non-stationary as proposed by Berg-Andreassen (1996) amongst others, then the results of the cointegration analysis will be spurious. In fact traditional maritime economic theory, such as outlined in Zannetos (1966), amongst others, argues that due to supply and demand fundamentals, freight rates are in fact mean reverting, while more recent research, such as Koekebakker, *et al.* (2006) argue that they are stationary. This being said, this research adds another dimension to the debate by arguing that, due to the lag between the ordering and delivery of new vessels, freight rates are in fact fractionally integrated. If this is the case, then the correct model to use would not be the GARCH model proposed by Kavussanos (1996), but the Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) model, a hypothesis explored here.

2.4.1.4 Concluding Comments

Therefore one can see that the overall consensus in the literature was that using Autoregressive Conditional Heteroscedasticity (ARCH) and Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models, developed by Engle (1982) and Bollerslev (1986), provided the most effective means of modelling volatility in shipping freight markets. These models were also all based on the assumption that freight rates followed a non-stationary process; however, this may not in fact be the case, as was illustrated by Koekebakker, *et al.* (2006), amongst others. One should note that this, and other short-comings, of these papers are discussed in fuller detail in the section below.

2.4.2 Critique of the Models for Modelling Freight Rate Volatility

The major flaws with the methodologies used above are two-fold. Firstly, these models were based upon the assumptions that the freight rate process was either stationary or non-stationary. As the research sets out to prove, this may not necessarily be the case, and in fact the freight rate process may follow a fractionally integrated process. The second problem with the models used in the papers above is that they ignore the problem of the potential existence long-memory, or persistence, in volatility. Koekebakker, *et al.* (2006) briefly mentioned the problem of persistence in levels of freight rates, discussed in the critique of models of levels of freight rates above, and therefore it may be of interest to look at the use of the Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) model, developed by Baillie, *et al.* (1996a) to further examine the issue of the persistence and structure of volatility in the freight rate series.

The first advantage of the FIGARCH model is that it allows for long memory in the second conditional moment of the process being examined in that it assumes a hyperbolic rate of decay for the volatility, unlike the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model, proposed by Bollerslev (1986) which assumes that volatility decays exponentially. This being said, unlike the Integrated Generalised Autoregressive Conditional Heteroscedasticity (IGARCH) model, developed by Engle and Bollerslev (1986), which assumes that volatility does not decay, the FIGARCH model does allow for the eventual decay of volatility. A further advantage of the FIGARCH model is that allows for the simultaneous comparisons of numerous potential models that can capture the features of the process, a characteristic outlined by Conrad and Karanasos (2005). For these reasons, the hypothesis that freight rate volatility is best modelled by a FIGARCH model is examined in Chapter 6 of the thesis.

2.5 A Review of Fractionally Integrated Processes

Having examined the literature on the first- and second-moments of the freight rate process in the previous sections, as well as the rationale behind this thesis' argument that freight rate levels and volatility follow a fractionally integrated process, this section examines the application of fractionally integrated process in other markets. The original work on fractionally integrated process was performed by Hurst (1951) who investigated this issue in the context of analysing river flow data. Following this, Adelman (1965) applied this concept to the financial markets, where Adelman proposed the use of long memory models to model long-run cycles in the macro economy. Since then, the use of long memory processes in finance has been extended to asset pricing models, exchange rates and interest rates.

2.5.1 Long Memory Processes in Asset Pricing Models

Traditionally, asset pricing models assumed that should there have been one unit root in the nominal price of an asset, then the continuously compounded rate of return, which is the fist difference of the price of the asset, could be expected to be stationary and was assumed to be uncorrelated so that it was well approximated as a martingale. Therefore, from the conventional asset pricing formula, if p_t was the price of the asset in question and x_t were the fundamentals in period t, then:

$$p_t = \sum_{j=0;\infty} \xi E_t x_{t+j} \qquad (2.34)$$

One should note that in Expression (2.34), above, ξ denoted the discount factor, where $0 < \xi < 1$.

Baillie (1989) and Campbell and Shiller (1987) rearranged this expression to reveal the following model:

$$\Delta p_t = \left(p_t - \xi x_t \right) \qquad (2.35)$$

The authors noted that if the price and fundamentals were both $I(1)$ processes, then Expression (2.35) implied that a cointegrating relationship existed between the asset price and the fundamentals. Any failure to find the regular form of CI(1;1) cointegration between prices and fundamentals may then not have necessarily been interpreted as a rejection of the asset price model, but the case may have been that a form of $CI(1;1-d)$ cointegration may have been apparent where residuals from the cointegrating vector were $I(d)$, where $0 < d < 1$, rather than $I(0)$. This would have implied a lower response to shocks and a longer time required to adjust back to equilibrium, i.e. a fractionally integrated long memory process.

2.5.2 Long Memory Processes and Stock Returns

Another financial application that has benefited from the introduction of long memory processes has been stock returns. For example, both Greene and Fielitz (1977) and Aydogan and Booth (1988) used the original rescaled range $\left(R_T / s_T \right)$ statistic, outlined in Hurst (1951), to test for long memory in common stock returns, where the components of this statistic were defined as follows:

$$R_T = \max_{0 \le j \le T} \left\{ \sum_{j=1}^{T} \left(y_j - j\bar{y} \right) \right\} - \min_{0 \le j \le T} \left\{ \sum_{j=1}^{T} \left(y_j - j\bar{y} \right) \right\} \qquad (2.36)$$

And:

$$s_T = \left\{ \left(\frac{1}{T} \right) \sum \left(y_t - \overline{y} \right)^2 \right\}^{\frac{1}{2}} \qquad (2.37)$$

One should note that in Expressions (2.36) and (2.37), above, R_T, s_T and $\overline{y}$ denote the sample range, sample standard deviation and sample mean, respectively.

In a later study, Lo (1991) compared the results between a modified rescaled range statistic and the original rescaled range statistic, described above, using returns from both value and equally weighted CRSP[11] indices for the period between July 1962 and December 1987. The modified rescaled range statistic used by Lo may be calculated in the following manner:

$$Q_T = \frac{R_T}{\sigma_T(q)} \qquad (2.38)$$

Where:

$$\sigma_T^2(q) = c_o + 2\sum_{j=1}^{q} w_j(q) c_j \qquad (2.39)$$

It is important to note that in Expression (2.39), above, c_j is the jth-order sample autocovariance of y_t; while w_j are the Bartlett window weights of the following:

$$w_j(q) = 1 - \left[\frac{j}{(q+1)} \right] \text{ for } q < T \qquad (2.40)$$

In this study, Lo (1991) found that the results when using the regular rescaled range statistic were statistic, while the results from using the modified rescaled range statistic were insignificant. Lo attributed this difference in results to the short-term persistence within the returns series, and also reported that there was no long-range persistence when using annual returns for the period between 1872 and 1986.

2.5.3 Long Memory Processes and Exchange Rates

Meese and Singleton (1982) and Baillie and Bollerslev (1989), amongst many others, provided evidence that the logarithm of nominal exchange rates contained a unit root

[11] One should note that the acronym CRSP stands for the Centre for Research in Security Prices, which is part of the Graduate School of Business of the University of Chicago.

and that the approximate rate of return was uncorrelated. This evidence indicated that a martingale model would be appropriate. Cheung (1993), however, provided contrary evidence that long memory existed in the French Franc / US Dollar exchange rate in addition to providing some marginal evidence of long memory in the UK Pound / US Dollar exchange rate.

A major issue in this literature was the speed of adjustment of exchange rates to shocks from disequilibrium. In their paper, Baillie and Bollerslev (1989) found that while seven nominal exchange rates exhibited non-stationary properties in their uni-variate time series representations, these rates also appeared to be tied together through one cointegrating vector. This being said, Hakkio and Rush (1991) and Sephton and Larsen (1991) found mixed results as to the existence of a cointegrating relationship between the same several exchange rates. Furthermore, Diebold, *et al.* (1994), using the same daily exchange rates over a five-year period, noted that the application of the Johansen procedure, outlined in Johansen (1991), to test for cointegration was sensitive to whether or not an intercept was included in the vector autoregression, and concluded that there was no cointegration between these spot exchange rates. Baillie and Bollerslev (1994a) provided evidence that a linear combination of these same exchange rates exhibited slowly decaying autocovariance, which is characteristic of long-range dependence.

Another area of interest in the literature on long memory processes in financial markets was the properties of real exchange rates and the potential validity of Purchasing Power Parity (PPP) as a long-run phenomenon; however, there has generally been little support for PPP. Kim (1990) reported some evidence of cointegration between nominal exchange rates and relative prices, having used the Johansen test, while Diebold, *et al.* (1991) estimated Autoregressive Fractionally Integrated Moving Average (ARFIMA) models for annual real exchange rate data using maximum likelihood estimation (MLE) methodology outlined by Fox and Taqqu (1986). The results in this study illustrated that shocks take a long time to return to equilibrium, however, this time-frame is finite, thereby being very supportive of the PPP doctrine. Further supportive evidence was provided by Cheung and Lai (1993) who tested for fractional cointegration between nominal exchange rates and relative prices, using annual data for the period between 1914 and 1972; and

Steigerwald (1996), who provided yet more evidence of PPP in a unit root framework. In contradistinction, Crato and Rothman (1994b) found evidence of mean reversion in UK real exchange rates, when estimating ARFIMA models by MLE. Baillie (1996) argued that the PPP issue was one of the best examples of researchers being misled by the low power of unit root tests, resulting in researchers abandoning PPP without paying sufficient attention to the econometric procedures.

2.5.4 Long Memory Processes and Interest Rates

Another area in which there have been some interesting applications of long memory processes is the interest rate markets. Shea (1991) estimated fractional process on a set of interest rates using the Geweke and Porter-Hudak (1983) Geweke and Porter-Hudak (1983) procedure. Shea also discussed the implication that long memory would have had on the variance bounds tests that would have result from the term structure. Some of the initial work done by Shea (1991) appeared to provide evidence of long memory in interest rate spreads and some interest rate levels. Backus and Zin (1993) found evidence of long memory using various time series, including, but not only, $\mathrm{AR}(1)$, unit root and fractional white noise processes. The authors also discussed the implications of the presence of fractional integration in the context of term structure, and upon comparing the implied forward rates and corresponding yields on maturities of n-period bonds, concluded that this assumption of long memory did not compare favourably with the alternatives. Backus and Zin noted that the estimation of various ARFIMA models for bond series was relatively inconclusive. Crato and Rothman (1994a) found a contrary result, when full Maximum Likelihood Estimation was used to estimate an $\mathrm{ARFIMA}(0;d;1)$ model for annual bond yields and the conclusion was reached that $d = 0.81$ and was significantly different from one, i.e. a fully non-stationary process.

2.5.5 Long Memory Processes and Volatility

A final application of long memory processes was concerned with the volatility of asset prices, where the work by Ding, *et al.* (1993), amongst other, provided an additional stylised fact for asset pricing. Following along this train of thought, Baillie, *et al.* (1996b) applied a Fractionally Integrated Generalised Autoregressive Heteroscedasticity (FIGARCH) model to exchange rates; Bollerslev and Mikkelsen (1996) applied a Fractionally Integrated Exponential Generalised Autoregressive Conditional Heteroscedasticity (FIEGARCH) model to stock prices; and Breidt, *et al.* (1993) and Crato and de Lima (1994) found evidence of long memory stochastic volatility in stock prices and exchange rates, respectively.

2.5.6 Concluding Comments

As one can see there has been a lot of interest in long memory processes in financial markets. The aim of the research is to introduce this concept to a different type of financial market, i.e. the shipping market, where the underlying is a service and not an asset. The presence of long memory can be defined as the persistence of the observed autocorrelations, where Baillie (1996) argued that "the extent of the persistence is consistent with an essentially stationary process, but where the autocorrelations take far longer to decay than the exponential rate associated with the ARMA class."

Classical maritime economic theory suggests that freight rates follow a mean reverting process, however, as Koekebakker, *et al.* (2006) pointed out, "the persistency of the spot freight rate process is caused by the fact that the supply cannot generally react to changes in demand with sufficient speed and magnitude to eliminate all demand shocks that bring the freight rate away from levels that yield a normal return to investment…". Koekebakker, e*t al.* gave two main reasons for this, i.e. that there is a lag between the ordering and delivery of new vessels; and, that from the theory of investment under uncertainty it is not always optimal for an investor to respond to a positive demand shock as investors would typically require an option premium to react, where freight rates would exceed long-term average costs. For these reasons, the research believes that freight rates may in fact follow neither a stationary process, where this was suggested by Adland and Cullinane (2006), amongst others; nor a fully non-stationary process, as was suggested by Berg-Andreassen (1996), amongst others, but may instead follow something in between, i.e. the argument freight rates follow a fractionally integrated process.

2.6 The Impact of Conditional Higher Moments

Having examined the literature regarding the characteristics of the first- and second-moments of freight rates in the shipping markets in the previous sections, this section moves on to examine the concept of conditional higher moments and its previous application in financial markets. As this is a relatively new concept to the financial markets, there is no literature on the topic in the shipping markets; for this reason, this literature review focuses on the impact of conditional skewness and kurtosis on other financial markets. Bearing this in mind, focuses on three seminal papers, i.e. those by Harvey and Siddique (1999), Brooks, *et al.* (2005) and León, *et al.* (2005), as well as looking at the practical application of these concepts by Bali, *et al.* (2008).

2.6.1 Harvey and Siddique (1999) and Conditional Skewness

Harvey and Siddique (1999) provided the first discussion regarding the estimation of conditional skewness. In this paper, the authors defined negative skewness as the phenomenon where, after the returns had been standardised by subtracting the mean, any negative returns of a given magnitude would have had a higher probability of occurring than positive returns of the same magnitude, or vice versa. Following this, Harvey and Siddique argued that by modelling the conditional skewness, one is better able to understand the performance of financial assets with skewed return distributions, whereas conventional models assume a normal distribution of asset returns. The authors highlighted the fact that this concept would be of particular use for pricing options, where it is well-known that the distributions of the returns tend to be negatively skewed.

In order to explore this phenomenon, Harvey and Siddique (1999) introduced, for the first time, the Generalised Autoregressive Conditional Heteroscedasticity with Skewness (GARCHS) model, where:

$$h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 \varepsilon_{t-1}^2 \qquad (2.41)$$

$$s_t = \gamma_0 + \gamma_1 s_{t-1} + \gamma_2 \varepsilon_{t-1}^2 \qquad (2.42)$$

One should note that in Expressions (2.41) and (2.42), above, $h_t = \mathrm{Var}_{t-1}\left(r_{M;t}\right)$ and $s_t = \mathrm{Skew}_{t-1}\left(r_{M;t}\right)$. Harvey and Siddique then tested the performance of this $\mathrm{GARCHS}(1;1;1)$ model and compared the results to those for the Generalised Conditional Heteroscedasticity in Mean (GARCH-M) and Exponential Generalised Conditional Heteroscedasticity in Mean (EGARCH-M) models, using $\mathrm{GARCH\text{-}M}(1;1)$ and $\mathrm{EGARCH\text{-}M}(1;1;1)$ specifications. This was done using returns from the S&P 400, DAX 30 and Nikkei 225 stock indices. The authors found that any asymmetry in the variance disappeared once the inclusion of conditional skewness was accounted for, thereby providing a better fit to the data series.

A limitation of Harvey and Siddique (1999)'s approach is that the model does not account for the presence of conditional kurtosis. One could therefore argue, as was the case in Brooks, *et al.* (2005) and León, *et al.* (2005), that a vital piece of the risk picture has been omitted and that while one has certainly improved on the assumption of normality, they do not have a understanding of the true risk position.

2.6.2 Brooks, *et al.* (2005) and Conditional Kurtosis

Brooks, *et al.* (2005) took a different approach to that of Harvey and Siddique (1999), where instead of focussing on conditional skewness, they instead focused on the issue of conditional kurtosis. They argued that the fact that assets returns are leptokurtic implies that extreme market movements, in either direction, will occur with greater frequency, thus leading to a systematic underestimation of the true riskiness of a portfolio. Brooks, *et al.* (2005) went further to propose that by modelling this conditional kurtosis, one should be able to better understand the distribution of asset returns and ensure that the portfolio construction is such that the risk structure will be optimal.

In order to examine this characteristic of the data, Brooks, *et al.* (2005) introduced the Generalised Autoregressive Conditional Heteroscedasticity and Kurtosis (GARCHK) model to allow for the estimation of this conditional kurtosis.

In Brooks, *et al.* (2005)'s GARCHK model:

$$r_t = \gamma_0 + \varepsilon_t^* \qquad (2.43)$$

$$\varepsilon_t^* = \lambda_t \varepsilon_t \;\; ; \;\; \varepsilon_t \sim t_{v_t} \qquad (2.44)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^{*\,2} + \alpha_2 h_{t-1} \qquad (2.45)$$

$$k_t = \beta_0 + \beta_1 \frac{\varepsilon_{t-1}^{*\,4}}{h_{t-1}^2} + \beta_2 k_{t-1} \qquad (2.46)$$

$$v_t = \frac{2(2k_t - 3)}{k_t - 3} \qquad (2.47)$$

$$\lambda_t = \left(\frac{h_t (v_{t-2})}{v_t} \right)^{\frac{1}{2}} \qquad (2.48)$$

In Expressions (2.43) to (2.48), above, $k_t = \mu_{4;t} \big/ h_t^2$, h_t and k_t are the conditional variance and kurtosis, respectively; while v_t denotes the degrees of freedom. Furthermore, one should note that:

$$\mu_{2;t} = \lambda_t^2 \frac{v_t}{v_t - 2} \qquad (2.49)$$

$$\mu_{4;t} = \lambda_t^4 \frac{3v_t^2}{(v_t - 2)(v_t - 4)} \qquad (2.50)$$

Having outlined this model, Brooks, *et al.* (2005) then estimated this GARCHK model using returns from the S&P 500 and FTSE 100 stock and the US and UK bond indices, and then tested the performance of these models using diagnostic tests. The results obtained indicated the strong presence GARCH style dependence in conditional kurtosis. In addition to this, the moment based specification tests indicated that there still remained some features within the data which had not been captured; however, they argued that since other studies had found similar results, this did not indicate a problem.

As was the case for Harvey and Siddique (1999)'s approach, above, a limitation of this methodology is that the model does not account for the joint presence of conditional skewness and kurtosis, and in fact ignores the presence of conditional skewness. One could therefore argue, as was the case in the papers by Harvey and

Siddique (1999) and León, *et al.* (2005), that a vital piece of the risk picture has been omitted and that while one has certainly improved on the assumption of normality, they do not have a understanding of the true risk position.

2.6.3 Leon, *et al.* and Joint Conditional Skewness and Kurtosis

León, *et al.* (2005) extended the concepts introduced by Harvey and Siddique (1999) and Brooks, *et al.* (2005) by combining the concepts of conditional skewness and kurtosis into one model. Furthermore, León, *et al.* (2005) argued that by using a Gram-Charlier series expansion of the normal density function, it became easier to estimate the respective likelihood function, thus making the model more easily applicable, in addition to the fact that this enabled one to account for both time-varying skewness and kurtosis. The authors proposed that this application would be useful should one wish to estimate the volatility, skewness and kurtosis, where these are unknown parameters in option pricing models, which account for non-central skewness and kurtosis.

In order to do this León, *et al.* (2005) introduced the Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis (GARCHSK) model, as well as the Non-Linear Asymmetric Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis (NAGARCHSK) model. The authors specified the NAGARCHSK model, which would nest the GARCHSK model when $\beta_3 = 0$, as follows:

$$r_t = E_{t-1}\left(r_t\right) + \varepsilon_t \;\; ; \;\; \varepsilon_t \sim N\left(0 \; ; \; \sigma_\varepsilon^2\right) \qquad (2.51)$$

$$\varepsilon_t = h^{1/2}\eta_t \;\; ; \;\; \eta_t \sim N\left(0 \; ; \; 1\right) \;\; ; \;\; \varepsilon_t \middle| I_{t-1} \sim \left(0 \; ; \; h_t\right) \qquad (2.52)$$

$$h_t = \beta_0 + \beta_1\left(\varepsilon_{t-1} + \beta_3 h_{t-1}^{1/2}\right)^2 + \beta_2 h_{t-1} \qquad (2.53)$$

$$s_t = \gamma_0 + \gamma_1\eta_{t-1}^3 + \gamma_2 s_{t-1} \qquad (2.54)$$

$$k_t = \delta_0 + \delta_1\eta_{t-1}^4 + \delta_2 k_{t-1} \qquad (2.55)$$

León, *et al.* (2005) noted that, in Expressions (2.51) to (2.55), above, $E_{t-1}\left(\bullet\right)$ denotes the conditional expectation on an information set till period $t-1$, which in turn was denoted as I_{t-1} in Expression (2.52). Furthermore, Leon *et al.* established that

$E_{t-1}(\eta_t)=0$, $E_{t-1}(\eta_t^2)=1$, $E_{t-1}(\eta_t^3)=s_t$ and $E_{t-1}(\eta_t^4)=k_t$, where both s_t and k_t are driven by a GARCH (1;1) structure. Therefore it could be stated that s_t and k_t represented the skewness and kurtosis, respectively, which corresponded to the conditional distribution of the standardised return $\eta_t = \varepsilon_t h^{1/2}$.

León, *et al.* (2005) then tested the performance of the GARCHSK and NAGARCHSK models using returns from the British Pound, Japanese Yen and German Mark versus the US Dollar exchange rates, as well as returns from the S&P 500, NASDAQ 100, DAX 30, IBEX 35 and MEXBOL stock indices. Having estimated these models, the authors then compared the results to those for the standard Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model and standard Non-Linear Asymmetric Generalised Autoregressive Conditional Heteroscedasticity (NAGARCH) models, respectively. This was done on the basis of a likelihood ratio test, the properties of the conditional variances, and the in-sample predictive ability, where the latter was evaluated on the basis of the median absolute error and median percentage absolute error of the respective forecasts. The results obtained indicated the significant presence of conditional skewness and kurtosis, and that the specifications allowing for time-varying skewness and kurtosis outperformed those with constant third and fourth moments.

As León, *et al.* (2005) appears to have addressed the main criticism of the GARCHS and GARCHK models, which were outlined in Harvey and Siddique (1999) and Brooks, *et al.* (2005), respectively, namely that the other two approaches did not take into account joint skewness and kurtosis, thereby potentially underestimating the true risk in the market. For this reason, this thesis applies this approach to examining these same concepts in the shipping market.

2.6.4 Bali, *et al.* (2008) and Value-at-Risk Estimation

The final paper examined in this section is Bali, *et al.* (2008), which provided a practical application of conditional skewness and kurtosis, in terms of how it could be applied to the concept of Value-at-Risk (VaR). The authors argued that the fact that the distribution of asset returns are generally skewed, fat-tailed and peaked around the mean, implies that the traditional value-at-risk methodology, which assumes a normal distribution for these, would result in an underestimation of the true VaR and hence result in an underestimation of the risk involved.

Using returns from the CRSP value-weighted index, Bali, *et al.* (2008) estimated nine different types of Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models with constant and then with autoregressive skewness and kurtosis parameters, using the skewed generalised *t*-distribution (SGT), first proposed by Theodossiou (1998). Having estimated the models and calculated the respective VaR for each, Bali, *et al.* found that the conditional SGT-GARCH models with time-varying skewness and kurtosis outperformed those with constant skewness and kurtosis. The authors continued to discuss the advantages of the SGT distribution over other distributions; however, this is not strictly relevant to this thesis and therefore will not be discussed further.

A limitation of this approach is that one must assume the SGT distribution, an assumption which in itself may be flawed, in addition to the fact that the use of this distribution makes the likelihood-function used to estimate the models computationally complicated, a fact discussed by León, *et al.* (2005). The advantage of approach outlined by Leon, *et al.* is that the use of the Gram-Charlier expansion solves this computational problem, thereby making the concepts of time-varying skewness and kurtosis easier to implement.

2.6.5 Concluding Comments

One can therefore conclude that there is evidence in support of the view that incorporating time-varying skewness and kurtosis of an asset returns series, as opposed to just assuming constant skewness and kurtosis, does have significant benefits in terms of reducing risk and correctly pricing assets. This thesis aims to contribute to the literature by giving it a practical dimension in a market in which the underlying asset is a service through the estimation of these models in the shipping freight market context.

2.7 Forecasting Levels and Volatility in the Shipping Context

Having examined the theoretical foundations and empirical considerations for modelling shipping freight rates, this section reviews the literature on the practical applications of these estimation procedures in the shipping freight markets. In this context, two seminal papers are considered, namely Batchelor, *et al.* (2007) and Angelidis and Skiadopoulos (2008), with the ultimate aim of outlining the benefits of forecasting freight rate levels and volatility, as well as outlining the possible limitations of past research on these topics.

2.7.1 Batchelor, et al. (2007) and Forecasting Freight Rate Levels

Batchelor, *et al.* (2007) extended the earlier work by Cullinane (1992), amongst others, by testing the performance of alternative uni-variate and bi-variate linear time-series models in terms of generating short-term forecasts of spot freight rates, and corresponding rates fixed in the Forward Freight Agreement (FFA) market. In this regard, Batchelor, *et al.* (2007) used both daily spot and FFA prices on Panamax Atlantic routes for the period between 16 January 1997 and 31 July 2000, and Panamax Pacific routes for the period between 16 January 1997 and 30 April 2001.

Before going any further, Batchelor, *et al.* (2007) first tested the data series for stationarity using Augmented Dickey-Fuller and Phillips-Perron and KPSS unit root tests, outlined in Dickey and Fuller (1981) and Phillips and Perron (1988), respectively. The results of these tests indicated that both spot and FFA prices followed an $I(1)$ process. This being said, Batchelor, *et al.* (2007) noted that the Augmented-Dickey Fuller and Phillips-Perron tests had been previously criticised for their lack of power in rejecting the null hypothesis of a unit root when it is false. In order to address this issue, the authors also implemented the KPSS test, developed by Kwiatkowski, *et al.* (1992), which confirmed the previous results. Having established this, Batchelor, *et al.* (2007) considered four different time series models in order to identify the model that provides the most accurate short-term forecasts of spot and FFA prices in the market. These were an Autoregressive Integrated Moving Average (ARIMA) model, developed by Box and Jenkins (1970); Sims (1980)'s Vector

Autoregression (VAR) model; a Vector Error Correction Model (VECM), outlined in Engle and Granger (1987); and a restricted VECM.

Batchelor, *et al.* (2007) initially estimated the models over the in-sample period between 16 January 1997 and 30 June 1998, where it was found that the VECM models provided the best in-sample fit for the data set. Having established this, the authors then generated independent N -period ahead forecasts over the out-of sample period between 1 July 1998 and 31 July 2000, for the Atlantic routes, and 1 July 1998 and 30 April 2001, for the Pacific routes. Batchelor, *et al.* then assessed the forecast accuracy of these models using the conventional Root Mean Squared Error (RMSE) metric, where it was found that while the random walk outperformed the ARIMA models, this was not the case for the VECM models. To provide a better picture of the performance of the models, the authors then tested the significance of any outperformance by applying the Diebold-Mariano statistic, outlined in Diebold and Mariano (1995), to test the null hypothesis that the RMSE metrics from two competing models were equal. On this basis, Batchelor, *et al.* (2007) concluded that, in terms of the out-of-sample forecasting performance, VECM models were not helpful in predicting forward rate behaviour, but do help predict spot rates, a finding that the authors argued was more consistent with market efficiency.

A limitation that may apply to this paper is that there are doubts regarding the efficiency of the tests used to establish the non-stationarity of the data series. Furthermore, classical maritime economic theory, such as outlined by Zannetos (1966), amongst others, as well as the partial equilibrium models proposed by Adland and Cullinane (2006), amongst other, suggest that freight rates follow a mean reversion process. This thesis adds another dimension to this debate by proposing that the delay in the mean reversion process, outlined by Adland and Cullinane (2006) and Koekebakker, *et al.* (2006), suggests the freight rate are instead fractionally integrated, and therefore tests the forecasting performance of Autoregressive Fractionally Integrated Moving Average (ARFIMA) models, against alternative models, which assume stationarity or non-stationarity, respectively.

2.7.2 Angelidis and Skiadopoulos (2008) and Forecasting Volatility

Angelidis and Skiadopoulos (2008) noted that the fluctuation of shipping freight rates, i.e. the freight rate risk, is an important source of market risk for all participants in the freight markets. In order to gain a better understanding of this risk, the authors examined which volatility models provided the best forecast of the true risk in the market by calculated the respective Value-at-Risk (VaR) for Moving Average (MA), Exponentially Weighted Moving Average (EWMA) and different types of Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models as well as standard non-parametric estimation techniques, namely, Historical Simulation (HS) and Filtered Historical Simulation (FHS).

Angelidis and Skiadopoulos (2008) argued that an accurate calculation of the VaR in the freight rate markets is important for at least three reasons. The first of the reasons given was that this would enable the market participants to quantify the level of the freight rate risk to which they are exposed to so as to develop effective hedging schemes to mitigate this risk. The second reason was that an understanding of the possible extreme fluctuations of freight rates is important since freights are currently viewed as an alternative investment by many hedge funds, which are now beginning to expand their presence in the market. The final reason given was that the VaR could be used to set the margin requirements in the freight exchange derivatives market so as to ensure it grows even further. The authors highlighted the fact that the ultimate aim of the research was to shed light on what method should be preferred to calculate VaR in the freight markets; where previous literature in other markets provided mixed results as to the best method, where this depends on the data set, the confidence level, and the period under scrutiny.

In order to evaluate the performance of the models in estimating the true VaR, Angelidis and Skiadopoulos (2008) followed the methodology outlined in previous papers by Lopez (1998) and Sarma, *et al.* (2003) and conducted the backtesting in two stages. The first of these stages involved the use of three formal statistical tests, which were outlined in Christoffersen (1998), in order to verify the accuracy of the VaR estimates; while the second stage involved constructing an appropriate loss function so as to choose the best VaR method among the ones that pass the statistical backtesting criteria from the first stage. Angelidis and Skiadopoulos (2008) constructed the loss function in the second stage using the Expected Shortfall (ES).

In order to perform this whole procedure, Angelidis and Skiadopoulos (2008) used daily price data from four indices published by the Baltic Exchange for the period between 1 March 1999 and 30 October 2006. Having estimated the volatility of the underlying price series and calculated the respective VaRs and ES, noted that only the estimates for the tanker index passed any of the statistical tests. In order to compare the loss functions for the different models, the authors implemented the Modified Diebold-Mariano (MDM) test, proposed by Harvey, *et al.* (1997), where these results indicated that in almost all cases, the simplest non-parametric models were preferred.

There appear to be a number of limitations in this paper, in that the results from only one of the four indices used passed the first stage of the backtesting process. Furthermore, Angelidis and Skiadopoulos (2008) do not appear to take the fact that freight rates may exhibit conditional skewness and kurtosis. A final limitation may be that the authors did not consider whether the underlying volatility exhibited any form persistence. This thesis aims to address these issues by using spot freight rate data from routes themselves, rather than indices, as well as incorporating models which account for both persistence and conditional skewness and kurtosis.

2.7.3 Concluding Comments

It becomes apparent from this research that it is definitely important for participants in the shipping markets to have an understanding of the future behaviour of freight rates, both in terms of levels and volatility. This being said, there is room for improvement on the previous research in that the assumption as to the degree of stationarity in the underlying spot freight rates was based on potentially limited test, a fact outlined by Schwert (1989), amongst others. Furthermore, papers by Adland and Cullinane (2006) and Koekebakker, *et al.* (2006) both outline that any mean reversion that may occur will occur with a delay due to the supply and demand dynamics within the shipping markets. This thesis addresses these issues by introducing the concept of fractional integration, thereby adding another dimension to the debate.

When looking at the research on forecasting the volatility of freight rates, this research appears to have been limited by the data selection process, as well as the fact that volatility in the freight markets may exhibit some persistence. Another limitation may be that the underlying spot freight rates may also exhibit conditional skewness and kurtosis. This thesis addresses these issues by using spot freight rate data from routes themselves, rather than indices, as well as incorporating models which account for both persistence and conditional skewness and kurtosis.

2.8 Summary and Contribution

This chapter reviews the somewhat limited set of literature on the structure of freight rates in the shipping market, as well as introduces some new concepts from other financial markets. This literature review began by examining the existing research on the structure of the first moment of these freight rates, where a conflict has arisen as to the exact degree of integration of these spot freight rate levels. Traditional maritime economic theory, such as was outlined by Hawdon (1978) and Beenstock and Vergottis (1989), amongst others, examined the structure of the demand and supply functions in this market, arguing that the fact that demand and supply continually re-adjust, over the long-run, to an equilibrium level would imply that freight rates are mean reverting. Following this, with the development on new test for unit roots and the concept of cointegration, the trend in the literature, for example papers by Berg-Andreassen (1996) and Glen and Rogers (1997), where the result of these tests indicated that instead freight rates followed a non-stationary process. As these tests came under criticism, for example in a paper by Schwert (1989), more recent research proposed the implementation of partial equilibrium models, where once again freight rates were assumed to be stationary, where the presence of statistically significant unit roots were attributed to the weakness of the respective unit root tests. This thesis provides an alternate dimension to the ongoing debate by proposing a middle ground in that it argues that freight rates follow a fractionally integrated process. The rationale behind this is that the dynamics of the supply and demand functions in the shipping markets are such that, as supply is fixed in the short-run, due to the delay in the delivery of new capacity, freight rates are capable of exhibiting long memory. This being said, as new tonnage is delivered in the longer-term, freight rate levels will revert to the mean, where this a characteristic of fractionally integrated processes. One should note that, as far as the author is aware, this is examined for the first time in a shipping context.

The literature on the second moment of freight rates, i.e. freight rate volatility, once again is flawed in that the models used were based upon the assumptions that the freight rate process was either stationary or non-stationary. As the research sets out to prove, this may not necessarily be the case, and in fact the freight rate process may follow a fractionally integrated process. A second problem with the models used in the papers above is that they ignore the problem of the potential existence long-memory, or persistence, in volatility. Koekebakker, *et al.* (2006) briefly mentioned the problem of persistence in levels of freight rates, discussed in the critique of models of levels of freight rates above, while Kavussanos (1997) also mentions this issue with regard to the volatility of freight rates. In order to address these issues, and thereby add a new dimension to the existing literature, this thesis examines, using the same rationale, in terms of the supply and demand dynamics in the market, as for the first moment the concept of persistence in freight rate volatility by introducing, as far as the authors are aware, for the first time, the concept of the Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) model, developed by Baillie, *et al.* (1996a), to the shipping literature with the ultimate aim of determining whether this provides a better understanding of the structure of volatility, and therefore risk, in this market.

The final section of literature, where this examined the third and fourth moments of a price series, introduced the concepts of conditional skewness and kurtosis to the financial literature. The rationale behind this was that the assumption of constant skewness and kurtosis resulted in a misunderstanding as to the true extent of the risk exposure of market participants. This thesis examines these issues for the first time in the shipping literature, and further contributes to the literature in that this is the first time, to the best of the author's knowledge, that these concepts are being applied in a market in which the underlying good provided is a service, thereby adding a new dimension to the debate.

Having outlined the three main hypotheses in Chapter 1, and thoroughly reviewed in the literature on these issues in this chapter, this thesis continues by outlining the methodologies and data to be used to test these hypotheses in Chapters 3 and 4, respectively. Following this, each hypothesis will be individually tested in Chapters 6, 7 and 8 of the thesis.

3 Methodology

3.1 *Introduction*

Traditional methods, which address the investment timing and decision making processes generally involve some form of forecasting with respect to the underlying data series. The ultimate aim of applying this in the shipping markets is to determine the future direction of freight rates, and the risk, using standard time series models. A wide variety of methodologies have been applied to the shipping markets with this aim in mind of providing a better understanding of the underlying freights and the risks associated with participating in these markets, where these are discussed in more detail in Chapter 2 of the thesis. The aim of this section is to outline and fully discuss the methodologies that will be used in this thesis, highlighting where these will extend on the existing literature, thereby providing a clearer understanding as to the true structure of the different moments of freight rates. This is essential for all participants in the shipping market as freight rates form the basis for all investment and planning decisions, both for the direct participants in the market, i.e. the ship-owners and charterers, and for auxiliary parties, such as banks, hedge funds and maritime lawyers.

The chapter begins by outlining the various methodologies used in the thesis to analyse the structure of the freight rate levels in Section 3.1. In Section 3.2, the chapter continues to highlight how one can analyse the structure of the volatility of the freight rate process, while Section 3.3 analyses the higher the moments, i.e. the skewness and kurtosis, of the freight rates by introducing the concept of conditional skewness and kurtosis and the methodology for analysing these. Section 3.4 concludes by outlining the methodologies that will be used to analyse the process of forecasting the different moments of freight rates.

3.2 *Determining the Structure of the First Moment*

As mentioned above, the ultimate aim of any form of research in finance is to be able to forecast the future direction of the underlying series. This section presents the methodology for understanding the processes defining the first moment of freight rates, i.e. the spot freight rate levels. To date, the shipping literature has mainly focussed on the ideas that freight rates are either stationary or non-stationary, hence, this section outlines one model of freight rate levels for each scenario, i.e. the Autoregressive Moving Average (ARMA) and the Autoregressive Integrated Moving Average (ARIMA). The ARMA model would conform to the general and partial equilibrium theories outlined in Chapter 2, as this model assumes that the underlying freight rates are stationary. In contrast, the ARIMA model would conform to the non-stationary theories, in that this model assumes that the underlying freight rates are non-stationary.

As discussed in Chapter 2, the shape of the supply curve in the shipping market, in addition to supply being fixed in the short-term, but not in the long-term, imply that freight rates should be fractionally integrated. To this end, this section also introduces the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model to the shipping literature, where, to the best of the author's knowledge, this is for the first time in this literature.

3.2.1 A Brief Discussion on AR and MA Processes

Following the structure in Greene (2003), before looking at the structure of the ARMA and ARIMA models, one first has to define the Autoregressive (AR) and Moving Average (MA) components of the models. Starting with the AR process, assume that one is given the following model:

$$y_t = \mu + \phi y_{t-1} + \varepsilon_t \qquad (3.1)$$

In the model in Expression (3.1), the variable y_t is considered to be autoregressive of order one, i.e. AR(1), as, under certain assumptions regarding the model, the following expression will hold true:

$$E\left[y_t \middle| y_{t-1}\right] = \mu + \phi y_{t-1} \qquad (3.2)$$

Following this, the model in Expression (3.1) can generalised to an $\mathrm{AR}(p)$ process by rewriting the model as follows:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} \qquad (3.3)$$

In Expression (3.3), above, p denotes the number of autoregressive, or lagged dependent variable, terms. Alternatively, one could rewrite this model more compactly by using the lag operator, hence:

$$C(L)y_t = \mu + \varepsilon_t \qquad (3.4)$$

In Expression (3.4), above, $C(L)$ denotes the polynomials in the lag operator, where $C(L) = (1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p)$ and denotes the number of autoregressive terms in the model. Moving on from this definition of the AR process, one can now look at the MA process.

Consider the following MA(1) specification:

$$y_t = \mu + \varepsilon_t - \theta\varepsilon_{t-1} \qquad (3.5)$$

The model in Expression (3.5) can be generalised to a MA(q) process in a manner similar to that for the $\mathrm{AR}(p)$ process, i.e.:

$$y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \ldots - \theta_q\varepsilon_{t-q} \qquad (3.6)$$

One should note that in Expression (3.6), above, q denotes the number of moving average, or lagged error, terms. Once again, one can rewrite the MA process more compactly using the lag operator, where:

$$y_t = \mu + \varepsilon_t - D(L)\varepsilon_t \qquad (3.7)$$

In Expression (3.7), above, $D(L)$ denotes the polynomials in the lag operator, where $D(L) = (\theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q)$ and q denotes the number of moving average terms in the model.

The discussion continues in the next section to establish how one can combine these two processes to form ARMA and ARIMA models of the appropriate orders.

3.2.2 The ARMA and ARIMA Models

Having examined the AR and MA process, one can combine these to form an $\mathrm{ARMA}(p;q)$ model, which has p autoregressive terms and q moving average terms.

The $\text{ARMA}(p;q)$ model, in its most general form, is thus specified as follows:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \ldots - \theta_q \varepsilon_{t-q} \qquad (3.8)$$

As with the AR and MA process, this model can be rewritten in a more concise form using the lag operator, hence:

$$C(L) y_t = \mu + \varepsilon_t - D(L) \varepsilon_t \qquad (3.9)$$

In Expression (3.9), above, $C(L)$ and $D(L)$ denote the polynomials in the lag operators, where these are defined in the notes for Expressions (3.4) and (3.7), respectively. One should note that one of the underlying assumptions for the ARMA model is that the data series being modelled is stationary.[12] Therefore, should one apply the ARMA model to a non-stationary data series, the results would be spurious. The reason for this is that is one is using a non-stationary data series, when one regresses one variable on another, and the two variables are related over time, then one could find that the model has a high measure of fit, even if the two variables are completely unrelated.

In order to rectify this issue, one could either make the data series stationary by taking the appropriate number of first differences until the data series becomes stationary, or more simply, one could run an ARIMA model. The ARIMA model is similar to an ARMA model in that, as with the ARMA model, it has *p* autoregressive terms and *q* moving average terms. However, where the ARIMA model differs is that it has an extra component, denoted *d*, where this illustrates the number of times the series would have to differences in order to make it stationary, also known as the order of integration, denoted $I(d)$. One should note that for the ARIMA model, *d* must be an integer. The $\text{ARIMA}(p;q)$ model can therefore be specified, in its most general form, as follows:

$$\Delta^d y_t = \mu + \phi_1 \Delta^d y_{t-1} + \ldots + \phi_p \Delta^d y_{t-p} + \varepsilon_t \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q} \qquad (3.10)$$

As with the previous models, this model can be written more compactly through the use of lag operators, where:

$$C(L)\left[(1-L)^d y_t\right] = \mu + \varepsilon_t - D(L) \qquad (3.11)$$

[12] One should note that stationarity is defined in Section C1 of Appendix C.

In Expression (3.11), above, $C(L)$ and $D(L)$ denote the polynomials in the lag operators, where these are defined in the notes for Expressions (3.4) and (3.7), respectively, and $(1-L)^d y_t = \Delta^d y_t$ is the d^{th} difference of y_t.

As far as the application of the ARMA and ARIMA models is concerned, both these models have found widespread use throughout the literature, with ARIMA models being used in a number of papers, such as Kavussanos and Nomikos (1999). The following section introduces to the concept of fractional integration and its application to modelling spot freights rates levels.

3.2.3 The Definition of Long Memory

Before going any further, one needs to define what is meant by long memory processes. Hurst (1951) first introduced the concept of long memory, in the area of hydrology, when seeking to understand the persistence of streamflow data and the design of reservoirs. A number of definitions exist for the property of 'long memory', such as those by Rosenblatt (1956) and Taqqu (1975), however, for simplicity, this thesis focuses only on the definition given by McLeod and Hipel (1978).

Given a discrete time series process, y_t, this process is defined as possessing long memory if the following quantity in non-infinite, where the quantity is:

$$\lim_{n\to\infty} \sum_{j=-n}^{n} \left|\rho_j\right| \qquad (3.12)$$

In Expression (3.12), above, ρ_j denotes the autocorrelation function at lag *j*. One should note that this time series process, y_t, is said to be a fractionally integrated $I(d)$ process, where $-0.5 < d < 1$, if:

$$(1-L)^d = \varepsilon_t \qquad (3.13)$$

In Expression (3.13), above, L denotes the lag operator, and ε_t is a stationary and ergodic process, where the spectrum is bounded and positively bounded at all frequencies. Where $0 < d < 0.5$, and ε_t is $I(0)$ and therefore covariance stationary, as the autocorrelations for the process are all positive and decay hyperbolically, one can argue that the process possesses long memory in that it satisfies the condition in Expression (3.12). In contrast, where $-0.5 < d < 0$, the sum of the absolute values of the autocorrelations tends to a constant, hence the process has short memory according to this condition. The final important case to note is where the $d = 0$, and the process is therefore a stationary and invertible ARMA process. In this case, the autocorrelations are geometrically bound, in that $|\rho_k| \leq cm^{-k}$, for large values of k and where $0 < m < 1$, therefore the process is a short memory process as it does not fulfil the condition in Expression (3.12).

3.2.4 The ARFIMA Model

As mentioned in Section 3.1 above, the first issue encountered when attempting to address the investment timing and decision making processes is how to forecast the future direction of prices. Traditionally, investors have used the ARMA and ARIMA processes to fulfil this purpose; however, the suggestion that freight rate processes are mean reverting would cause this method of forecasting to be inaccurate. In order to correct this, this research proposes, for the first time in the shipping literature, the use of ARFIMA processes to fulfil any forecasting needs.

The existence of cycles in the freight rate markets suggest that the series are long-term dependent, with persistence being exhibited in the mean return generating process. This suggests that one should use an ARFIMA$(p;d;q)$ model, where $-0.5 < d < 1$, rather than an ARIMA$(p;d;q)$ model, where $d \geq 1$ and an integer. One should note that d here is the order of integration, or, in other words, number of differences one has to take, in order to make the process stationary and hence enable one to use a normal ARMA model.

Characterising the returns generating process is a crucial element of asset and risk management, the asset pricing process and correct portfolio allocation. Contrary to the random walk hypothesis, i.e. that returns follow a random walk and therefore cannot be predicted, several studies, such as Lo (1991), find that there is evidence of long-horizon predictability in stock returns. Lo argued that such evidence may be symptomatic of a long-range dependent, long-memory, component in stock market prices, allowing asset returns to exhibit significant autocorrelation between distant observations. There is no reason that this should not apply to the shipping market, where freight rate processes display many of the features of stock returns.

In order to account for this long memory, this research introduces the ARFIMA, for the first time, to the shipping literature. The ARFIMA model, first introduced by Granger and Joyeux (1980), Granger (1980), Granger (1981) and Hosking (1981), parameterises the conditional mean of the returns generating process as an ARFIMA$(p;d;q)$ process, where this is specified as follows:

$$\phi(L)(1-L)^{d}\left(y_{t}-\mu\right)=\theta(L)\varepsilon_{t} \qquad (3.14)$$

In Expression (3.14), above, *d* denotes the fractional differencing parameter, where $-0.5 < d < 1$, *L* denotes the lag operator and $\phi(L)$ and $\theta(L)$ denote the polynomials of the lag operators, where all the roots of $\phi(L)$ and $\theta(L)$ lie outside the unit root circle, and ε_t is white noise. One should note that the Wold decomposition and the autocorrelation coefficients for this process will all exhibit a very slow rate of hyperbolic decay, where the higher the value of *d*, the slower the rate of decay. In addition, one should note that where $-0.5 < d < 0.5$, the process is covariance stationary, and that as long as $d < 1$, the process will exhibit mean reversion, while should $0.5 < d < 1$ then the process would be fractionally integrated by effectively non-stationary in terms of the covariance.

Baillie (1996) highlights the fact that a number of different methodologies have been proposed for estimating the parameters of the ARFIMA model, for example, Geweke and Porter-Hudak (1983) suggest a semi-parametric estimator of the fractional differencing parameter, *d*, in the frequency domain, while Robinson (1990) also suggests a semi-parametric estimator of *d*, but this time it is in the time domain. This being said, this thesis uses the maximum likelihood estimation (MLE) approach,[13] where Sowell (1986) and Sowell (1992) propose an exact MLE of the ARFIMA process with unconditional Normally distributed disturbances, ε_t.

[13] Should one require more information, a much more detailed discussion of the various MLE methodologies is provided by Baillie (1996).

The proposed log-likelihood function is then:

$$\ell = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\log|\Omega| - \frac{1}{2}Y'\Omega^{-1}Y \qquad (3.15)$$

In Expression (3.15), above, $\{\Omega\}_{i;j} = \gamma_{|i-j|}$, where γ denotes the autocovariances of the ARFIMA process, and *Y* represents a *T*-dimensional vector of the observations on the process y_t.[14]

In contradistinction, Whittle (1951) finds that the autocovariance matrix, Ω, from Expression (3.15), may be diagonalised by transforming the vector *Y* in this log-likelihood function into the frequency domain, hence one could approximate the log-likelihood by using the following log-likelihood function:

$$\ell = \sum_{j=1}^{T-1}\log\left[(2\pi) f(\omega_j)\right] + \sum_{j=1}^{T-1}\left[\mathrm{I}_T(\omega_j) \big/ f(\omega_j)\right] \qquad (3.16)$$

In Expression (3.16), above, $\mathrm{I}_T(\omega_j)$ denotes the periodogram evaluated at frequency ω_j, and *T* denotes the sample period.

Fox and Taqqu (1986) provide an alternate frequency domain approximation of the MLE, which is the one used in this thesis, where they numerically minimise the following quantity:

$$\sum_{j=1}^{m}\left\{\mathrm{I}(\omega_j) \big/ f(\omega_j;\theta)\right\} \qquad (3.17)$$

In Expression (3.17), above, $\mathrm{I}(\omega_j)$ denotes the periodogram evaluated at frequency ω_j, and the summation is over *m* frequencies.

ARFIMA models are preferred to the more traditional ARIMA or ARMA models used in the literature as these two other models assume that the price process is stationary and non-stationary, respectively. As mentioned before, this research believes that the price process is neither stationary nor non-stationary, but is in fact a

[14] Sowell (1992)'s full MLE requires the inversion of a $T \times T$ matrix of non-linear functions of the hypergeometric function at each iteration of the maximisation of the likelihood. The method to do this requires that all the roots of the autoregressive polynomial must be distinct, and that the theoretical mean parameter, μ, must be either zero or known. This means that, although it is theoretically appealing, it is very computationally demanding, therefore rendering it undesirable for this research.

fractionally integrated process. This means that, should the process indeed be fractionally integrated, one was to use either an ARIMA or ARMA model, the price process would be misspecified, and one's forecasts would be inaccurate, rendering them useless. This may be the reason for the low levels of accuracy in forecasts traditionally achieved for freight rates.

ARFIMA models have been used to model long memory price processes in a wide range of financial markets as well as well as for their original purpose in the geophysical sciences. Crato and Rothman (1994a) use an $\text{ARFIMA}(0;d;1)$ process to model annual bond yields and find that this process follows a statistically significant fractionally integrated process. Sowell (1992) uses an $\text{ARFIMA}(3;d;2)$ process to model US real GNP; Baillie, *et al.* (1996b) model inflation rates using ARFIMA models; and Baillie and Bollerslev (1994b) use an $\text{ARFIMA}(2;d;0)$ process to model the forward premium in exchange rates. To the author's knowledge, there has been no previous work on the use of ARFIMA models in the shipping freight markets, thereby making this a ripe area for further research.

Having determined the different structures of the ARMA, ARIMA and ARFIMA models, the following section examines the tests to determine whether a data series is fractionally integrated or not.

3.2.5 Tests for Fractional Integration

As standard unit root tests [15] can only distinguish between $I(0)$ stationary and $I(1)$ non-stationary processes, the need arose for a test which would enable one to determine if a process is $I(d)$, where $0<d<1$. Several tests arose, the most commonly used of which is the test proposed by Robinson (1994). This being said, this research uses the LM test first proposed by Nielsen (2005) as it has a number of advantages over other tests in that it is a time domain test, as opposed to a frequency domain test, and can be used for multivariate models. The main objective of this

[15] Three standard unit root tests, namely the Augmented Dickey-Fuller, Phillips-Perron and KPSS tests, proposed by Dickey and Fuller (1981), Phillips and Perron (1988) and Kwiatkowski, *et al.* (1992), respectively, are outlined in detail in Section C2 of Appendix C.

Lagrange multiplier (LM) test is to test whether a time series, where the time series is denoted y_t, is $I(d)$, against the alternate hypothesis that y_t is $I(d+\theta)$, where $\theta \neq 0$. Therefore, if one was to difference the observed time series, this would be equivalent to testing whether $x_t = (1-L)^d y_t$ is $I(0)$ as opposed to $I(\theta)$.

This test may be used as a tool for preliminary data analysis, such as testing for stationarity, etc., where this test may indicate what transformation of the data would be required to make the data stationary. However, this test may also be applied after modelling to ensure that the fractional difference implemented is sufficient to render the process stationary, or $I(0)$, where this is the context in which it shall be used in this thesis. Therefore, this thesis uses this test to determine if the data series are fractionally integrated prior to implementing the ARFIMA models, and then tests the residuals from the AFRIMA models for stationarity using this test after implementing the models.

This test is carried out, in general terms, as follows, where one was to observe a time series $\{y_t\,; t=1;\ldots;n\}$, which is generated by:

$$(1-L)^{d+\theta} y_t = \varepsilon_t \mathrm{I}(t \geq 1) \;\; ; \;\; t = 0 \; ; \;\; \pm 1 \; ; \;\; \pm 2 \; ; \; \ldots \; ; \qquad (3.18)$$

In Expression (3.18), above, $\mathrm{I}(\cdot)$ denotes the indicator function while ε_t is I (0), which means that ε_t is covariance stationary with a spectral density that is bounded away from zero at the origin. One should also note that the process y_t generated by Expression (3.18) is well defined for all values of d and θ. An important point is that the process outlined in this expression allows for a uniform definition, valid for all values of both d and θ, whereas the alternative definition, without the truncation included in the expression, would only be valid for $d+\theta \in (-1/2\,;1/2)$ with partial summation being required in order to generate a process with an order of integration outside this range.

In order to perform the test, one would assume a value of d, which is known a priori, and then test the following hypothesis:

$$H_0 : \theta = 0 \qquad (3.19)$$

This hypothesis is tested against the alternative $H_1 : \theta \neq 0$, where an example of this would be to test the unit root hypothesis, which is done by setting $d = 1$ in Expressions (3.18) and (3.19). It is important to note that the assumption that the value of d is known *a priori,* is made without any loss of generality and that the specification of a particular value of d exactly specifies the null hypothesis since $\theta = 0$ in Expression (3.19).

One should note that the Gaussian log-likelihood function of the model in Expression (3.18) would be:

$$\ell(\theta;\Sigma) = -\frac{n}{2}\ln\left(2\pi|\Sigma|\right) - \frac{1}{2}\sum_{t=1}^{n}(1-L)^{d+\theta} y_t^{'}\Sigma^{-1}(1-L)^{d+\theta} y_t \qquad (3.20)$$

Hence, the score would be:

$$\frac{\partial\ell(\theta;\Sigma)}{\partial\theta}\Big|_{\theta=0;\Sigma=\hat{\Sigma}} = -\sum_{t=1}^{n}\left(\ln(1-L)\right)x_t^{'}\hat{\Sigma}^{-1}x_t = tr\left(\hat{\Sigma}^{-1}S_{10}\right) \qquad (3.21)$$

In Expression (3.21), above, $x_t = (1-L)^d y_t$, $S_{1;0} = \sum_{t=2}^{n} x_{t-1}^{*} x_t^{'}$, $x_{t-1}^{*} = \sum_{j=1}^{t-1} j^{-1} x_{t-j}$, and $\hat{\Sigma} = n^{-1}\sum_{t=1}^{n} x_t x_t^{'}$ is a consistent estimate of $\Sigma = E\left(\varepsilon_t \varepsilon_t^{'}\right)$ under the null hypothesis. In the case of a univariate time series, the score in Expression (3.21), normalised by $\sqrt{n}$, reduces to Tanaka (1999)'s univariate time domain score statistic:

$$s_n = \sqrt{n}\sum_{j=1}^{n-1} j^{-1}\rho(j) \qquad (3.22)$$

In Expression (3.22), above, $\rho(j)$ is the j th order sample autocorrelation with respect to the process y_t.

The Nielsen (2005) test has a number of advantages over similar tests used in the literature. The first of these advantages is that the model can be extended to allow for deterministic terms, a different value of d and θ for each variable. In addition, this test works in almost exactly the same manner as the Tanaka (1999) and Robinson (1994) tests for fractional integration. A further advantage of this test over others is that Nielsen (2005) argues that the proposed test is a time domain test, as opposed to a frequency domain test, where Tanaka (1999) suggests that time domain tests are superior in terms of finite sample properties. One should also note that Nielsen (2005) compared the finite sample properties of the Nielsen test with those of

Breitung and Hassler (2002)'s $\Lambda_0(d)$ test, finding that the Nielsen test had higher finite sample power than the Breitung and Hassler test when testing for fractional integration, where should one wish to examine the issue further, a detailed discussion can be found in Nielsen (2005).

3.3 Determining the Structure of the Second Moment

Having outlined a selection of models to be used to determine the structure of the first moment of freight rates, thereby enabling one to determine the structure of the underlying spot freight rate levels process, the focus now changes to the issue of the risk inherent in the market. In order to be able to correctly gauge the market risk exposure for participants in the shipping market, one must have some idea of the structure of the second moment, or the volatility, of the underlying freight rate process. This section outlines a number of methodologies that will be used in this thesis, with the ultimate goal of being able to reduce this market risk exposure.

This section begins by outlining the standard Autoregressive Conditional Heteroscedasticity (ARCH) and Generalised Autoregressive Conditional Heteroscedasticity (GARCH) methodologies, proposed by Engle (1982) and Bollerslev (1986), respectively, describing these in detail and briefly discussing examples of their application, in both the shipping literature and the broader finance literature as a whole. Following this, this section moves on to discuss the issue of persistence in volatility, introducing the Integrated Generalised Conditional Heteroscedasticity (IGARCH) model, first outlined in Engle and Bollerslev (1986), and then concludes by introducing, for the first time in the shipping literature, the Fractionally Integrated Generalised Conditional Heteroscedasticity (FIGARCH) model, introduced by Baillie, *et al.* (1996a), and discussing the methodology for implementing this model to model the volatility of shipping freight rates.

3.3.1 The ARCH, GARCH and IGARCH Models

In the financial markets, most studies on the volatility of the underlying series focus on the ARCH family of models, where the ARCH model was first introduced by Engle (1982). ARCH models have been used to model volatility in inflation, as in

Coulson and Robins (1985), amongst others; the term structure of interest rates, as illustrated in Engle, *et al.* (1987), amongst many others, and the behaviour of exchange rates, see Domowitz and Hakkio (1985), amongst others; as well as many other markets. The ARCH methodology has also been used in the shipping freight markets, where Kavussanos (1996), Kavussanos (1997) and Kavussanos and Nomikos (2000b), amongst others, use this model to examine the volatility of the underlying freight rates.

The simplest of the ARCH family of models is the simple ARCH model, first proposed by Engle (1982), where the ARCH process, denoted $\{\varepsilon_t\}$ may be specified as follows:

$$\varepsilon_t \equiv z_t \sigma_t \qquad (3.23)$$

One should note that, in Expression (3.23), above, z_t is distributed as a standard normal distribution, i.e. $E_{t-1}(z_t) = 0$ and $\mathrm{VAR}_{t-1}(z_t) = 1$, where $E_{t-1}(\cdot)$ and $VAR_{t-1}(\cdot)$ denote the conditional expectation and variance with respect to the same information set. In addition, σ_t is a positive time-varying and measurable function with respect to the information set at time $t-1$. Therefore, one can state that, by definition, the ARCH, or $\{\varepsilon_t\}$ process is serially uncorrelated, with mean zero; however, the conditional variance of the process, denoted σ_t^2, changes over time.

A restriction of this $\mathrm{ARCH}(q)$ model, is that, in its classical form, the conditional variance is postulated to be a linear function of the lagged squared innovations. This implies that Markovian dependence will only date back q periods, i.e. ε_{t-i}^2, where $i = 1;\ldots;2$. Bollerslev (1986) solves this problem by allowing for a more flexible lag structure in the $\mathrm{GARCH}(p;q)$ model, where this is formally defined as follows:

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2 \qquad (3.24)$$

In Expression (3.24), above, L denotes the lag operator, hence one can see that $\alpha(L) \equiv \alpha_1 L + \alpha_2 L^2 + \ldots + \alpha_q L^q$ and $\beta(L) \equiv \beta_1 L + \beta_2 L^2 + \ldots + \beta_p L^p$. In order to ensure that the $\{\varepsilon_t\}$ process is stable and covariance stationary, all the roots of

$\left[1-\alpha(L)-\beta(L)\right]$ and $\left[1-\beta(L)\right]$ are constrained such that they lie outside the unit circle. This stationarity condition implies that the effect of the past squared innovations on the current conditional variance will decay exponentially with the lag length, i.e. the impact of past shocks will decay exponentially over time.

Engle and Bollerslev (1986) extend this model by arguing that should the autoregressive lag polynomial, $1-\alpha(L)-\beta(L)$, contain a unit root, then the $\text{GARCH}(p;q)$ process will be integrated in variance. Therefore, the corresponding $\text{IGARCH}(p;q)$ model will be given, succinctly, by:

$$\phi(L)(1-L)\varepsilon_t^2 = \omega + \left[1-\beta(L)\right]v_t \qquad (3.25)$$

In Expression (3.25), above, $\phi(L) \equiv \left[1-\alpha(L)-\beta(L)\right](1-L)^{-1}$, where L, $\alpha(L)$ and $\beta(L)$ are defined above, is of order $m-1$. It is worth noting that, while shocks to the conditional variance for the $\text{GARCH}(p;q)$ model defined in Expression (3.24) decay exponentially, these will persist indefinitely for this $\text{IGARCH}(p;q)$ model.

The following section examines the concept of fractional integrated variance, using the FIGARCH model, and discusses how this model would be implemented.

3.3.2 The FIGARCH Model

As mentioned in Sections 3.2.3, above, the concept of long memory was first introduced by Hurst (1951). Following this, Granger (1980), Granger (1981), Granger and Joyeux (1980) and Hosking (1981) proposed that this concept could be implemented, in terms of modelling the levels of the underlying series, through the use of the ARFIMA model. This section examines the implementation of this concept in terms of modelling the variance of the underlying data series, through the use of Baillie, *et al.* (1996a)'s FIGARCH model. This means, that in contrast to the $\text{GARCH}(p;q)$ model, where shocks dissipate exponentially, and the $\text{IGARCH}(p;q)$ model, where shocks persist indefinitely, the response of the conditional variance to past shocks decays at a slower, hyperbolic rate.

Baillie, *et al.* (1996a) extend the ARCH literature by proposing that, in addition to being able $\text{GARCH}(p;q)$ processes, which are integrated in variance, one can model $\text{GARCH}(p;q)$ process, which are fractionally integrated in variance. Therefore, in a manner analogous to the $\text{ARFIMA}(p;d;q)$ process for the mean, one can define the $\text{FIGARCH}(p;d;q)$ process for $\{\varepsilon_t\}$ as follows:

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + [1-\beta(L)]v_t \qquad (3.26)$$

In Expression (3.26), above, the first difference operator in Expression (3.25) has been replaced with the fractional differencing operator, denoted *d*, where $0 < d < 1$. In addition, in Expression (3.26), all the roots of $\phi(L)$ and $[1-\beta(L)]$, where *L*, $\phi(L)$ and $\beta(L)$ are defined above, lie outside the unit root circle. An alternative representation for the $\text{FIGARCH}(p;d;q)$ model is attained by rearranging the parameters in Expression (3.26), where:

$$[1-\beta(L)]\sigma_t^2 = \omega + \left[1-\beta(L)-\phi(L)(1-L)^d\right]\varepsilon_t^2 \qquad (3.27)$$

Therefore, the conditional variance of ε_t is simply given by:

$$\sigma_t^2 = \omega[1-\beta(1)]^{-1} + \left\{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\right\}\varepsilon_t^2 \qquad (3.28)$$

As with all of the ARCH-type models, in order for the $\text{FIGARCH}(p;d;q)$ to be well-defined and the conditional variance to be positive, all the coefficients in Expression (3.28) must be non-negative. As mentioned above, unlike for the $\text{GARCH}(p;q)$ and $\text{IGARCH}(p;q)$ models, shocks to the conditional variance, in the case of the $\text{FIGARCH}(p;d;q)$ model, will decay at a hyperbolic rate.

One can now move on from the development of the model to the actual implementation. As with the other ARCH-type models, the $\text{FIGARCH}(p;d;q)$ is estimated using the MLE approach, where estimates of the parameters are obtained by maximising the following log-likelihood function:

$$\ell(\theta;\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_T) = -0.5T\log(2\pi) - 0.5\sum_{t=1}^{T}\left[\log(\sigma_1^2) + \varepsilon_1^2\sigma_t^{-2}\right] \qquad (3.29)$$

In the log-likelihood function in Expression (3.29), above, $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T\}$ denotes the sample and $\theta' \equiv (\omega; d; \beta_1, \beta_2, \ldots, \beta_p; \phi_1, \phi_2, \ldots, \phi_q)$. In addition, the MLE approach assumes conditional normality of the process. One should note that, where the standardised innovations $\varepsilon_t \equiv z_t \sigma_t^{-1}$ are leptokurtic and not i.i.d. normally distributed through time, the robust Quasi-Maximum Likelihood Estimation (QMLE) procedures, proposed by Weiss (1986) and Bollerslev and Wooldridge (1992), may be used to allow for asymptotically valid inference.

FIGARCH models have been widely used in the financial market literature, with areas ranging from crude oil markets, see for example Kang, *et al.* (2009), amongst others, to the exchange rate markets, as in Kilic (2007), amongst others. However, to the best of the author's knowledge, this is the first time that this model is being implemented in the shipping literature.

3.4 Determining the Structure of the Higher Moments

Having examined the structure of the first and second moments of shipping freight rates in the previous sections, this section extends the analysis of the structure of freight rates by outlining the methodology to examine the higher moments of the underlying data series. The concepts of time-varying skewness and kurtosis are relatively new to the financial markets literature, and this is, to the best of the author's knowledge, the first application of these concepts in the shipping literature. Despite this novelty, the literature, as discussed in Chapter 2 of the thesis, seems to be unanimous in outlining the importance of understanding the risk structure for participants in these markets.

This sections begins by analysing the initial analysis of conditional skewness in the form of Harvey and Siddique (1999)'s Generalised Autoregressive Conditional Heteroscedasticity with Skewness (GARCHS) model. It then moves on to examine the issue of conditional kurtosis using Brooks, *et al.* (2005)'s Generalised Autoregressive Conditional Heteroscedasticity and Kurtosis (GARCHK) model. This section concludes by examining the Generalised Autoregressive Conditional Heteroscedasticity, Skewness and Kurtosis (GARCHSK), first proposed by León, *et* *al.* (2005), and indicating some previous empirical applications of these concepts in the financial markets literature.

3.4.1 The GARCHS Model

As mentioned above, the focus of the methodology used here changes from merely modelling the conditional volatility to modelling the higher moments as well. Harvey and Siddique (1999) introduce the concept of autoregressive conditional skewness through the introduction of the GARCHS model. The assumption here is that excess returns, which they denote as $r_{M;t+1}$, have a non-central conditional *t*-distribution, which allows one to estimate time-varying skewness of either sign. This distribution is defined by two time-varying parameters, i.e. the degrees of freedom, denoted ν_{t+1}, and the non-centrality parameter, denoted δ_{t+1}, where the conditional variance is used as the scale parameter controlling the dispersion of the data. Harvey and Siddique use the conditional variance to standardise the returns to have unit variance, with a non-zero mean, and then use the conditional mean and skewness to calculate the respective ν_{t+1} and δ_{t+1} for the series.

The sample likelihood function for this non-central *t*-distribution, with unit variance, can be calculated as follows:

$$\ell\left(\varepsilon_{t+1} \middle| \mathrm{Z}_t;\ \Theta\right) = \prod_{t=1}^{T} \frac{\nu_{t+1}^{\frac{\nu_{t+1}}{2}}}{\Gamma\left(\frac{\nu_{t+1}}{2}\right)} \frac{\exp^{\frac{-\delta_{t+1}^2}{2}}}{\sqrt{\pi}\left(\nu_{t+1} + \frac{\varepsilon_{t+1}^2}{1}\right)^{\frac{\nu_{t+1}+1}{2}}}$$
$$\times \sum_{i=0}^{\infty} \Gamma\left(\frac{\nu_{t+1}+i+1}{2}\right)\left(\frac{\delta_{t+1}^i}{i!}\right)\left(\frac{2\frac{\varepsilon_{t+1}^2}{1}}{\nu_{t+1}+\frac{\varepsilon_{t+1}^2}{1}}\right)^{\frac{1}{2}} \tag{3.30}$$

In Expression (3.30), above, Γ denotes the Gamma function, while ν_{t+1} and δ_{t+1} denote the degrees of freedom and the non-centrality parameter, respectively, as discussed above. One should note that δ_{t+1} determines the shape and, therefore, the skewness of the distribution.

The GARCHS model was developed so as to allow for the specification of the conditional variance and skewness as an autoregressive process as follows:

$$h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 \varepsilon_{t-1}^2 \qquad (3.31)$$

$$s_t = \gamma_0 + \gamma_1 s_{t-1} + \gamma_2 \varepsilon_{t-1}^2 \qquad (3.32)$$

In Expressions (3.31) and (3.32), above, $h_t = \mathrm{Var}_{t-1}\left(r_{M\,;\,t}\right)$ and $s_t = \mathrm{Skew}_{\,t-1}\left(r_{M\,;\,t}\right)$. One should note that the variance and skewness in the GARCHS model needs to be constrained in order to ensure that they are stationary, and in the case of the variance, positive. In order to achieve this, Harvey and Siddique impose the constraints that $0 < \beta_1 < 1$, $0 < \beta_2 < 1$, $\beta_1 + \beta_2 < 1$, $-1 < \gamma_1 < 1$, $-1 < \gamma_2 < 1$ and $-1 < \gamma_1 + \gamma_2 < 1$.

In order to allow for the estimation of the model, Harvey and Siddique estimate the central conditional variance and then use the recurrence relation, proposed by Kendall, *et al.* (1991), to obtain the non-central skewness and variance from the central moments as follows:

$$\mu_3 = \mu_3^{'} - 3\mu_2^{'} + 2\mu_1^3 \qquad (3.33)$$

$$\mu_2 = \mu_2^{'} - \mu_1^2 \qquad (3.34)$$

In Expressions (3.33) and (3.34), above, μ_2 and μ_3 are the central moments, about the mean, while $\mu_2^{'}$ and $\mu_3^{'}$ are the non-central moments, about zero.

Following this, Harvey and Siddique (1999) calculate v_{t+1} and δ_{t+1} by solving the following system of nonlinear equations:

$$\mu_1 = \left(\frac{1}{2}v\right)^{\frac{1}{2}} \frac{\Gamma\left[\frac{1}{2}(v-1)\right]}{\Gamma\left(\frac{1}{2}v\right)} \delta \qquad (3.35)$$

$$\mu_3 = \mu_1 \left[\frac{v\left(2v - 3 + \delta^2\right)}{(v-2)(v-3)} - 2 \right] \qquad (3.36)$$

Harvey and Siddique then set the initial conditional variance and skewness, denoted h_1 and s_1, respectively, to the conditional variance and skewness and estimate the parameter set, where $\Theta = [\alpha; \beta; \gamma]$. This is done by maximising the log-likelihood function outlined in Expression (3.30), above.

3.4.2 The GARCHK Model

Brooks, *et al.* (2005) look at a different aspect of the higher moments, where they argue that, following the research by Mandelbrot (1963), it is almost universally accepted that asset returns are leptokurtic, as opposed to normally distributed. For this reason, they introduce the GARCHK model, which allows for the kurtosis to develop over time in a manner that is not fixed with respect to the variance, in order to examine the impact that conditional kurtosis has on asset returns.

This model proposes that if one was to let ε_t, where $t = 1; 2; \ldots; T$, be independently distributed as central Student's *t* variates, with ν_t degrees of freedom, then one could extend Bollerslev (1986)'s GARCH model, in which Bollerslev considers a time-varying transformation of ε_t, denoted λ_t. This extension would result in a new process, which may have any desired variance, denoted h_t, and kurtosis, denoted k_t.

This transformation would be given by the following:

$$\varepsilon_t^* = \lambda_t \varepsilon_t \qquad (3.37)$$

In Expression (3.37), above, ε_t^* are the analogues of the disturbances of a *t*-GARCH model. Following this, one can define the time-varying transformation as a function of the conditional variance and kurtosis of the data series, i.e.:

$$\lambda_t = \left(\frac{k_t h_t}{2k_t - 3} \right) \qquad (3.38)$$

In Expression (3.38), above, h_t and k_t denote the variance and kurtosis, respectively. Following this, Brooks, *et al.* (2005) define the conditional variance, denoted $h_t = \mu_{2;t}$, as follows:

$$\mu_{2;t} = \lambda_t^2 \frac{\nu_t}{\nu_t - 2} \qquad (3.39)$$

In Expression (3.39), above, λ_t and ν_t denote the time-varying transformation and the degrees of freedom, respectively. The conditional fourth moment, denoted $k_t = \mu_{4;t}$, is defined, in turn, as:

$$\mu_{4;t} = \lambda_t^4 \frac{3\nu_t^2}{(\nu_t - 2)(\nu_t - 4)} \qquad (3.40)$$

In Expression (3.40), above, λ_t and ν_t denote the time-varying transformation and the degrees of freedom, respectively. It is worth noting that Expressions (3.39) and (3.40), above, arise from the moment-generating function for a central *t*-distribution, in which all odd moments are zero, by definition. Brooks, *et al.* then rearrange Expression (3.39), above, to obtain the time-varying transformation as a function of the conditional variance and the time-varying degrees of freedom, hence:

$$\lambda_t = \left[\frac{h_t(\nu_t - 2)}{\nu_t} \right]^{1/2} \qquad (3.41)$$

In Expression (3.41), above, λ_t and ν_t, once again, denote the time-varying transformation and the degrees of freedom, respectively, while h_t denotes the conditional variance. Following this, Brooks, *et al.* define the conditional kurtosis as $k_t = \mu_{4;t} / h_t$, and then substitute Expression (3.41) into (3.40) to attain the conditional kurtosis as a function of the degrees of freedom at time *t*, therefore:

$$k_t = \frac{3(\nu_t - 2)}{(\nu_t - 4)} \qquad (3.42)$$

In Expression (3.42), above, ν_t denotes the degrees of freedom and k_t denotes the conditional variance. Following this, Expression (3.42) is rearranged to determine the degrees of freedom as a function of the conditional kurtosis, hence:

$$\nu_t = \frac{2(2k_t - 3)}{k_t - 3} \qquad (3.43)$$

In Expression (3.43), above, ν_t and k_t denote the degrees of freedom and the conditional kurtosis, respectively. Expressions (3.39) to (3.43) illustrate that there is no fixed relationship between the conditional variance and kurtosis and therefore these may vary freely over time as the conditional kurtosis depends only on the degrees of freedom, whereas the conditional variance also depends on the time-varying transformation. This process means that, as they are not directly functionally related, one is able to parameterise the conditional variance and kurtosis terms individually, as desired. One should note that there is a degrees of freedom restriction in that $\nu_t > 4$, if the requirement for the existence of a second and fourth conditional moment is to be met.

Brooks, *et al.* (2005) highlight that in order to estimate the parameters of these terms, one should note that the Jacobian of the transformation $\varepsilon_t^* / \lambda_t = \varepsilon_t$ is:

$$J = \frac{\partial \varepsilon_t}{\partial \varepsilon_t^*} = \frac{1}{\lambda_t} \qquad (3.44)$$

One should recall that, in Expression (3.44), above, λ_t and ε_t^* denote the time-varying transformation and the analogues of the disturbances of a *t*-GARCH model, respectively. The density of ε_t^* is then obtained by taking the Student's *t*-density for ε_t, into which one would substitute $\varepsilon_t^* / \lambda_t = \varepsilon_t$, and then multiplying this by the Jacobian, therefore:

$$f\left(\varepsilon_t^*\right) = \frac{1}{\lambda_t} \frac{\Gamma\left[\left(\nu_t + 1\right)/2\right]}{\pi^{1/2} \nu_t^{1/2} \Gamma\left[\nu_t/2\right] \left(1 + \varepsilon_t^{*2} / \lambda_t^2 \nu_t\right)^{\left(\nu_t + 1\right)/2}} \qquad (3.45)$$

One should note, once again, that, in Expression (3.45), ε_t^*, λ_t and ν_t denote the analogues of the disturbances of a *t*-GARCH model, the time-varying transformation and the degrees of freedom, respectively, while Γ denotes the Gamma function. Following this, one can determine the log-likelihood function for the t^{th} observation by substituting for λ_t in Expression (3.45), and then taking logarithms of the resultant function. The log-likelihood function is therefore:

$$\begin{aligned} \ell = \log\left[\Gamma\left\{\left(\nu_t + 1\right)/2\right\}\right] - \log\left[\Gamma\left\{\nu_t/2\right\}\right] - \frac{1}{2}\log\left[h_t\right] - \frac{1}{2}\log\left[\nu_t - 2\right] \\ - \frac{\nu_t + 1}{2} \log\left[1 + \frac{\varepsilon_t^*}{h_t\left(\nu_t - 2\right)}\right] \end{aligned} \qquad (3.46)$$

In Expressions (3.46), above, ε_t^*, Γ and ν_t denote the analogues of the disturbances of a *t*-GARCH model, the Gamma function and the degrees of freedom, respectively. In addition, the degrees of freedom are a function of the conditional kurtosis, hence, should one maximise the log-likelihood function, this would yield the MLE estimates for all the parameters of the model.

Brooks, *et al.* (2005) finally formalise the GARCHK model, which is described using the following series of expressions:

$$y_t = \gamma_0 + \varepsilon_t^* \qquad (3.47)$$

$$\varepsilon_t^* = \lambda_t \varepsilon_t \ ; \ \varepsilon_t \sim t_{\nu_t} \qquad (3.48)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^{*2} + \alpha_2 h_{t-1} \qquad (3.49)$$

$$k_t = \beta_0 + \beta_1 \frac{\varepsilon_{t-1}^{*2}}{h_{t-1}^2} + \beta_2 k_{t-1} \qquad (3.50)$$

$$\nu_t = \frac{2\left(2k_t - 3\right)}{k_t - 3} \qquad (3.51)$$

$$\lambda_t = \left(\frac{h_t \left(\nu_t - 2\right)}{\nu_t} \right)^{1/2} \qquad (3.52)$$

In Expressions (3.47) and (3.52), ε_t^*, λ_t, h_t, k_t and ν_t denote analogues of the disturbances of a *t*-GARCH model, the time-varying transformation, the conditional variance, the conditional kurtosis and the degrees of freedom, respectively. One should note that the parameters in this model are estimated using QMLE estimation. The first point to note regarding the model is that, as a result of Expression (3.42), there is a degrees of freedom restriction in that $\nu_t > 4$, if the requirement for the existence of a fourth moment is to be met. In addition to this, Expression (3.42) also implies that $k_t \to 3$ as $\nu_t \to \infty$, while $k_t \to \infty$ as $\nu_t \to 4$ in the model. One should also note that it is sufficient that $\alpha_0 > 0$, $\alpha_1 > 0$, $\beta_0 > 0$, $\beta_1 > 0$, $\alpha_2 \geq 0$, $\beta_2 \geq 0$, $h_t > 0 \forall t$ and $k_t > 3 \forall t$. The final issue to note for the GARCHK model, is that the initial values for the conditional variance and kurtosis series are set such that every element is equal to their respective unconditional values.

The next section looks at the joint estimation of conditional skewness and kurtosis, with the idea that if one can incorporate both of these into one model, one can generate a better understanding of the underlying data series.

3.4.3 The GARCHSK Model

A disadvantage of the GARCHS and GARCHK models is that, while the GARCHS model allows one to model the conditional skewness of a data series, it does not allow one to model the conditional kurtosis, and vice versa for the GARCHK model. This means that, essentially, one is unable to jointly estimate the higher moments of a data series using either of these models. This issue is resolved by León, *et al.* (2005), whose GARCHSK model enables one to jointly estimate the conditional variance, skewness and kurtosis of the underlying data series. An additional advantage of the GARCSK model is that the likelihood function is based on a Gram-Charlier (GC) expansion of the normal density function, in a manner similar to that suggested by Gallant and Tauchen (1989), which makes it easier to estimate than the likelihood function based on the non-central *t*-distribution used by Harvey and Siddique (1999).

León, *et al.* (2005)'s GARCHSK model is given by:

$$r_t = E_{t-1}\left(r_t\right) + \varepsilon_t \ ; \ \varepsilon_t \sim \left(0;\sigma_\varepsilon^2\right) \qquad (3.53)$$

$$\varepsilon_t = h^{1/2}\eta_t \ ; \ \eta_t \sim \left(0;1\right) \ ; \ \varepsilon_t \middle| I_{t-1} \sim \left(0;h_{t-1}\right) \qquad (3.54)$$

$$h_t = \beta_0 + \beta_1\varepsilon_{t-1}^2 + \beta_2 h_{t-1} \qquad (3.55)$$

$$s_t = \gamma_0 + \gamma_1\eta_{t-1}^3 + \gamma_2 s_{t-1} \qquad (3.56)$$

$$k_t = \delta_0 + \delta_1\eta_{t-1}^4 + \delta_2 k_{t-1} \qquad (3.57)$$

In Expressions (3.53) to (3.57), above, $E_{t-1}\left(\cdot\right)$ denotes the conditional expectation on an information set till period $t-1$, where this in turn is denoted I_{t-1}. Leon, *et al.* establish that $E_{t-1}\left(\eta_t\right)=0$, $E_{t-1}\left(\eta_t\right)^2=1$, $E_{t-1}\left(\eta_t\right)^3=s_t$, and $E_{t-1}\left(\eta_t\right)^4=k_t$, where both s_t and k_t are driven by a $\text{GARCH}\left(1;1\right)$ structure. This means that s_t and k_t represent to the skewness and kurtosis corresponding to the conditional distribution of the standardised residuals, denoted η_t, where $\eta_t = \varepsilon_t h^{-1/2}$, respectively.

Leon, *et al.* go on to obtain the density function for the standardised residuals, denoted η_t, which is conditional on the information available at time $t-1$, by using a

GC series expansion of the normal density function and truncating this at the fourth moment. Therefore, the density function will be as follows:

$$g\left(\eta_t \left| I_{t-1}\right.\right)=\phi\left(\eta_t\right)\left[1+\frac{s_t}{3!}\left(\eta_t^3-3\eta_t\right)+\frac{k_t-3}{4!}\left(\eta_t^4-6\eta_2^2+3\right)\right]=\phi\left(\eta_t\right)\psi\left(\eta_t\right) \quad (3.58)$$

In Expression (3.58), above, $\phi(\cdot)$ denotes the probability density function (pdf) corresponding to the standard normal distribution, while $\psi(\cdot)$ is the polynomial part of the fourth order, corresponding to expression between the brackets. One can argue that this is not really a density function in that, for some of the parameter values in Expressions (3.53) to (3.57), the density function $g(\cdot)$ might be negative, and, similarly, the integral of $g(\cdot)$ on $\Re$ is not equal to one.

To solve this issue, Leon, *et al.* propose a true pdf, denoted $f(\cdot)$, where this is obtained by transforming the density function $g(\cdot)$ using the method outlined in Gallant and Tauchen (1989). Looking at the specifics, to obtain this well defined density everywhere, the polynomial part of the density function, i.e. $\psi(\cdot)$, is squared, and then divided by the integral of $g(\cdot)$ over $\Re$, where the latter is to ensure that the integral of $g(\cdot)$ is equal to one. This means that the resulting pdf will be as follows:

$$f\left(\eta_t \left| I_{t-1}\right.\right)=\frac{\phi\left(\eta_t\right)\left[1+\frac{s_t}{3!}\left(\eta_t^3-3\eta_t\right)+\frac{k_t-3}{4!}\left(\eta_t^4-6\eta_2^2+3\right)\right]^2}{\Gamma_t}=\frac{\phi\left(\eta_t\right)\psi^2\left(\eta_t\right)}{\Gamma_t} \quad (3.59)$$

In Expression (3.59), above, all the terms are as defined above, with the exception of the term Γ_t, which is defined as follows:

$$\Gamma_t=1+\frac{s_t^2}{3!}+\frac{\left(k_t-3\right)^2}{4!} \quad (3.60)$$

In Expression (3.60), above, s_t and k_t are as defined above.

Following this, the log-likelihood function, after omitting unessential constants, can be defined as the following:

$$\ell_t=-\frac{1}{2}\ln h_t-\frac{1}{2}\eta_t+\ln\left[\psi^2\left(\eta_t\right)\right]-\ln\left[\Gamma_t\right] \quad (3.61)$$

One should note that the log-likelihood function defined in Expression (3.61), above is for one observation corresponding to conditional distribution $\varepsilon_t = h^{1/2}\eta_t$, where the pdf for this conditional distribution is $h^{1/2} f\left(\eta_t \left| I_{t-1}\right.\right)$. As one can see, this likelihood function is clearly much easier to estimate than the density function proposed by Harvey and Siddique (1999), which is based on a non-central *t*-distribution. An additional advantage is that the pdf in Expression (3.59) nests the normal density function, which would occur when $s_t = 0$ and $k_t = 3$, while that based on the non-central *t*-distribution does not.

3.5 Conclusion

This chapter examined the various methodologies that will be used in this thesis to gain a better understanding of the underlying structure of the different moments of freight rates in the shipping markets. As outlined above, this thesis introduces the concepts of fractional integration and conditional skewness and kurtosis to the shipping literature, and therefore may provide for a greater understanding of the structure of freight rates. In addition, this chapter introduces the methodology that will be used to gain a better understanding of the future direction of spot freight rates across the different vessel types, and a greater insight into the potential risk exposure faced by participants in these markets. Given this, the following chapters will provide an empirical application of these methodologies and provide an insight into the meaning of the results of these empirical studies.

Appendix 3.A – A Discussion on Stationarity

Brooks (2002) and Tong (1990) highlight the fact that there are two main forms of stationarity for a data series. A data series is defined as a strictly stationary process if the distribution of the values of the data series remains constant across time, i.e.:

$$P\left\{y_t\right\} = P\left\{y_{t+k}\right\} \forall k \qquad (3.A.1)$$

In other words, for a data series to be strictly stationary, the probability that the variable y_t falls within a particular interval must be the same now as any other point in time. In contrast, a data series is defined as a weakly stationary process if it satisfies three conditions, the first of which is that it has a constant mean, i.e.:

$$E\left(y_t\right) = \mu \qquad (3.A.2)$$

The second condition is that the data series a constant variance, hence:

$$E\left(y_t - \mu\right)\left(y_t - \mu\right) = \sigma^2 < \infty \qquad (3.A.3)$$

The third, and final, condition is that the data series constant autocovariance for each lag, in other words:

$$E\left(y_{t_1} - \mu\right)\left(y_{t_2} - \mu\right) = \gamma_{t_2 - t_1} \quad \forall t_1; t_2 \qquad (3.A.4)$$

One should note that the autocovariance is defined as the extent to which the value of y_t is related to its previous values, where, for a weakly stationary series, this will depend only on the difference between t_1 and t_2, hence one could state that the covariance between y_t and y_{t-1} will be the same as that for y_{t-4} and y_{t-5}.

This thesis uses three main tests to determine whether the underlying process is stationary or non-stationary, where the first of these is the Augmented Dickey-Fuller (ADF), first proposed by Dickey and Fuller (1981). For this ADF test, assume that one is given the following model:

$$\Delta y_t = \alpha y_{t-1} + x_t^{'}\delta + \sum_{i=1}^{p} \phi_1 \Delta y_{t-i} + \varepsilon_t \qquad (3.A.5)$$

In Expression (3.A.5), above, $\Delta y_t = y_t - y_{t-1}$ and $\alpha = p - 1$. One then estimates this model and performs the following hypothesis test:

$$\begin{aligned} H_0 &: \alpha = 0 \\ H_1 &: \alpha < 0 \end{aligned} \qquad (3.A.6)$$

One should note that in Expression (3.A.6), above, the null hypothesis implies that the underlying process is non-stationary, as opposed to the alternate hypothesis, where the underlying process would be stationary. This test is performed using the standard *t*-ratio for a regression coefficient, i.e.:

$$t_\alpha = \frac{\hat{\alpha}}{SE(\hat{\alpha})} \qquad (3.A.7)$$

In Expression (3.A.7), above, $\hat{\alpha}$ denotes the estimate of the coefficient α, while $SE(\hat{\alpha})$ denotes the standard error of the coefficient. One should note that this test statistic does not follow the conventional Student's *t*-distribution, and must be compared with critical values outlined in Dickey and Fuller (1979).

The second, alternative, test is the Phillips-Perron (PP) test, first proposed by Phillips and Perron (1988), who argue that this is an improvement on the ADF test in that it is non-parametric. For the PP test, assume that one is given the following model:

$$\Delta y_t = \alpha y_{t-1} + x_t^{'} \delta + \varepsilon_t \qquad (3.A.8)$$

In Expression (3.A.8), above, $\Delta y_t = y_t - y_{t-1}$ and $\alpha = p - 1$. One then estimates this model and performs the following hypothesis test:

$$\begin{aligned} H_0 &: \alpha = 0 \\ H_1 &: \alpha < 0 \end{aligned} \qquad (3.A.9)$$

Once again, one should note that in Expression (3.A.9), above, the null hypothesis implies that the underlying process is non-stationary, as opposed to the alternate hypothesis, where the underlying process would be stationary. This hypothesis test is based on the following test statistic:

$$\bar{t}_\alpha = t_\alpha \left(\frac{\gamma_0}{f_0} \right) - \frac{T(f_0 - \gamma_0) SE(\hat{\alpha})}{2 f_0^{1/2} s} \qquad (3.A.10)$$

In Expression (3.A.10), above, $\hat{\alpha}$, $SE(\hat{\alpha})$ and t_α denote the estimate, standard error and *t*-ratio of the coefficient α, respectively, where the *t*-ratio is calculated in the same was as outlined in Expression (3.A.7), and *s* denotes the standard error of the test regression as a whole. In addition, γ_0 is a consistent estimate of the error variance in Expression (3.A.8), where $\gamma_0 = (T-k)s^2 / T$ and k denotes the number of parameters, and f_0 denotes an estimator of the residual spectrum frequency zero.

The final test of stationarity discussed here is the KPSS tests, first proposed by Kwiatkowski, *et al.* (1992), where the KPSS test statistic, which is a Lagrange multiplier statistic, is based on the residuals from the following model:

$$y_t = x_t^{'}\delta + \varepsilon_t \qquad (3.A.11)$$

One then tests the following hypotheses:

$$\begin{aligned} &H_0 : \text{Underlying process is stationary} \\ &H_1 : \text{Underlying process is non-stationary} \end{aligned} \qquad (3.A.12)$$

This test is performed using the following LM statistic, where:

$$\text{LM} = \sum_t S(t)^2 / \left(T^2 f_0\right) \qquad (3.A.13)$$

In Expression (3.A.13), f_0 denotes an estimator of the residual spectrum frequency zero and $S(t)$ is a cumulative residual function, where $S(t) = \sum_{i=1}^{t} \hat{\varepsilon}_t$ and $\hat{\varepsilon}_t = y_t - x_t^{'}\hat{\delta}(0)$. This statistic is then compared to the reported critical values for the LM test, where these are presented in Table 1 on p. 166 of Kwiatkowski, *et al.* (1992).

4 Description of Data

4.1 Introduction

Having outlined the methodology to be used in the thesis in Chapter 3, the aim of this chapter is to present the data that will be used to analyse the structure of freight rates. These data series were collected from Clarkson Shipping Intelligence Network, a database of all relevant data series for the shipping industry. In order to model the structure of freight rates, prevailing weekly spot rates on shipping routes for five different types of vessels were selected, with the sample period extending from 13 January 1989 to 26 June 2009, comprising 1,068 observations. Of the five types of vessels, three tankers, i.e. VLCC, Suezmax and Aframax tankers, and two dry bulk vessels, i.e. the Capesize and Panamax dry bulk vessels, were selected in order to give a balanced perspective of the tramp shipping market.

In addition to evaluating the spot freight levels, proxies for freight rate returns were calculated for each of the respective vessel data series, using the formula in Expression (4.1) below:

$$r_t = \frac{\Delta^d FR_t}{FR_t} \qquad (4.1)$$

One should note that FR_t denotes the freight rate in period t, while d denotes the order of fractional integration, which is estimated using Autoregressive Fractionally Integrated Moving Average (ARFIMA) models. The reason for using these proxies for freight rate returns is that should one wish to incorporate an ARFIMA model into the mean equation for the volatility model, using standard returns would automatically render the series stationary as one inherently differences the series to generate standard returns. This would therefore imply that any fractionally integrated properties would have been removed.

One small point to note before going any further is that the manner in which freight rates are quoted differs across sectors. In the tanker sector, freight rates are quoted in terms of the Worldscale (WS), which can then be converted into US$ per tonne, and according to the route, using the Worldscale Book, which is revised annually. In contrast, dry-bulk sector freight rates are simple quoted in terms of US$ per tonne, regardless of the route on which the goods are transported.

In the case of the VLCC data series, these are based on the spot freight rates for the transportation of crude oil on a 270,000DWT VLCC tanker, where the port of loading is Ras Tanura (Saudi Arabia) and the port of discharge is Rotterdam (Netherlands). For the Suezmax data series, these are based on the spot freight rates for the transportation of crude oil on a 130,000DWT Suezmax tanker, where the port of loading is Bonny (Nigeria) and the port of discharge is Off the Coast of Philadelphia (United States). As regards the Aframax data series, these are based on the spot freight rates for the transportation of crude oil on an 80,000DWT Aframax tanker, where the port of loading is Sullom Voe (United Kingdom) and the port of discharge is Bayway (United States). In the case of the bulk carriers, for the Capesize data series, these are based on the spot freight rates for the transportation of ore on a 165,000DWT Capesize tanker, where the port of loading is Tubarao (Brazil) and the port of discharge is Rotterdam (Netherlands); and finally for the Panamax data series, these are based on the spot freight rates for the transportation of coal on a 70,000DWT Panamax dry bulk carrier, where the port of loading is Hampton Roads (United States) and the port of discharge is in the Antwerp-Rotterdam-Amsterdam range (Belgium and Netherlands).

One should note that voyage freight rates, as opposed to trip-charter rates were used as limitations regarding the availability of the data meant that should trip-charter freight rates have been used, the sample would have been dramatically smaller, for example, data across all Capesize trip-charter routes is only available on Clarkson Shipping Intelligence Network, at the weekly frequency, for the period between 17 July 2009 to the present. Furthermore, trip-charter freight rates are not available for the tanker sector, therefore, so as to ensure comparability and consistency between sectors, voyage freight rates, which are reported both for the tanker and dry-bulk sectors are preferred. A second possible issue regards the route selection process. The routes used in this thesis were selected on the grounds that they maximised the size of the sample set, while ensuring that there was sufficient liquidity on these routes so as to justify their inclusion. A final possible issue regards the use of raw spot freight rates as opposed to the log of these freight rates. This is not necessarily a concern as the models in this thesis were also run on the natural logarithms of the respective data series in order to take account of the fact that freight rates can never be negative. One should note that the results from these estimations did not differ significant from those presented here[16].

Having outlined the routes and data sample period, the next section of the chapter examines the characteristics of the spot freight rate levels, while the subsequent section looks at the descriptive statistics for the freight returns. The chapter then investigates the conditional properties of the data set using Ljung-Box statistics and, finally, reviews the findings.

4.2 Description of Spot Freight Rates

Examining firstly the evolution of the tanker spot freight rates, as illustrated in Figure 4.1, freight rates for the tanker market were relatively stable for the ten year period up to 2001 when they experienced the first of a series of peaks which followed over the next seven or eight years. The first peak coincides with the initiation and the process of accelerated phasing out of single-hulled tankers in favour of the double-hulled alternates, as a result of the amendment to the Marpol Convention. This led to reduction in the number of vessels in the tanker fleet and a consequent decrease in the supply of tanker services and a resultant dramatic increase in freight rates as a result of the shape of the supply curve. The second peak corresponds to the second Gulf War in 2003 as well as a further amendment to the Marpol Convention which increased the scrapping schedule of single-hulled tankers, with corresponding effects on the price of oil and supply of shipping services, and therefore freight rates. This peak then leads on to the further peaks resulting from an increased demand for oil, increased oil prices and the development boom in China, and therefore an increased demand for transportation services for the oil needed. These series of peaks were followed by an unprecedented collapse in the freight rate market in late 2008, caused by both the world undergoing arguably the most severe economic slowdown since the Great Depression, combined with massive over-ordering during the previous peaks resulting in a huge number of new tankers entering the fleet thus increasing the supply

[16] The results from these estimations are not presented here due to space constraints and are available from the author upon request.

Figure 4.1 – Evolution of Tanker Spot Freight Rates (13 January 1989 to 26 June 2009)

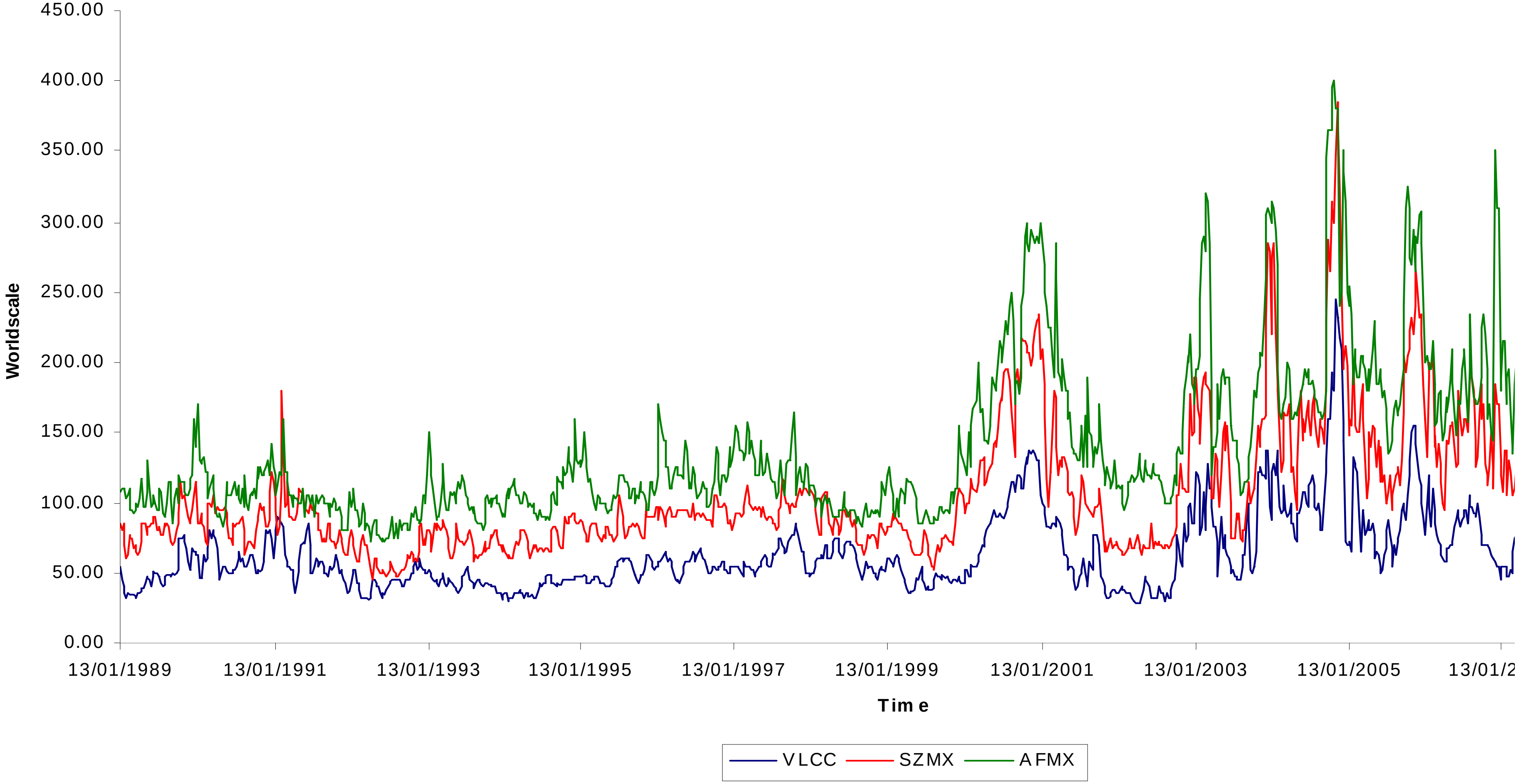
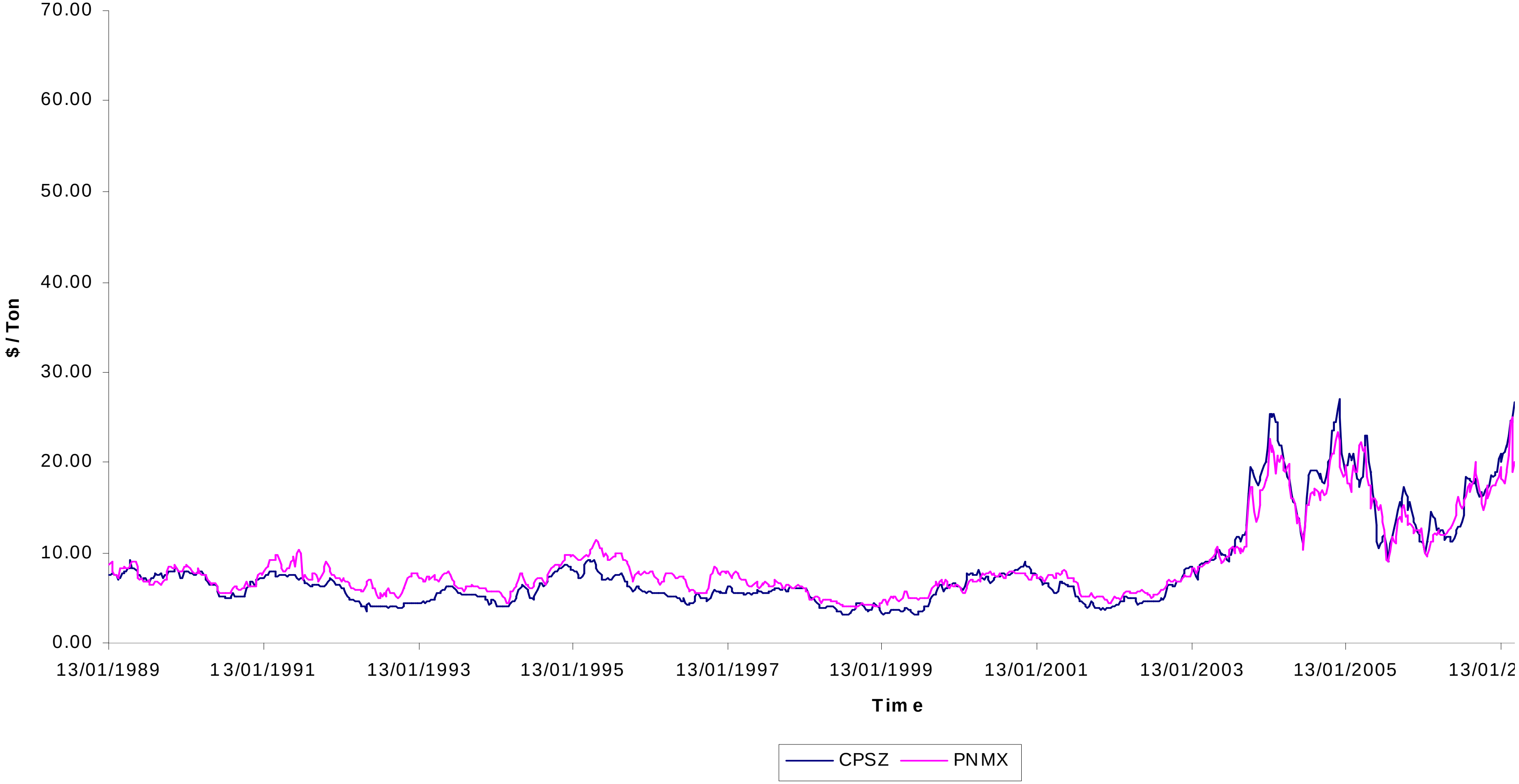


Figure 4.2 – Evolution of Dry-Bulk Spot Freight Rates (13 January 1989 to 26 June 2009)



of vessels to record levels and causing an extremely and sudden fall in freight rate levels. One could also argue that these peaks may be attributed to the rescaling of the Worldscale rates in January, however, this is not felt to be a major factor as many of the peaks occur in the middle of the year.

When changing focus to look at the dry bulk market, the picture is somewhat more tranquil, as illustrated in Figure 4.2, in that freight rates remained relatively stable at fairly low levels until 2003. During the period between 2004 and 2005, a first peak is found, corresponding with increased demand for commodities driven by the growth of the Chinese economy. The correction in the market was then followed by a massive increase in freight around the end of 2006 and beginning of 2007, driven by a rapid increase in the demand for commodities by China, congestion in world ports leading to tonnage being tied up, and a dramatic increase in the price of commodities. However, as was the case in the tanker market, a dramatic slowdown in world economic growth, as well as over-ordering, led to as extreme a fall in freight rates, although in this case, continued demand for commodities such as coal and iron ore, by China led to a much faster and somewhat greater recovery.

Having examined the evolution of freight rates, one can now begin to focus on the descriptive statistics in Panel A of Table 4.1 to obtain a more detailed picture of the dynamics of the spot freight rates. Studying the mean statistics, the mean spot freight rate was found to range between WS64.436 and WS142.682 across the tanker market, while in the dry bulk market, mean freight rates were found to range between $10.38 and $10.65 per ton. A size effect is observed here; where larger vessels have lower mean freight rate levels in both markets. This is logical given that larger vessels are expected to benefit from cost-advantages, since they are able to carry much larger cargoes and therefore incur lower costs per ton of cargo. In addition to this, there again appears to be a size effect in terms of the standard deviation of spot freight rates in the tanker market. This may be due to the fact that larger vessels are only able to operate on fewer routes owing to cargo and port restrictions, hence the supply of vessels is fairly constant and freight rates relatively stable. In the case of smaller vessels, they are free to choose the routes on which they operate, and as such freight rates will fluctuate on routes as the supply of vessels alters. In the case of the dry bulk market, the size effect in the standard deviations is reversed. In this case, restrictions in terms of the cargo they can carry and the ports in which they can operate entail that vessels may spend time unemployed, resulting in fluctuations in supply and hence freight rates. This is not necessarily the case with smaller vessels as they have greater flexibility with respect to cargoes they can carry and ports from which they can operate and therefore the supply of vessels is much more stable.

Moving on to the spot freight rate distribution in Panel A of Table 4.1, all five data series exhibit large and significant positive skewness. This implies that there would be a higher probability of earning a freight rate in excess of the mean level than below it. A possible reason for this positive skewness is that the shape of the supply curve in the shipping market is so shaped that it is relatively flat for a large part, before bending steeply upwards. This indicates that in periods of excess supply, freight rates are unlikely to fall far below the mean level, whereas in periods of supply shortage, freight rates are expected to be well above the mean, thus positively skewing the freight rate distribution. In addition to this positive skewness, spot freight rates are also found to exhibit significant excess kurtosis, hence there will be fat-tails in the distribution, which implies a higher probability of extremely high or low freight rates, which can once again be explained by the shape of the supply function in the shipping markets, in a manner similar to that for the skewness. Finally, the Jarque-Bera statistic confirms this in that one can reject the null hypothesis of normally distributed spot freight rates at all conventional levels of significance for all five data series.

This analysis of the structure of spot freight rates is concluded by examining the question of whether the data series are stationary or not by performing three standard unit root tests on the spot freight rate levels, the results of which are presented in Table 4.2. These tests are performed on the basis that there is a constant but no trend, as there appeared to be no trend in the data, hence it could not be trend stationary. In the dry bulk market the Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and KPSS tests developed by Dickey and Fuller (1981), Phillips and Perron (1988) and Kwiatkowski, *et al.* (1992), respectively, are unanimous in concluding that the CPSZ and PNMX data series are non-stationary; however, in the tanker market, the picture is not as clear. In this case, the ADF and PP tests both agree that the VLCC, SZMX and AFMX data series are stationary, a conclusion in direct conflict with that of the KPSS test, which indicates that these same spot freight rate series are non-stationary.

Table 4.1 – Descriptive Statistics for Spot Freight Rates and Returns

Panel A – Descriptive Statistics for Spot Freight Rates	VLCC	SZMX	AFMX	CPSZ	PNMX
Observations	1068.000	1068.000	1068.000	1068.000	1068.000
Mean	64.436	108.743	142.682	10.377	10.647
Variance	906.249	2498.052	3451.005	86.856	64.482
Standard Deviation	30.104	49.981	58.745	9.320	8.030
Skewness	1.938	1.802	1.551	2.490	2.444
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	8.453	6.909	5.351	9.666	9.136
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	1991.701	1257.871	674.112	3080.334	2739.112
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Panel B – Descriptive Statistics for Freight Rate Returns	VLCC	SZMX	AFMX	CPSZ	PNMX
Observations	1068.000	1068.000	1068.000	1068.000	1068.000
Mean	-0.026	-0.005	-0.002	0.063	0.039
Variance	0.684	0.713	0.598	0.194	0.178
Standard Deviation	0.827	0.845	0.773	0.440	0.422
Skewness	0.315	0.521	0.482	0.047	-0.433
	(0.000)	(0.000)	(0.000)	(0.527)	(0.000)
Kurtosis	8.474	6.185	8.267	11.882	8.018
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	1351.253	499.797	1275.624	3511.140	1154.020
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note 1: VLCC denotes the weekly spot freight rates for a 270,000 DWT VLCC tanker carrying crude oil from Ras Tanura (Saudi Arabia) to Rotterdam (Netherlands).
SZMX denotes the weekly spot freight rates for a 130,000 DWT Suezmax tanker carrying crude oil from Bonny (Nigeria) to off the coast of Philadelphia (USA).
AFMX denotes the weekly spot freight rates for an 80,000 DWT Aframax tanker carrying crude oil from Sullom Voe (UK) to Bayway (USA).
CPSZ denotes the weekly spot freight rates for a 145,000 DWT Capesize bulk-carrier carrying iron ore from Tubarao (Brazil) to Rotterdam (Netherlands).
PNMX denotes the weekly spot freight rates for a 55,000 DWT Panamax bulk-carrier carrying grain from the Hampton Roads (USA) to Antwerp-Rotterdam-Amsterdam (Benelux).

Note 2: The sample period for the data used for this table extends from 13 January 1989 to 26 June 2009, with a total of 1,068 observations.

Note 3: The mean, variance and standard deviation in Panel A are weekly figures; whereas the mean, variance and standard deviation in Panel B are annualised.

Note 4: The spot freight rates for the VLCC, Suezmax and Aframax tankers' data series are denoted in Worldscale units.

Note 5: The spot freight rates for the Capesize and Panamax bulk-carriers' data series are denoted in US$ per metric tonne.

Note 6: The data used for this table is all sourced from the Clarkson Shipping Intelligence Network (www.clarskons.net).

Note 7: Figures in parentheses denote the respective *p*-values, where for the skewness and kurtosis tests the null hypothesis is that these statistics are equal to zero and for the Jarque-Bera test it is that the data series is normally distributed.

Table 4.2- Unit Root Test for Spot Freight Rate Levels and First-Differences

Panel A – Unit Root Tests for Spot Freight Rate Levels	**VLCC**	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
ADF Test	-5.064	-4.597	-4.232	-2.116	-2.196
t-value at 1%	-(3.436)	-(3.436)	-(3.436)	-(3.436)	-(3.436)
t-value at 5%	-(2.864)	-(2.864)	-(2.864)	-(2.864)	-(2.864)
t-value at 10%	-(2.568)	-(2.568)	-(2.568)	-(2.568)	-(2.568)
Phillips-Perron Test	-5.163	-5.166	-4.998	-2.827	-1.918
t-value at 1%	-(3.436)	-(3.436)	-(3.436)	-(3.436)	-(3.436)
t-value at 5%	-(2.864)	-(2.864)	-(2.864)	-(2.864)	-(2.864)
t-value at 10%	-(2.568)	-(2.568)	-(2.568)	-(2.568)	-(2.568)
KPSS Test	1.321	1.946	2.233	1.959	1.819
t-value at 1%	(0.739)	(0.739)	(0.739)	(0.739)	(0.739)
t-value at 5%	(0.463)	(0.463)	(0.463)	(0.463)	(0.463)
t-value at 10%	(0.347)	(0.347)	(0.347)	(0.347)	(0.347)

Panel B – Unit Root Tests for Spot Freight Rate 1st Differences	**VLCC**	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
ADF Test	-17.042	-22.905	-23.938	-8.911	-16.923
t-value at 1%	-(3.436)	-(3.436)	-(3.436)	-(3.436)	-(3.436)
t-value at 5%	-(2.864)	-(2.864)	-(2.864)	-(2.864)	-(2.864)
t-value at 10%	-(2.568)	-(2.568)	-(2.568)	-(2.568)	-(2.568)
Phillips-Perron Test	-32.372	-34.196	-35.253	-26.181	-27.732
t-value at 1%	-(3.436)	-(3.436)	-(3.436)	-(3.436)	-(3.436)
t-value at 5%	-(2.864)	-(2.864)	-(2.864)	-(2.864)	-(2.864)
t-value at 10%	-(2.568)	-(2.568)	-(2.568)	-(2.568)	-(2.568)
KPSS Test	0.029	0.028	0.031	0.028	0.038
t-value at 1%	(0.739)	(0.739)	(0.739)	(0.739)	(0.739)
t-value at 5%	(0.463)	(0.463)	(0.463)	(0.463)	(0.463)
t-value at 10%	(0.347)	(0.347)	(0.347)	(0.347)	(0.347)

Note 1: See notes from Table 4.1.

Note 2: ADF test denotes the results from conducting the Augmented Dickey-Fuller test for stationarity (Dickey and Fuller (1981)).
Phillips-Perron test denotes the results from conducting the Phillips-Perron test for stationarity (Phillips and Perron (1988)).
KPSS test denotes the results from conducting the Kwiatkowski, Phillips, Schmidt and Shin test for stationarity (Kwiatkowski, *et al.* (1992)).

Note 3: The null hypotheses for the ADF and Phillips-Perron tests are that the data series contains a unit root, i.e. is non-stationary.
The null hypothesis for the KPSS test is that the data series is stationary.

Note 4: The figures in parentheses denote the *t*-values for the respective data series at the given levels of significance.

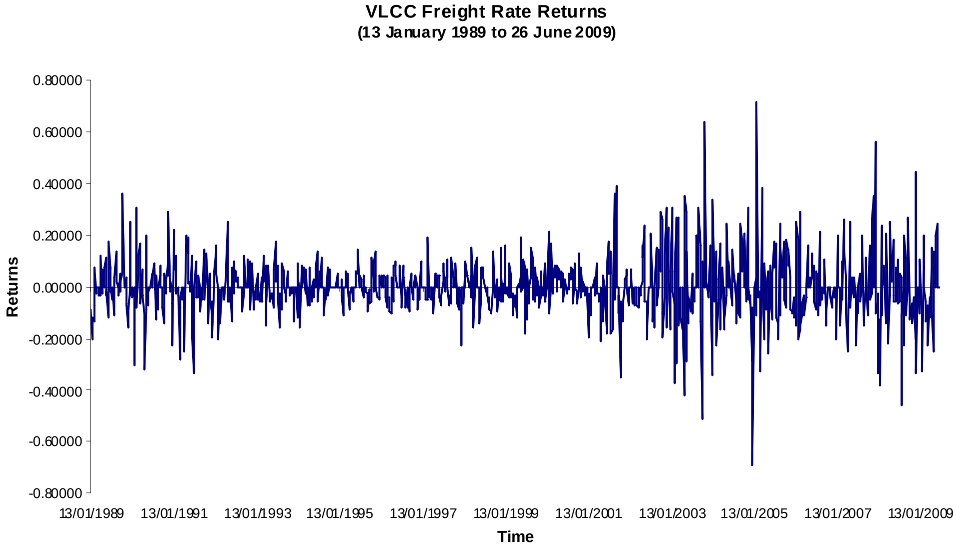
This conflict in the results of the unit root tests for the tanker market illustrates the need for further examination and lends support to the proposal that a third alternative, i.e. fractional integration needs to be investigated. Finally, to check the order of the non-stationarity, should freight rates indeed be non-stationary, these three unit root tests were performed on the first-differences of the spot freight rates, where all three tests were unanimous in both markets and across all data series in concluding that the first differences of the spot freight rates are stationary.

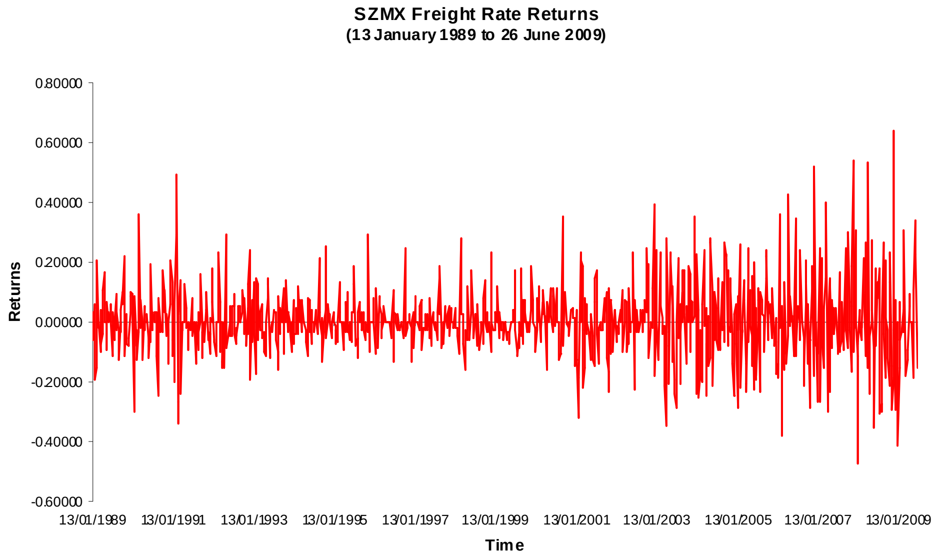
4.3 Description of Freight Rate Returns

Moving on to examine the higher moments of the spot freight rate series, the focus now changes to the characteristics of the freight returns. These freight rate returns are the returns on spot freight rates for the data series described in Section 4.1 above and are calculated using the formula in Expression (4.1), above. Figure 4.3 and Figure 4.4 portray the evolution of the freight rate returns over the sample period, with higher volatility in returns being observed during the periods coinciding with the peaks in spot freight rates discussed in Section 4.2 above. One should also note that it evident from the figures that in the latter part of the sample period there is significant volatility clustering, which gives a preliminary indication that there may be GARCH effects within the freight rate returns.

To be more specific, Panel B of Table 4.1 presents the descriptive statistics for the freight rate returns series. The mean returns for the tanker market are found to range between -2.6% and -0.2% p.a., while the standard deviation of these returns is found to range between 84.5% and 77.3% p.a. This is counterintuitive as economies of scale should imply that larger vessels earn higher returns due to cost savings. A possible reason for this observed anomaly may be that larger vessels are only able to operate on fewer routes because of cargo and port restrictions. This reduced flexibility in comparison with smaller vessels may mean that they were hit to a greater extent by the global economic slowdown and fall in spot freight rates, thus causing periods of unemployment and uncertainty in returns. Another reason for negative returns may be that the delivery of large numbers of brand new vessels, ordered at the peak of the market, led to an oversupply of vessels across the market, thus driving down freight rates and returns. A final possible explanation for these negative returns in the tanker

Figure 4.3 – Tanker Freight Rate Returns (13 January 1989 to 26 June 2009)





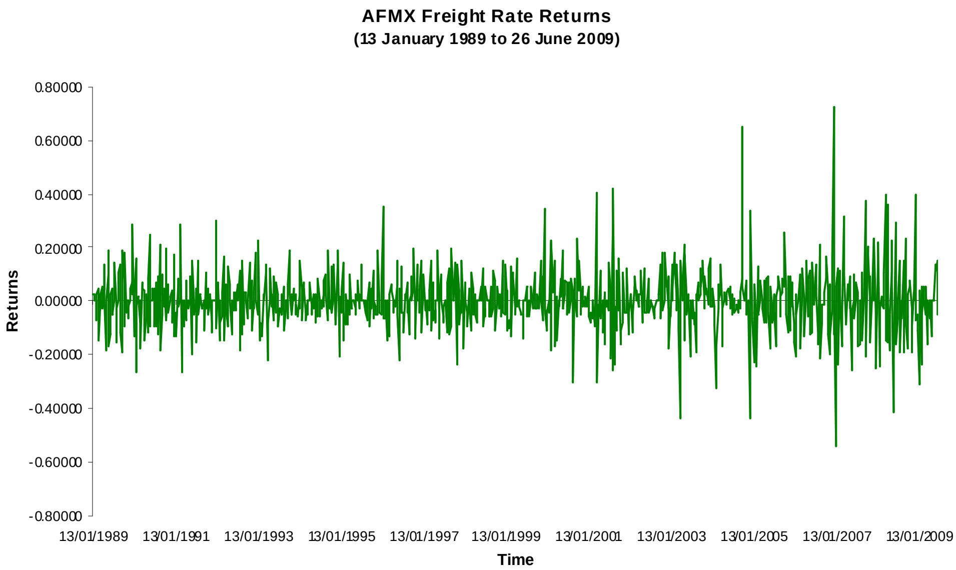
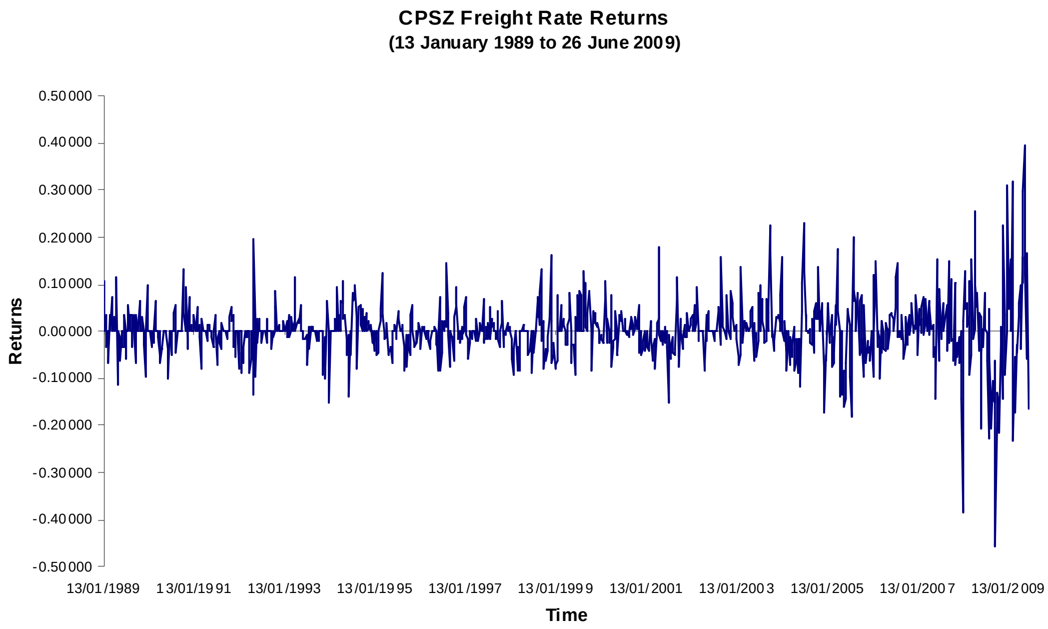
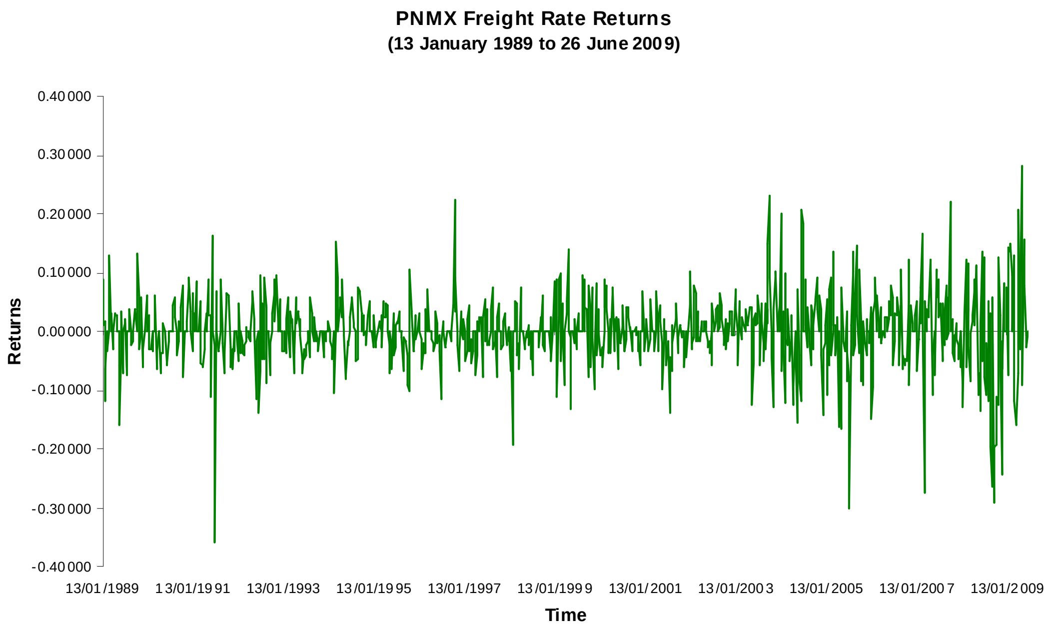


Figure 4.4 – Dry-Bulk Freight Rate Returns (13 January 1989 to 26 June 2009)





market may be the rescaling of the Worldscale measure, where these are rebased annually, therefore smoothing out the effect of large short-term peaks in the market. In the dry bulk market, the picture is once again reversed in that the expected economies of scale entail that CPSZ vessels earn returns of 6.3% p.a. that are much higher than the 3.9% p.a. earned by PNMX vessels. This result may differ from that for the tanker market in that the increased flexibility of bulk carriers in terms of the cargoes they carry means that they are less susceptible to price changes in a particular underlying cargo and any over-supply can be spread across different markets. Nevertheless, the same effects regarding the standard deviations of these returns are observed in the dry bulk market as in the tanker market, where decreased flexibility in terms of cargoes and ports means that larger vessels are more exposed to the prevailing freight rates on those routes on which they operate.

As a word of caution, one should note that mean returns in Panel B of Table 4.1 may be biased upwards as a result of taking arithmetic as opposed to log returns. The reason for using arithmetic returns, as opposed to log returns, is that, should one wish to incorporate an ARFIMA model into the mean equation for the volatility model, taking the differencing the logs of the respective series to generate standard returns would render the series stationary.

Having determined the basic statistics regarding location and dispersion of the freight rate returns, it is worth noting that all tanker data series exhibit significant positive skewness, ranging from 0.315 to 0.521, at all conventional levels of significance, implying that there is a higher probability observing positive than negative returns. This is logical in that the shape of the supply curve for shipping services is such that it is relatively flat for a large part, before bending steeply upwards indicating that in periods of high demand, huge returns can be expected, thus positively skewing the distribution. The picture is somewhat different in the dry bulk market, where although the distribution of the CPSZ returns is found to positively skewed, this figure is insignificant at all conventional levels of significance, while the distribution of the PNMX returns is found to be significantly negatively skewed, thus implying a greater probability of negative than positive returns. In addition to this, significant excess kurtosis was observed for both market and all data series, with levels of kurtosis ranging between 6.185 and 8.474 in the tanker market, and 8.018 and 11.882 in the dry bulk market. This implies that the probability of extreme positive or negative returns is greater than would be the case under a normally distributed returns series. Once again, this is logical given the shape of the supply curve and the long duration of cycles in the shipping markets. To conclude this analysis, Jarque-Bera statistics were calculated for all data series, where these unanimously rejected the null of

Figure 4.5 – Distribution of VLCC Returns vs. Normal Distribution

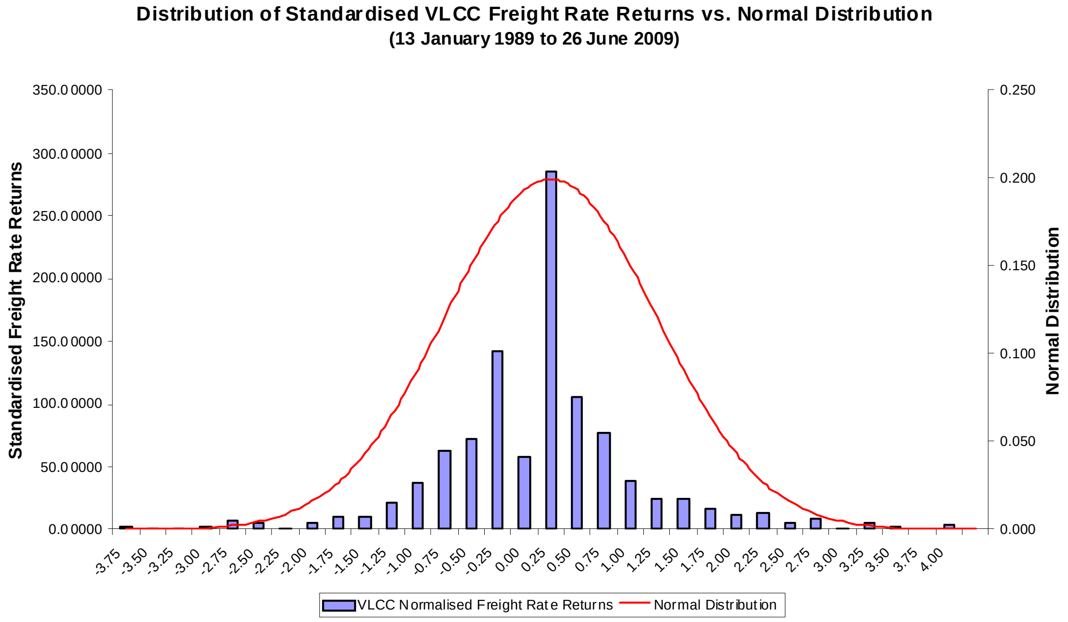


Figure 4.6 – Distribution of SZMX Returns vs. Normal Distribution

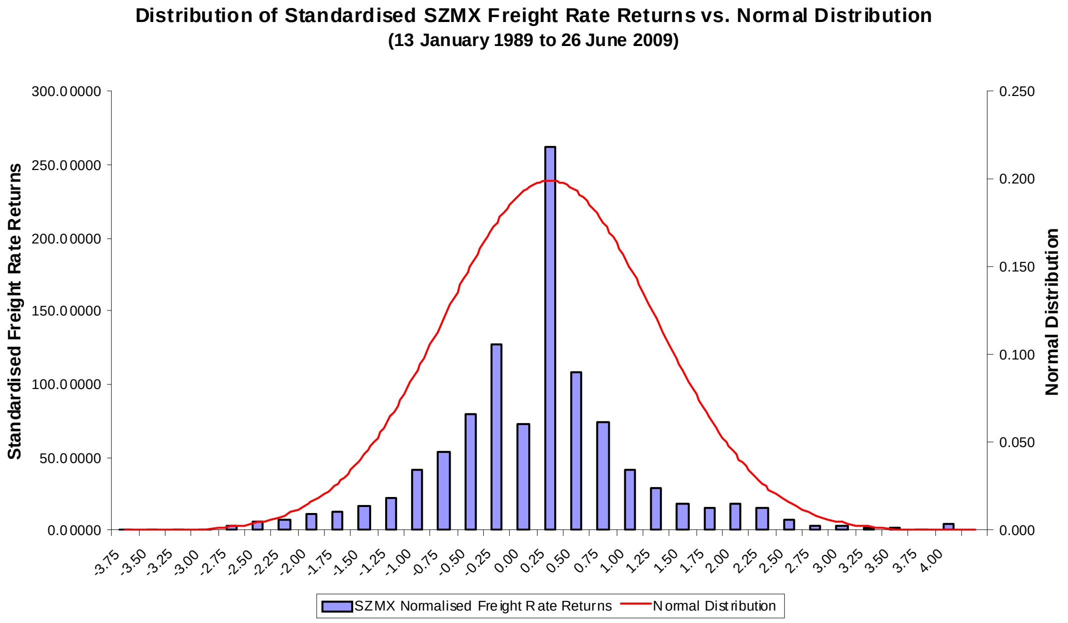


Figure 4.7 – Distribution of AFMX Returns vs. Normal Distribution

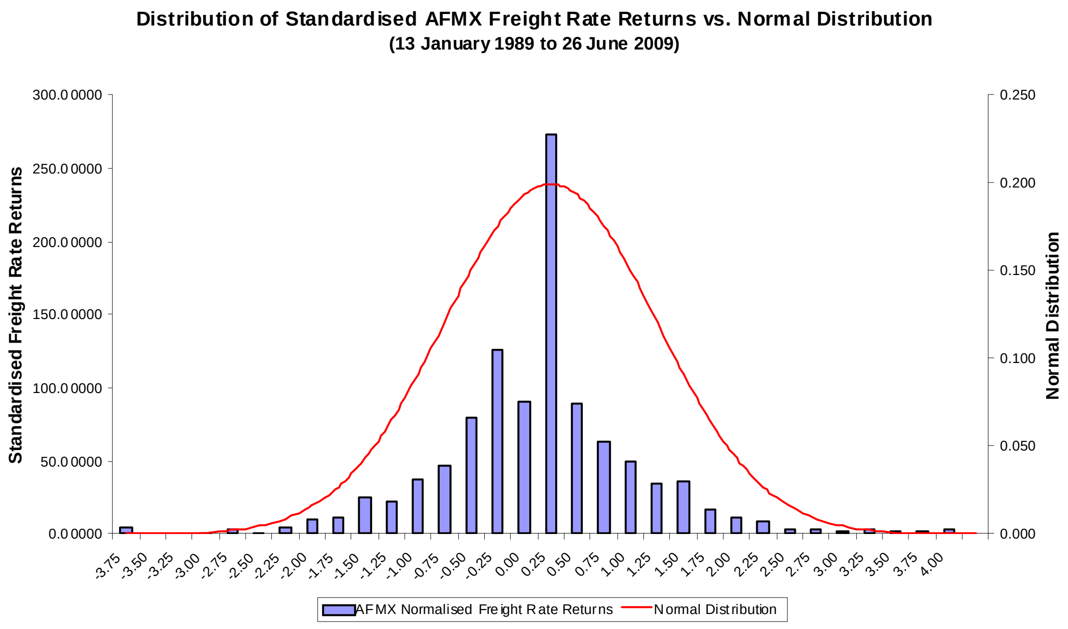


Figure 4.8 – Distribution of CPSZ Returns vs. Normal Distribution

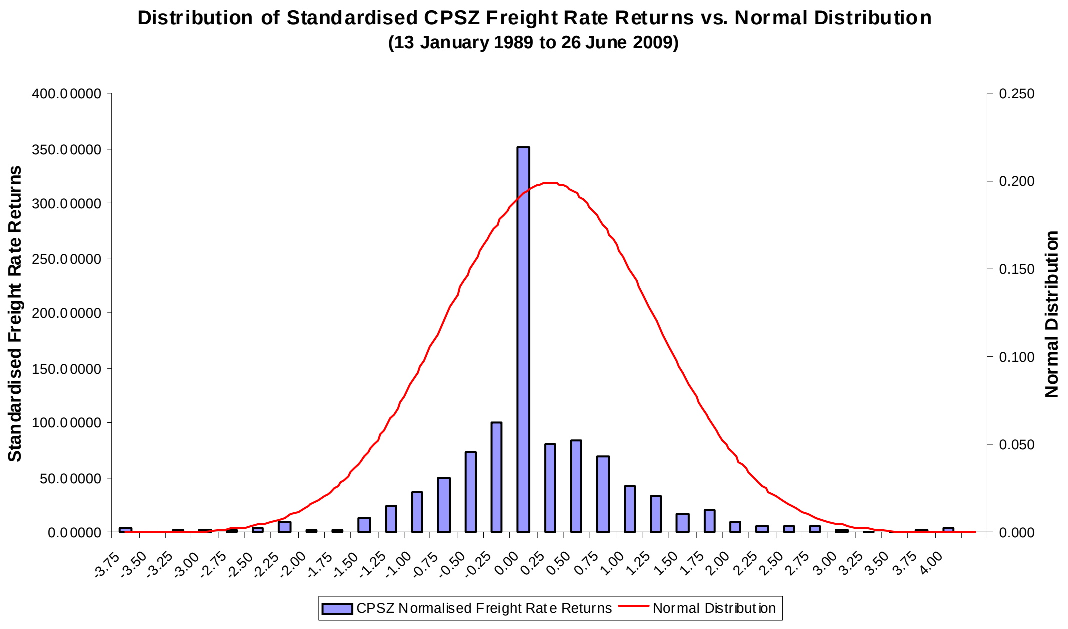
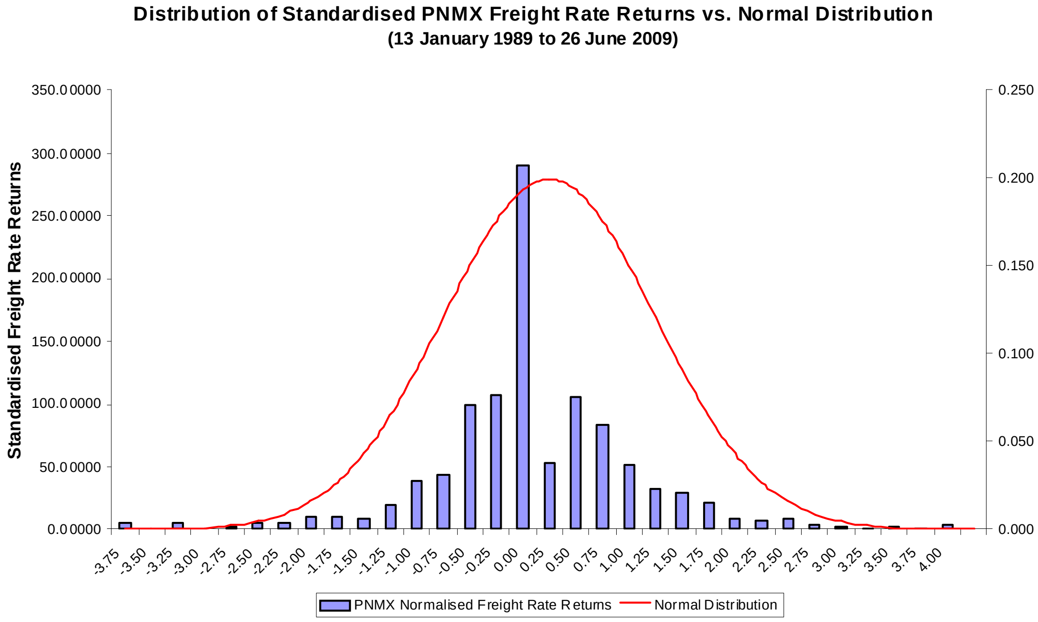


Figure 4.9 – Distribution of PNMX Returns vs. Normal Distribution



normality in favour of the alternative of non-normally distributed returns across all arket and data series and at all conventional levels of significance. One should note that to make a graphical comparison of the distribution of the returns against the standard normal distribution, the returns were standardised as follows:

$$sr_t = \frac{r_t - \mu}{\sigma} \qquad (4.2)$$

In Expression (4.2), sr_t and r_t denote the standardised returns and returns at time *t*, respectively, while μ and σ denote the mean return and the standard deviation of the returns, respectively. These results are also presented in graphical form in Figure 4.5 to Figure 4.9.

4.4 Ljung-Box Statistics and Conditional Moments

Having performed the descriptive analysis above, Ljung-Box statistics were calculated for the respective freight returns with respect to the first, second, third and fourth moments, the results of which are presented in Table 4.3. These were calculated on the standard returns of each data series and at the first, twelfth and twenty-fourth lags. Examining these results, and in particular Panel A, one can see that, at the first lag, there is no significant autocorrelation in spot freight rates for the SZMX and AFMX data series, whereas the test indicates that there is significant autocorrelation for all data series in the dry bulk market at all conventional levels of significance and for the VLCC data series at the 5% and 10% levels of significance. However, at the twelfth and twenty-fourth lags, the results become unanimous in rejecting the null of no autocorrelation at all conventional levels of significance.When looking at the second moment and the issue of conditional variance in Panel B of Table 4.3, the Ljung-Box are almost unanimous across markets, data series and lags in rejecting the null of no ARCH effects at all conventional levels of significance, where the only exception was the $Q^2(1)$ for the VLCC data series, where any ARCH effects are only significant at the 5% and 10% levels of significance. This implies the existence of conditional volatility, in addition to the suitability of GARCH-type models with respect to model the variance of the data series.

Panel C of the table presents the results for the examination of the existence of conditional skewness. This argues that skewness is not constant, as assumed in Bollerslev (1986)'s traditional GARCH model, Baillie, *et al.* (1996a)'s FIGARCH model or Brooks, *et al.* (2005)'s GARCHK model, but varies across time. The results indicate that in the tanker market, there is no evidence of conditional skewness at the first lag; as opposed to the dry bulk market; however, when looking at higher orders the results are unanimous in rejecting the null hypothesis of no conditional skewness in favour of the alternative at all conventional levels of significance, thus providing strong evidence of conditional skewness at greater lags. This would indicate the appropriateness of GARCH-type models that incorporate conditional skewness, such as the GARCHS model proposed by Harvey and Siddique (1999) or the GARCHSK model used here and originally proposed by León, *et al.* (2005).

Finally, Panel D of Table 4.3 presents the results for the tests for conditional kurtosis, where in a manner similar to that for conditional skewness, this would argue that kurtosis is time-varying as opposed to static as assumed in the traditional GARCH-type models, as well as in Harvey and Siddique (1999)'s GARCHS model. The results for these Ljung-Box tests are somewhat mixed in that the null hypothesis of no conditional kurtosis for the first lag cannot be rejected at any level of

Table 4.3 – Ljung-Box Tests for the Freight Rate Series

Panel A – Tests for Autocorrelation	**VLCC**	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
$Q(1)$	8.846	2.059	0.722	98.137	54.290
	(0.003)	(0.151)	(0.396)	(0.000)	(0.000)
$Q(12)$	65.942	78.062	97.525	173.910	66.061
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Q(24)$	83.206	106.926	167.642	200.813	91.068
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Panel B – Tests for Conditional Volatility	**VLCC**	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
$Q^2(1)$	6.316	16.520	7.945	60.645	33.227
	(0.012)	(0.000)	(0.005)	(0.000)	(0.000)
$Q^2(12)$	200.721	147.685	66.478	265.776	230.484
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Q^2(24)$	251.708	258.020	80.545	612.058	327.206
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Panel C – Tests for Conditional Skewness	**VLCC**	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
$Q^3(1)$	0.102	0.023	0.145	16.041	11.027
	(0.750)	(0.880)	(0.703)	(0.000)	(0.001)
$Q^3(12)$	182.547	44.359	48.365	34.849	41.065
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Q^3(24)$	187.337	63.773	52.679	102.206	54.058
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)

Panel D – Tests for Conditional Kurtosis	**VLCC**	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
$Q^4(1)$	0.000	0.188	0.044	8.152	1.835
	(0.992)	(0.665)	(0.834)	(0.004)	(0.176)
$Q^4(12)$	201.575	35.158	42.038	19.052	23.211
	(0.000)	(0.000)	(0.000)	(0.087)	(0.026)
$Q^4(24)$	202.052	60.729	42.527	108.766	25.803
	(0.000)	(0.000)	(0.011)	(0.000)	(0.363)

Note 1: See the respective notes from Table 4.1.

significance for the first lag of all but the CPSZ data series. At the level of the twelfth lag, the picture is somewhat clearer in that the tests for the tanker data series unanimously reject the null hypothesis at all conventional levels of significance, while those for the CPSZ and PNMX data series reject the null at the 10%, and 5% and 10%, levels of significance, respectively. At the twenty fourth lag, the waters are somewhat muddied again in that the null hypothesis can be rejected at any levels of significance for the VLCC, SZMX and CPSZ data series, can only be rejected at the

5% and 10% level of significance for the AFMX data series, and cannot be rejected at any levels of significance for the PNMX data series. Although the picture may not be as clear as desired, there does appear to be some evidence suggesting the appropriateness of a model that incorporates conditional kurtosis, such as Brooks, *et al.* (2005)'s GARCHK model or, once again, León, *et al.* (2005)'s GARCHSK model used in this research.

4.5 Concluding Comments

In summary, the unit root tests performed provided mixed results, which support the need for further examination as conducted in further chapters. In addition, it is proposed that these results lend support to the argument that, due to the nature of supply and demand in the shipping freight rates, these may illustrate non-stationary characteristics in the short-term, as supply is fixed in this period, however, as supply is able to adjust in the long-term so freight rates become mean reverting; a characteristic of fractionally integrated processes.

When looking at the specifics (second, third and fourth moments) of freight rate returns, there appears to be strong evidence of conditional volatility, skewness and kurtosis. This indicates the need to employ a model, such as the GARCHSK model, that incorporates conditional volatility, conditional skewness and conditional kurtosis when examining the second, third and fourth moments of these returns.

Having examined the characteristics of the data to be used in the empirical analysis, the thesis continues to begin this analysis. To this end, the following chapter examines the hypothesis that spot freight rate levels are fractionally integrated.

5 Dynamics of the First Moment

5.1 Introduction

Having outlined the methodologies to be used and the characteristics of the relevant data series in Chapter 3 and 4 of the thesis, respectively, this chapter seeks to provide an empirical analysis of the structure of the first moment of spot freight rates. The correct structure of freight rates is of great interest in that freight rates play a pivotal role and form the basis of almost every function, from the determination of the price of the transport service through to the price of second-hand vessels. Therefore, a correct model for freight rates is vital for all participants in the shipping market, from the ship-owners and charterers themselves, right on down through the market to ship-brokers, maritime lawyers and other auxiliary parties.

As discussed above and in Chapter 1, the argument that freight rates are fractionally integrated is supported by the characteristics of the supply and demand functions in the shipping markets. In the short-run, the supply curve is relatively inelastic as ship-owners may only reduce capacity through the laying-up of vessels and are unable to increase the overall supply capacity of the market due to the time needed to introduce new tonnage to the market. As a result, in the short-run, the supply of vessels is unable to adapt to short-term increases or decreases in demand, thus resulting in freight rates exhibiting non-stationary characteristics. However, when looking at the long-run, the picture changes quite dramatically in that ship-owners are now able to reduce the overall capacity in the face of falling long-term demand through the process of scrapping vessels and new tonnage may now be ordered thus enabling them to increase the overall capacity of supply should long-run demand increase. This would imply that freight rates in the long-term would illustrate stationary characteristics as freight rates revert to mean levels as demand and supply adjust. If one was to look at the overall process, this combination of short-run non-stationary characteristics and long-run stationarity would resemble a fractionally integrated series. This would entail that, although shocks may persist, a fact contrary to the intuition of the stationary hypothesis, they will eventually revert to the long-term mean, contrary to the non-stationary hypothesis.

This chapter therefore presents the results of various models used to ascertain certain of the dynamics of freight rates in the shipping freight markets, in particular as regards the first moment of freight rates and its characteristics. Using the data presented in Chapter 4, Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Fractionally Integrated Moving Average (ARFIMA) models are estimated to determine the true level of integration of freight rates. This is done to provide an alternative hypothesis to the partial equilibrium theorem, which argues that freight rates are stationary, such as, for example, Adland and Cullinane (2006) and Koekebakker, *et al.* (2006), amongst others, and the non-stationary and cointegration theorems, as outlined in, for example, Hale and Vanags (1989) and Berg-Andreassen (1997), which are presented in the literature review in Chapter 2.

As the ultimate aim of any empirical analysis in finance is to be able to forecast the future direction of the underlying series, following the initial analysis above, the sample was sub-divided into in-sample and out-of-sample periods, with the aim of forecasting the spot freight rate levels. The forecasting of freight rates has been a source of academic interest for a number of years, for instance Batchelor, *et al.* (2007) test the performance of the Autoregressive Integrated Moving Average (ARIMA), Vector Autoregressive (VAR) and various forms of Vector Equilibrium Correction models in predicting daily spot and Freight Forward Agreement (FFA) prices on Panamax Atlantic and Pacific routes. In terms of forecasting the spot prices, they find that all models outperform the random walk, with the possible exception of the ARIMA model on one of the routes. This ties in with the results from the study by Adland and Cullinane (2006), which reports success in forecasting spot freight rates with ARIMA models. In contradistinction, Kavussanos and Nomikos (1999) compare joint forecasts of spot freight rates and BIFFEX freight futures[17] for VECM, ARIMA, VAR and random walk models, finding that the VECM model gives the most accurate forecast of spot freight rates.[18] For this purpose, forecasts from ARMA, ARIMA and

---

[17] Baltic International Freight Futures Exchange (BIFFEX) contracts were futures contracts representing the expected future value of a respective freight index traded on the London International Financial Futures Exchange. However, due to sustained low trading interested, trading on these contracts was terminated in April 2002 (Kavussanos and Nomikos, 2003).

[18] A more extensive review of the forecasting literature can be found in Chapter 2 of the thesis.

ARFIMA models are performed, the accuracy of which are then compared using standard forecast evaluation techniques.

Following this, section two of the chapter provides a brief outline of the methodology used, while section three summarises the characteristics of the relevant spot freight rate levels. Section four presents the results for the models across the entire sample period, where the forecasting results and comparisons of accuracy between models are presented in section five, and section 6 concludes.

5.2 Methodology

This section provides a brief overview of the methodology used in this chapter, where a more detailed description can be found in Chapter 3 of the thesis. The shipping literature proposes two alternative hypotheses regarding the stationarity of freight rates, where one set of literature argues for stationarity and the other for non-stationarity. Depending on which hypothesis is being followed, the standard technique is to use either ARMA models, for the stationary hypothesis, or ARIMA models, for the non-stationary hypothesis, to provide a theoretical structure for the price series. A problem with using these methodologies is that, as is proposed here, the data is neither stationary nor non-stationary in the traditional sense, but is fractionally mean reverting in that the mean reversion process is delayed, a concept briefly discussed in Koekebakker, *et al.* (2006). This research, for the first time in shipping research, uses an ARFIMA model to model freight rates.

The ARMA model, in its most general form, is specified as follows:

$$\phi(L) y_t = \mu + \theta(L) u_t \qquad (5.1)$$

One should note that if the underlying process to the ARMA model is non-stationary, then this model becomes inappropriate. In order to rectify this problem, one could either make the process stationary by taking the appropriate number of first differences until it becomes stationary, or run an ARFIMA model.

The ARIMA model, in its most general form, is specified as follows:

$$(1-L)\phi(L) y_t = \theta(L) u_t \qquad (5.2)$$

In the ARIMA model, the integrated component is determined by the level of integration, which is equivalent to the number of times that one has to difference the series in order to achieve stationarity. ARMA and ARIMA models have found widespread use throughout the literature, with ARIMA models being used in a number of shipping papers, such as Kavussanos and Nomikos (1999), amongst others.

A third alternative is the ARFIMA model, where this can be expressed, in its most general form, as follows:

$$(1-L)^{d}\phi(L)y_t=\theta(L)u_t \qquad (5.3)$$

In Expressions (5.1) to (5.3), the lag operators for the autoregressive and moving average parameters in the respective models are $\phi(L)=1-\phi_1L-\phi_2L^2-\ldots-\phi_pL^p$ and $\theta(L)=\theta_1L+\theta_2L^2+\ldots+\theta_qL^q$, respectively. Furthermore, L in Expressions (5.2) and (5.3), denotes the normal lag operator. This provides the framework for the hypothesis that freight rates are fractionally integrated in that it allows one to estimate series in which the mean reversion process is delayed. The fractional difference parameter measures this delay, where the higher the value for *d*, where $-0.5<d<1$, the longer this delay in mean reversion.

The final piece of methodology to be covered is the Nielsen test for stationarity, developed by Nielsen (2005). This tests the hypothesis that the residuals from an ARFIMA model are integrated of order $d+\theta$, i.e. $I(d+\theta)$. The null hypothesis here is that $\theta=0$, which implies by taking the fractional difference, where *d* indicates the level of fractional integration, of the residuals from that ARFIMA model, this series will become stationary. An advantage of this test over similar alternatives is that this test is a time domain test, as opposed to a frequency domain test, where this has been suggested to provide superior results in terms of the finite sample properties, as discussed in Tanaka (1999).

One can now move on to look at the methodologies that will be used to evaluate the accuracy of the forecasts, and to compare these across models. To this end, five methods are used, namely: 1) the Mean Absolute Error (MAE), 2) the Mean Absolute Percentage Error (MAPE), 3) the Percentage of Correction Direction Predicted

(CDIR), 4) the Root Mean Sqaured Error, and 5) the Theil's U, also known as Theil's inequality coefficient, as outlined in Theil (1958). One should note that Theil's U, is constructed such that the statistic is confined to lie between zero and one. The closer the statistic is zero, the better the forecast, such that if the statistic is equal to zero, there is a perfect forecast as the actual and forecast values are equal; whereas, if the statistic is equal to one, the forecast value is in no way close to the actual value.

The first of these, i.e. the MAE, is calculated in the following manner:

$$MAE = \frac{1}{M}\sum_{i=1}^{M}\left|\left(y_i^a - y_i^f\right)\right| \qquad (5.4)$$

Following this, one can calculate the MAPE as follows:

$$MAPE = 100\times\left\{\frac{1}{M}\sum_{i=1}^{M}\left|\left(\frac{y_i^a - y_i^f}{y_i^a}\right)\right|\right\} \qquad (5.5)$$

In order to calculate the CDIR, one must first determine whether the sign of the forecast matches the sign of the actual values, i.e. when an increase (decrease) in the forecast matches an increase (decrease) in the actual values, where if this occurs, it is denoted C_i here, therefore:

$$\text{CDIR} = \frac{1}{M}\sum_{i=1}^{M}C_i \qquad (5.6)$$

Moving on, the RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(y_i^a - y_i^f\right)^2} \qquad (5.7)$$

Finally, Theil's U may be calculated as follows:

$$\text{Theil's U} = \frac{\sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(y_i^a - y_i^f\right)^2}}{\sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(y_i^a\right)^2} + \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(y_i^f\right)^2}} \qquad (5.8)$$

In Expressions (5.4) to (5.8), y_i^a denotes the actual observed value at time *i*, y_i^f denotes the forecasted value at time *i*, and *M* denotes the forecast horizon.

Having laid out the foundations of the methodologies to be used, one can now move on to examine the data to be used and estimate the models described.

5.3 Description of Data

Having established the methodology to be used in this empirical analysis in the previous section, this section provides a very brief summary of the relevant descriptive statistics for the respective data series, where the complete analysis can be found in Chapter 4 of the thesis. The data set used in this chapter comprises five data series of spot freight rate for five different vessel types across the tanker and dry-bulk sectors, namely VLCC, Suezmax (SZMX) and Aframax (AFMX) tankers, and Capesize (CPSZ) and Panamax (PNMX) dry-bulk vessels. The sample extends from 13 January 1989 to 26 June 2009, thus comprising 1,068 observations, where all data was collected from Clarksons Shipping Intelligence Network. To enable ex-post forecasts to be made, each series is further sub-divided into an in-sample period, extending from 13 January 1989 to 26 September 2003 and thus comprising 768 observations, and an out-of-sample period, extending from 3 October 2003 to 26 June 2009, giving a forecast horizon of 300 observations, which equates to roughly one third of the sample.

The descriptive statistics for the respective data series indicate that there is a size effect in terms of the mean freight rates and standard deviations. In addition, all data series exhibit large and significant positive skewness as well as significant excess kurtosis, a fact supported by the fact that the series are found not follow a normal distribution. To conclude, the Ljung-Box statistics indicate the presence of autocorrelation in the data series up to the 24th lag, across all data series, with the exception of the 1st lag of the Suezmax and Aframax data series, thus illustrating the appropriateness of autoregressive models for modelling the structure of freight rates in the shipping markets.

One should note that these models were also run on the natural logarithms of the respective data series in order to take account of the fact that freight rates can never be negative. One should note that the results from these estimations did not differ significant from those presented here[19].

[19] The results from these estimations are not presented here due to space constraints and are available from the author upon request.

5.4 Empirical Results for the In-Sample Period

Having outlined the methodology to be used and established the characteristics of the data in the sections above, this section presents the empirical results from the estimation of the ARMA, ARIMA and ARFIMA models over the entire sample period. This is done with the intention of throwing light on the structure of the first moment of the freight rate dynamics, i.e. the mean equation, and, in particular, on the issue of the level of integration in freight rate levels.

One should note that, prior to these estimations, Augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and KPSS tests for unit roots, where these are attributed to Dickey and Fuller (1981), Phillips and Perron (1988) and Kwiatkowski, *et al.* (1992), respectively, are performed on the five spot freight rates levels, where the results of these tests are presented in Panel A of Table 4.2 in Chapter 4, and re-presented for convenience in Table 5.1. The results for these three unit root tests are somewhat mixed. In the case of the tanker data series, the results of the ADF and PP tests indicate that one can reject the null hypothesis of a unit root at all conventional levels of significance, while, in contrast, the results for the KPSS test indicate that one should reject the null hypothesis of stationarity at all conventional levels of significance. When looking at the results for the dry-bulk market the picture is somewhat clearer with all three tests agreeing that freight rates are non-stationary, with the exception of the PP test on the PNMX data where one can reject the null hypothesis of a unit root at the 10% level of significance.[20]

Beginning the main analysis, in order to test which of the hypotheses regarding the order of integration for freight rates holds, ARMA, ARIMA and ARFIMA models are estimated. Beginning with the ARMA and ARIMA models, the appropriate lag structure for these models is established by estimating models with different lag structures using Maximum Likelihood Estimation, where the lag structure varies from an $\mathrm{ARMA}(0;0)$ to an $\mathrm{ARMA}(3;3)$, in the cases of the ARMA models, and from an $\mathrm{ARIMA}(0;1;0)$ to an $\mathrm{ARIMA}(3;1;3)$, for the ARIMA models. The best lag structure for each respective data series and model-type is then established using the respective

[20] These results are discussed in more detail in Chapter 4 of the thesis.

Table 5.1 – Results for Unit Root Tests of Spot Freight Rate Levels

	VLCC	SZMX	AFMX	CPSZ	PNMX
ADF Test	-5.064	-4.597	-4.232	-2.116	-2.196
t-value at 1%	-(3.436)	-(3.436)	-(3.436)	-(3.436)	-(3.436)
t-value at 5%	-(2.864)	-(2.864)	-(2.864)	-(2.864)	-(2.864)
t-value at 10%	-(2.568)	-(2.568)	-(2.568)	-(2.568)	-(2.568)
Phillips-Perron Test	-5.163	-5.166	-4.998	-2.827	-1.918
t-value at 1%	-(3.436)	-(3.436)	-(3.436)	-(3.436)	-(3.436)
t-value at 5%	-(2.864)	-(2.864)	-(2.864)	-(2.864)	-(2.864)
t-value at 10%	-(2.568)	-(2.568)	-(2.568)	-(2.568)	-(2.568)
KPSS Test	1.321	1.946	2.233	1.959	1.819
t-value at 1%	(0.739)	(0.739)	(0.739)	(0.739)	(0.739)
t-value at 5%	(0.463)	(0.463)	(0.463)	(0.463)	(0.463)
t-value at 10%	(0.347)	(0.347)	(0.347)	(0.347)	(0.347)

Note 1: VLCC denotes the weekly spot freight rates for a 270,000 DWT VLCC tanker carrying crude oil from Ras Tanura (Saudi Arabia) to Rotterdam (Netherlands).
SZMX denotes the weekly spot freight rates for a 130,000 DWT Suezmax tanker carrying crude oil from Bonny (Nigeria) to off the coast of Philadelphia (USA).
AFMX denotes the weekly spot freight rates for an 80,000 DWT Aframax tanker carrying crude oil from Sullom Voe (UK) to Bayway (USA).
CPSZ denotes the weekly spot freight rates for a 145,000 DWT Capesize bulk-carrier carrying iron ore from Tubarao (Brazil) to Rotterdam (Netherlands).
PNMX denotes the weekly spot freight rates for a 55,000 DWT Panamax bulk-carrier carrying grain from the Hampton Roads (USA) to Antwerp-Rotterdam-Amsterdam (Benelux).

Note 2: The sample period for the data used for this table extends from 13 January 1989 to 26 June 2009, with a total of 1,068 observations.

Note 3: The data used for this table is all sourced from the Clarkson Shipping Intelligence Network (www.clarskons.net).

Note 4: ADF test denotes the results from conducting the Augmented Dickey-Fuller test for stationarity (Dickey and Fuller (1981)).
Phillips-Perron test denotes the results from conducting the Phillips-Perron test for stationarity (Phillips and Perron (1988)).
KPSS test denotes the results from conducting the Kwiatkowski, Phillips, Schmidt and Shin test for stationarity (Kwiatkowski, *et al.* (1992)).

Note 5: The null hypotheses for the ADF and Phillips-Perron tests are that the data series contains a unit root, i.e. is non-stationary, whereas the null hypothesis for the KPSS test is that the data series is stationary.

Note 6: The figures in parentheses denote the critical *t*-values for the respective data series at the given levels of significance.

Log-Likelihoods, Akaike Information Criteria (AIC) and Schwartz-Bayesian Information Critreria (SBIC), where the latter two methods were first developed by Akaike (1974) and Schwarz (1978), respectively.

In the case of the ARMA model, the best lag specification, across all data series, is found to be an $\text{ARMA}(1;1)$, the results for which are presented in Table 5.2. One should note that this implies that, should freight rates be stationary, as implied by an

Table 5.2 – Results for Final ARMA Models of Spot Freight Rate Levels

$$\phi(L)\,y_t = \mu + \theta(L)\,u_t$$

	VLCC ARMA (1;1)	SZMX ARMA (1;1)	AFMX ARMA (1;1)	CPSZ ARMA (1;1)	PNMX ARMA (1;1)
Q(24)	8844.083	10411.634	11103.872	17435.569	19196.988
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	64.123	108.601	142.604	11.771	11.545
	(0.000)	(0.000)	(0.000)	(0.005)	(0.002)
AR(1)	0.942	0.933	0.935	0.990	0.990
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
MA(1)	0.089	0.106	0.085	0.252	0.162
	(0.006)	(0.001)	(0.009)	(0.000)	(0.000)
LL	-3900.947	-4499.333	-4670.215	-1629.667	-1486.964
AIC	4.477	5.599	5.919	0.220	-0.048
SBIC	4.497	5.618	5.938	0.239	-0.028
Sum AR	0.942	0.933	0.935	0.990	0.990
Sum MA	0.089	0.106	0.085	0.252	0.162

Note 1: See notes from Table 5.1.
Note 2: LL, AIC and SBIC denote the log-likelihoods, Akaike Information Criteria (Akaike (1974)) and Schwartz-Bayesian Information Critera (Schwarz (1978)).
Note 3: The figures in parentheses denote the respective *p*-values for the null hypothesis that the coefficient is equal to zero.

Table 5.3 – Results for Final ARIMA Models of Spot Freight Rate Levels

$$(1-L)\phi(L)\,y_t=\theta(L)\,u_t$$

	VLCC ARIMA (1;1;2)	SZMX ARIMA (1;1;1)	AFMX ARIMA (1;1;2)	CPSZ ARIMA (1;1;1)	PNMX ARIMA (1;1;1)
Q(24)	94.339	104.381	207.054	316.864	129.421
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
AR(1)	0.864	-0.310	0.515	0.687	-0.714
	(0.000)	(0.350)	(0.000)	(0.000)	(0.000)
MA(1)	-0.822	0.387	-0.511	-0.481	0.853
	(0.000)	(0.230)	(0.000)	(0.000)	(0.000)
MA(2)	-0.111	---	-0.180	---	---
	(0.001)	(---)	(0.000)	(---)	(---)
LL	-3905.064	-4511.362	-4660.085	-1622.091	-1481.669
AIC	4.492	5.628	5.908	0.207	-0.056
SBIC	4.518	5.648	5.934	0.227	-0.037
Sum AR	0.864	0.310	0.515	0.687	0.714
Sum MA	0.933	0.387	0.691	0.481	0.853

Note 1: See notes from Table 5.2.

Table 5.4 – Results for ARFIMA (0; d; 0) Models of Spot Freight Rates

$$(1-L)^d y_t = u_t$$

	VLCC ARFIMA (0;d;0)	SZMX ARFIMA (0;d;0)	AFMX ARFIMA (0;d;0)	CPSZ ARFIMA (0;d;0)	PNMX ARFIMA (0;d;0)
d	0.500	0.499	0.499	0.500	0.500
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LL	-4138.234	-4664.164	-4812.768	-2238.102	-2011.735
AIC	7.751	8.736	9.015	4.193	3.769
SBIC	7.752	8.737	9.016	4.194	3.770

Note 1: See notes from Table 5.2.

ARMA model, then the previous period's freight rates and white noise disturbances do indeed have an impact on the current level of freight rates. In the case of the ARIMA model, the lag structures which best described the underlying series are an $\mathrm{ARIMA}(1;1;2)$ for the VLCC and Aframax data series, and an $\mathrm{ARIMA}(1;1;1)$ for the Suezmax, Capesize and Panamax data series, the results for which are described in Table 5.3. This would suggest that the difference between the current and previous periods' freight rates does have an impact on the current freight rate level, and that for the Suezmax, Capesize and Panamax data series, only the previous period's white noise disturbance term impacts on this current freight rate, whereas, for the VLCC and Aframax data series, the previous two periods' errors have an impact on the current freight rate. One should note that the constraint that the sum of the autoregressive and the sum of the moving average coefficients must be less than one is met for all cases.

As far as the ARFIMA models are concerned, given the mixed results from the unit roots tests presented in Table 5.1, above, a preliminary $\mathrm{ARFIMA}(0;d;0)$ model is run on each data series to determine if the data series were fractionally integrated, the results for which can be found in Table 5.4, where this specification for the models was chosen so as to ensure that any real lag dynamics did not interfere with this process. This process is in essence merely checking whether the series are fractionally integrated white noise. This model would be expressed as follows:

$$(1-L)^d = u_t \qquad (5.9)$$

In Expression (5.9), *L* denotes the lag operator, while *d* denotes the order of fractional integration. The results for these models are unanimous in that all five data series are found to be fractionally integrated of the order $d = 0.5$, thus providing a preliminary indication that freight rates follow a fractionally integrated process. This being said, this result could not be confirmed until the stationarity of the corresponding residuals series has been established. To this end, the Nielsen test, first developed by Nielsen (2005), is used, where this tests whether the residual series is integrated of order $d + \theta$, where the null hypothesis is that $\theta = 0$.[21] Therefore, this would correspond to a stationary series should the estimated value for the *d*-parameter be close to zero under the null hypothesis that $\theta = 0$.

The results of these tests are presented in Panel A of Table 5.5, where in essence what this test does is establish whether the residuals of each data series are stationary or not. To establish this, the null hypothesis that $d + \theta = d$ is tested, where if this holds, i.e. $\theta = 0$, then the residuals will be integrated of order d. The procedure involves changing the values of d, where in this case the value of *d* changes by 0.01, and then calculates the corresponding *p*-value for the respective value of *d*. The range of values given for each of the respective levels of significance represents the range of values for *d* at which one cannot reject the null hypotheses. One can thus see that these results indicate that the residuals for all series, with the exception of those for the Panamax data series, are found to be non-stationary. This is indicative of two facts: first, that the data series are not white noise and therefore do indeed contain some information; and second, that there is a probably a need to include some lag dynamics to the model.

In order to address the latter issue, ARFIMA models with lag structures are now estimated. However, due to the complexity of the computations, as well as the fact that the order of the autoregressive coefficients for the ARMA and ARIMA models was never more than one, while the order of the moving average coefficients was more than one in only two cases, $\text{ARFIMA}(0;d;1)$, $\text{ARFIMA}(1;d;0)$ and $\text{ARFIMA}(1;d;1)$ models were estimated. As with the ARMA and ARIMA models, the correct lag specification for each data series was selected on the basis of the

[21] This test is discussed in more detail in Chapter 3 of the thesis.

Table 5.5 – Results for Nielsen (2005) Tests on Residuals of ARFIMA Models

Panel A – Results for Preliminary ARFIMA (0;d;0) Models					
	VLCC	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
Change in *d* parameter	0.01	0.01	0.01	0.01	0.01
Value of *d* with highest *p*-value	0.50	0.41	0.38	0.71	0.00
Values of *d* where cannot reject $H_0 : \theta = 0$ at 1%	0.43 < d < 0.57	0.34 < d < 0.48	0.31 < d < 0.45	0.64 < d < 0.77	0.00 < d < 0.05
Values of *d* where cannot reject $H_0 : \theta = 0$ at 5%	0.44 < d < 0.55	0.36 < d < 0.46	0.33 < d < 0.43	0.66 < d < 0.75	0.00 < d < 0.03
Values of *d* where cannot reject $H_0 : \theta = 0$ at 10%	0.45 < d < 0.54	0.37 < d < 0.45	0.34 < d < 0.42	0.67 < d < 0.75	0.00 < d < 0.02

Panel B – Results for Final ARFIMA (p;d;q) Models					
	VLCC	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
Change in *d* parameter	0.01	0.01	0.01	0.01	0.01
Value of *d* with highest *p*-value	0.00	0.01	0.01	0.00	0.00
Values of *d* where cannot reject $H_0 : \theta = 0$ at 1%	0.00 < d < 0.05	0.00 < d < 0.07	0.00 < d < 0.07	0.00 < d < 0.05	0.00 < d < 0.05
Values of *d* where cannot reject $H_0 : \theta = 0$ at 5%	0.00 < d < 0.03	0.00 < d < 0.05	0.00 < d < 0.05	0.00 < d < 0.04	0.00 < d < 0.04
Values of *d* where cannot reject $H_0 : \theta = 0$ at 10%	0.00 < d < 0.02	0.00 < d < 0.04	0.00 < d < 0.05	0.00 < d < 0.03	0.00 < d < 0.02

Note 1: See notes from Table 5.2.

Note 2: $H_0 : \theta = 0$, i.e. the given value of *d* for the respective data series is the true value of *d*, where the series would be stationary if $d = 0$ and $\theta = 0$.

Table 5.6 – Results for ARFIMA (p; d; q) Models of Spot Freight Rates

	VLCC ARFIMA (1;d;0)	SZMX ARFIMA (1;d;0)	AFMX ARFIMA (1;d;0)	CPSZ ARFIMA (1;d;1)	PNMX ARFIMA (1;d;1)
AR(1)	0.183	0.205	0.156	0.514	0.471
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
MA(1)	---	---	---	-0.268	-0.334
	(---)	(---)	(---)	(0.000)	(0.000)
d	0.855	0.809	0.831	0.879	0.929
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LL	-948.494	-1255.488	-1337.915	-342.800	-260.112
AIC	1.782	2.357	2.511	0.649	0.495
SBIC	1.785	2.360	2.514	0.653	0.498

Note 1: See notes from Table 5.2.

Log-Likelihood, AIC and SBIC criteria. For the tanker series, the best model specification was found to be an $\text{ARFIMA}(1;d;0)$, while for the dry-bulk series, this was an $\text{ARFIMA}(1;d;1)$, where the results for all these series are summarised in Table 5.6.

The results suggest that only the previous period's freights rates have an impact on the current freight rate levels in the tanker sector, whereas, in the case of the dry bulk-sector, both the previous period's freight rate levels and the white noise disturbance terms have an impact on the prevailing freight rate levels. Furthermore, all series are found to have $0 < d < 1$, indicating fractional integration, where the values for *d* range from 0.831, for the Aframax series, to 0.929, for the Panamax data series. It appears from these results that while all series are fractionally integrated, the dry-bulk series appear to exhibit a higher order of fractional integration than the tanker series, which may be as a result of the fact that there is more inter-changeability in terms of the cargoes that these vessels carry, whereas, the cargo is standard across tankers. This would imply that imbalances in supply and demand in the dry-bulk sector would persist for longer as vessels may change routes. As before, however, no conclusions regarding fractional integration in the series can be drawn until the residuals have been tested for stationarity. To this end, the results of the Nielsen test for the final ARFIMA models are presented in Panel B of Table 5.5. These indicate that one cannot reject the null hypothesis, across all data series, that the true value of *d* ranges between 0.00 and 0.07, at the 1%, 5% or 10% levels of significance. In addition, the value of d with the highest p-value is 0.00 for the VLCC, Capesize and Panamax data series and 0.01 for the Suezmax and Aframax series. The results thus provide strong evidence as to the stationarity of the residuals.

Having estimated the models, these models are then compared using the same methods as for the forecasts, where the results of these in-sample comparisons are presented in Table 5.7. The results for the VLCC data series suggest that although the $\mathrm{ARMA}(1;1)$ provides the best MAE and RMSE, the $\mathrm{ARFIMA}(1;d;0)$ model provides the best MAPE and CDIR, while one is unable to distinguish between the models on the basis of Theil's U. As far as the Suezmax data series is concerned, the $\mathrm{ARIMA}(1;1;1)$ model provides the best MAE, MAPE and CDIR as opposed the $\mathrm{ARFIMA}(1;d;0)$ model, which provides the best RMSE, and one is once again unable to make a comparison between the models on the basis of Theil's U. To conclude the comparison of the tanker series, the results for the Aframax data series indicate that the $\mathrm{ARIMA}(1;1;2)$ model provides the best MAE and MAPE, while the $\mathrm{ARFIMA}(1;d;0)$ provides the best CDIR and RMSE, where one is unable to distinguish between the $\mathrm{ARIMA}(1;1;2)$ and $\mathrm{ARFIMA}(1;d;0)$ models on the basis of the Theil's U, although this statistic indicates that both these models fare better than the $\mathrm{ARMA}(1;1)$.

In terms of the dry-bulk sector, the results for the Capesize data series suggest that the $\mathrm{ARFIMA}(1;d;1)$ model fares best across all statistics, with the exception of the CDIR, where this is equal for the $\mathrm{ARFIMA}(1;d;1)$ and $\mathrm{ARIMA}(1;1;1)$ models, although once again these are better than the $\mathrm{ARMA}(1;1)$ model. To conclude this comparison, the results from the Panamax data series suggest that the $\mathrm{ARFIMA}(1;d;1)$ model provides the best MAPE, while the $\mathrm{ARMA}(1;1)$ gives the best MAE and CDIR figures, and the $\mathrm{ARIMA}(1;1;1)$ model the best RMSE and Theil's U statistics.

Although the mixed results from this comparison are somewhat disappointing, they do provide something of an answer to the hypothesis in that, except for the case of a few

Table 5.7 – In-Sample Model Comparison Results

Panel A - VLCC In-Sample Model Comparison	ARMA	ARIMA	ARFIMA
Mean Absolute Error	5.258	5.294	5.273
Mean Absolute Percentage Error	7.773%	7.766%	7.713%
Percentage of Correct Direction Predicted	57.371%	56.995%	57.934%
Root Mean Squared Error	9.365	9.434	9.399
Theil's U	0.066	0.066	0.066

Panel B - SZMX In-Sample Model Comparison	ARMA	ARIMA	ARFIMA
Mean Absolute Error	9.835	9.707	9.809
Mean Absolute Percentage Error	8.237%	8.027%	8.125%
Percentage of Correct Direction Predicted	53.333%	53.803%	52.958%
Root Mean Squared Error	16.416	16.382	16.360
Theil's U	0.069	0.069	0.069

Panel C - AFMX In-Sample Model Comparison	ARMA	ARIMA	ARFIMA
Mean Absolute Error	11.481	11.297	11.428
Mean Absolute Percentage Error	7.567%	7.320%	7.404%
Percentage of Correct Direction Predicted	50.704%	50.798%	51.549%
Root Mean Squared Error	19.269	19.156	19.153
Theil's U	0.063	0.062	0.062

Panel D - CPSZ In-Sample Model Comparison	ARMA	ARIMA	ARFIMA
Mean Absolute Error	0.487	0.475	0.473
Mean Absolute Percentage Error	3.854%	3.643%	3.630%
Percentage of Correct Direction Predicted	58.122%	59.531%	59.531%
Root Mean Squared Error	1.115	1.108	1.102
Theil's U	0.040	0.040	0.040

Panel E - PNMX In-Sample Model Comparison	ARMA	ARIMA	ARFIMA
Mean Absolute Error	0.491	0.492	0.493
Mean Absolute Percentage Error	3.921%	3.943%	3.903%
Percentage of Correct Direction Predicted	58.873%	57.277%	57.653%
Root Mean Squared Error	0.975	0.971	0.979
Theil's U	0.037	0.036	0.037

Note 1: See notes from Table 5.2.
Note 2: Figures for the MAE are measured in Worldscale units for the VLCC, Suezmax and Aframax data series, and $ / tonne for the Capesize and Panamax data series.
Note 3: Comparative AIC, SBIC and Log-Likelihood results can be found in Table 5.2, Table 5.3 and Table 5.5, respectively.

statistics, the ARIMA and ARFIMA models outperform the ARMA specification, thereby suggesting that freight rates do not follow a stationary process. In addition, the fact that the data is somewhat limited, both in terms of length and frequency, may also suggest that, should the sample length have been able to be greater, the ARFIMA model may have been preferred as Maddala and Kim (1998) point out that there is a greater chance of mean reversion as the sample length increases.

5.5 Forecasting Spot Freight Rate Levels

Having examined the in-sample performance of the different models, and provided a preliminary conclusion as to the order of integration of spot freight rate levels, this chapter now focuses on the forecasting performance of the ARMA, ARIMA and ARFIMA models, in terms of forecasting spot freight levels in the tanker and dry-bulk sector. Having established the best specifications for each of these models in the section above, each of these models is used to create one-period ahead forecasts of freight rate levels for each data series, over the out-of-sample period, i.e. 3 October 2003 to 26 June 2009. In order to make this comparison, corresponding MAE, MAPE, CDIR, RMSE and Theil's U statistics were calculated for each model type across each data series, the results of which are presented in Table 5.8.

Beginning with the tanker sector, the results for the VLCC data series indicate that, on the basis of the MAE, MAPE, RMSE and Theil's U metricis, the $\mathrm{ARMA}(1;1)$ model provides the best forecasts, whereas the CDIR indicates that the $\mathrm{ARIMA}(1;1;2)$ provides the best forecast of direction. Continuing, the results for the Suezmax data series illustrate that $\mathrm{ARMA}(1;1)$ once again provides the best MAE, MAPE and RMSE figures, although the $\mathrm{ARIMA}(1;1;1)$ model provides the best Theil's U figure and one is unable to distinguish between the models using the CDIR. To conclude the comparison of the tanker series, the results for the Aframax data series show that the $\mathrm{ARIMA}(1;1;2)$ provides the best forecasts, on the basis of the MAE, RMSE and Theil's U metrics, even though the $\mathrm{ARMA}(1;1)$ provides the best MAPE and one is unable to distinguish between the $\mathrm{ARMA}(1;1)$ and $\mathrm{ARIMA}(1;1;2)$ on the basis of the CDIR. There therefore seems to be something of a size effect in the tanker market, where freight rates for larger vessels exhibit more persistence, in terms of their autocorrelation. This is logical in that these vessels are limited as to the routes on which they can operate and therefore are more susceptible to market conditions.

Table 5.8 – Comparison of Forecasting Performance

Panel A - VLCC Out-of-Sample Model Comparison			
	ARMA	**ARIMA**	**ARFIMA**
Mean Absolute Error	9.425	9.708	9.661
Mean Absolute Percentage Error	10.998%	11.141%	11.125%
Percentage of Correct Direction Predicted	66.221%	57.333%	58.333%
Root Mean Squared Error	15.292	15.532	15.474
Theil's U	0.081	0.082	0.081

Panel B - SZMX Out-of-Sample Model Comparison			
	ARMA	**ARIMA**	**ARFIMA**
Mean Absolute Error	19.708	19.753	20.506
Mean Absolute Percentage Error	12.566%	12.654%	13.062%
Percentage of Correct Direction Predicted	50.167%	49.667%	50.333%
Root Mean Squared Error	27.518	27.532	28.414
Theil's U	0.084	0.084	0.086

Panel C - AFMX Out-of-Sample Model Comparison			
	ARMA	**ARIMA**	**ARFIMA**
Mean Absolute Error	18.804	18.674	18.825
Mean Absolute Percentage Error	9.321%	9.411%	9.345%
Percentage of Correct Direction Predicted	55.853%	56.000%	55.000%
Root Mean Squared Error	30.289	29.696	30.242
Theil's U	0.074	0.073	0.074

Panel D - CPSZ Out-of-Sample Model Comparison			
	ARMA	**ARIMA**	**ARFIMA**
Mean Absolute Error	1.293	1.280	1.295
Mean Absolute Percentage Error	6.180%	6.074%	6.164%
Percentage of Correct Direction Predicted	58.528%	61.000%	59.667%
Root Mean Squared Error	2.079	2.061	2.081
Theil's U	0.043	0.042	0.043

Panel E - PNMX Out-of-Sample Model Comparison			
	ARMA	**ARIMA**	**ARFIMA**
Mean Absolute Error	1.212	1.213	1.222
Mean Absolute Percentage Error	6.154%	6.150%	6.160%
Percentage of Correct Direction Predicted	54.000%	51.333%	51.667%
Root Mean Squared Error	1.776	1.775	1.783
Theil's U	0.040	0.040	0.040

Note 1: See notes from Table 5.2.
Note 2: Figures for the MAE are measured in Worldscale units for the VLCC, Suezmax and Aframax data series, and $ / tonne for the Capesize and Panamax data series.

Changing focus to the dry-bulk sector, the results are more uniform here in that the $\text{ARIMA}(1;1;1)$ model is preferred across all metrics and data series, with the exception of the CDIR for the Capesize data series and the MAE for the Panamax data series, where the $\text{ARMA}(1;1)$ model is preferred. There therefore appears to a sector effect in that, with the exception of the Aframax data series, the tanker data series indicate that the $\text{ARMA}(1;1)$ model is preferred, while for the dry-bulk sector, the $\text{ARIMA}(1;1;1)$ model is preferred. This coincides with the argument above that there is more inter-changeability in terms of the cargoes that dry-bulk vessels carry, whereas, the cargo is standard across tankers. This would imply that imbalances in supply and demand in the dry-bulk sector would persist for longer as vessels may change routes. One should note that a possible reason for the poor performance of the ARFIMA models in this forecasting exercise may be that simpler models tend to forecast better than more complex ones. However, forecasting freight rates has been such a source of academic interest exactly because this process is exceptionally difficult. Furthermore, the poor forecasting performance of the ARFIMA models should not take away from the fact that they provide another dimension to the analysis of freight rates, and therefore further attempts using different sample periods and a different underlying series may yet provide better results.

One should further note that, in the interests of robustness, further one-step, two-step and four-step ahead forecasts were generated for the sub-periods between 3 October 2003 and 11 August 2006 and 18 August 2006 and 26 June 2009, corresponding to the first- and second-halves of the total out-of-sample horizon; as well as the sub-periods between 3 October 2003 and 16 May 2008 and 23 May 2008 and 26 June 2009, corresponding to the periods pre and post the credit crisis. The start of the credit crisis was calculated as the date at which the Baltic Dry Index reached its record maximum before beginning to dramatically fall, i.e. 20 May 2008. The results of the first two sub-periods and pre-crisis analysis were fairly consistent with those for the total forecast period presented above, across the one-step, two-step and four-step ahead forecasts, while those for post-crisis period indicated that the ARIMA specification was generally preferred across all data series, with the ARFIMA specification fairing slightly better.

5.6 Conclusion

The chapter examined the first moment of freight rates and, in particular, examined the hypothesis that spot freight rate levels are fractionally integrated. This hypothesis argues that freight rates exhibit long-memory, as the short-run dynamics of the supply and demand functions for shipping services are such that supply is unable to sufficiently adjust for changes in demand, thus resulting in freight rates exhibiting non-stationary characteristics. Nonetheless, when looking at a more long-term time frame, supply is able to adjust to changes in demand, and freight rates revert to the mean, thus illustrating behaviour more characteristic of a stationary series. This chapter presented a third alternative, i.e. that of fractionally integrated freight rates, whereby freight rates would exhibit long-memory in that shocks to freight rates would persist, in a manner contrary to stationary processes, but would eventually revert to the mean, unlike non-stationary processes.

In order to test this hypothesis fully, standard unit root tests were performed on five data series, where three of these data series were from the tanker market and two from the dry-bulk market, and where the results of these tests were inconclusive leading to the conclusion that further examination was necessary. In order to determine whether the data series were merely fractionally integrated white noise, $\text{ARFIMA}(0;d;0)$ models were estimated, and the residuals tested using the Nielsen test for stationarity. The results of these tests indicated that the residuals of the models were non-stationary, hence it was concluded that the data was not white noise and therefore had some information content within it.

Given these findings, ARMA and ARIMA models of different orders were estimated, with the best model for each class being selected on the basis of the respective log-likelihoods and AICs and SBICs. It was observed that across all five data series and two models, one should never need more than an $\text{ARMA}(1;1)$ or $\text{ARIMA}(1;1;1)$, therefore, when running the ARFIMA models, the conclusion was drawn that this type of lag dynamic should be sufficient. The results for these ARFIMA models were unanimous in determining significant *d*-parameter, which measure the level of fractional integration, where $0.5 < d < 1.0$, thus indicating that the data series exhibit long-memory. In order to test the stationarity of these residuals, and therefore double-check that these results were not spurious, Nielsen tests were performed on the respective residuals, unanimously indicating that the residuals were stationary, and, as a result, that the ARFIMA model results were not spurious.

The models were then compared using MAE, MAPE, CDIR, RMSE and Theil's U statistics, for both the in-sample and out-of-sample periods. The results for in-sample comparisons were somewhat inconclusive, however it is proposed that both the ARMA and ARFIMA models outperformed the ARMA models, where it is postulated that limitations in the size of the sample may have contributed to a lack of a conclusive results. In terms of the forecasting performance of the models, the ARMA models were found to outperform the ARIMA and ARFIMA models for the VLCC and Suezmax data, however, as these results are based on the assumption that freight rates are stationary, these results may be somewhat flawed. For the Aframax, Capesize and Panamax data series, the ARIMA models were found to outperform the ARMA and ARFIMA models.

This research thus concludes that there are some grounds for the hypothesis regarding the long-memory nature of freight rates thereby providing an alternative dimension to debate as to the true nature of the structure of the first moment of freight rates. As mentioned in the introduction to this chapter, above, this has a profound impact not only on the primary users of ships, i.e. ship-owners and charterers, but also on the wide number of auxiliary parties in the shipping markets. In addition to this, it may also provide an insight into other markets, for instance, the real estate market, in which the underlying asset in the market is also a real asset, or other such service-based industries.

In the following chapter, the issue of fractional integration in terms of the volatility of freight rate returns is examined. To this end, Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) models are compared with more standard techniques.

6 Volatility and the Dynamics of the Second Moment

6.1 Introduction

Having examined the nature of the first moment of freight rates in Chapter 5 of the thesis, this chapter expands on these concepts as well as on the traditional models of freight rate volatility through the use of Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) models. The results obtained are then compared to more standard models of volatility from the shipping literature. By doing this, a better understanding of the structure of freight rates, with particular reference to the second moment of spot freights and the degree of persistence therein, is obtained. The structure of freight rate volatility and the degree of persistence, in terms of this volatility, is considered to be one of the most crucial issues in the shipping industry. By accurately modelling this volatility, one is able to better understand their potential risk exposure and the period for which this exposure will exist.

In the shipping market, freight rates play a pivotal role, and form the basis of almost every function, from the determination of the price of the transport service through to the price of second-hand vessels. Therefore, a correct model for freight rate volatility is vital for all participants in the shipping market, from the ship-owners and charterers themselves, right down to ship-brokers, maritime lawyers and other auxiliary parties. This follows because, by reducing the risk exposure of the ship owners, one is passing that risk reduction down the line to the ancillary parties concerned.

Therefore, any model that can accurately forecast freight dynamics and volatility, and then the transition between periods of increasing and decreasing dynamics and volatility, will be of significant value.

Looking at the shipping literature, discussed in more detail in Chapter 2, traditional fundamentals models have suggested that freight rates are mean reverting, where these models are outlined by Hawdon (1978) and Beenstock and Vergottis (1989), amongst others. However, research in the 1990s, including the research by

Berg-Andreassen (1996), Glen (1997) and Kavussanos and Nomikos (2003), amongst others, found that freight rates were not mean reverting but followed a random walk process, and therefore were non-stationary in levels. In contradistinction, the most recent research, such as Adland and Cullinane (2006) and Koekebakker, *et al.* (2006), propose that the original fundamentals models of Hawdon and Beenstock and Vergottis were in fact correct and propose that freight rates are stationary, and that contrary conclusions were as a result of the application of an incorrect test. As discussed in Chapter 5, this research puts forward, for the first time in a market in which real assets are traded, the proposal that the answer may in fact lie somewhere between the two rival conclusions, i.e. that freight rates are fractionally integrated. The rationale behind this statement is that, the fact that the supply of shipping is fixed in the short-term and that the demand function is relatively inelastic combine to create a situation where freight rates do not immediately adjust to the equilibrium level, but do adjust eventually. This is in essence the definition of a long memory, or fractionally integrated process. Having established that the spot price levels of freight rates are fractionally integrated, as illustrated in Chapter 5, Bollerslev (1986)'s Generalised Autoregressive Conditional Heteroscedasticity (GARCH), Engle and Bollerslev (1986)'s Integrated Generalised Autoregressive Conditional Heteroscedasticity (IGARCH), as well as Baillie, *et al.* (1996a)'s FIGARCH models are run to determine whether the same can be said with respect to the volatility of the underlying freight rates.

Once again, if this hypothesis is true, it enables market participants to better forecast freight rate volatility, which in turn can lead to increased profits for market participants due to better investment-making capability and greater risk reduction. This can also have run-on effects on other markets, as most of the commodities traded in the world are transported by sea. A better understanding of the transport costs involved enables charterers to better forecasts their costs and potentially pass on these cost-savings to other participants in other markets. Better forecasting of freight rates enables a better understanding of investment timing, which could then be applied to other markets in which real assets are traded. In addition to this, there are of course policy and decision making implications.

In order to provide body to the hypotheses, section two of this chapter presents the methodologies applicable to the research question. Section three of the chapter presents the data as well as descriptive statistics. Section four examines the implementation of FIGARCH models in the shipping markets and section five provides a conclusion.

6.2 Methodology

Having outlined the rationale and the aims of this chapter in the previous section, this section provides a brief account of the methodology to be used in this chapter, where a more detailed discussion can be found in Chapter 3 of the thesis. These are three main methodologies used to estimate the volatility of freight rates, the first of which is the GARCH model, first proposed by Bollerslev (1986). This GARCH model may be specified as follows:

$$\left[1-\alpha(L)-\beta(L)\right]\varepsilon_t^2=\omega+\left[1-\beta(L)\right]v_t \qquad (6.1)$$

In Expression (6.1), L denotes the lag operator, hence $\alpha(L)\equiv\alpha_1 L+\alpha_2 L^2+\ldots\alpha_q L^q$ and $\beta(L)=\beta_1 L+\beta_2 L^2+\ldots\beta_p L^p$, where $m\equiv\max\{p;\ q\}$ and $v_t\equiv\varepsilon_1^2-h_t$ is mean zero serially uncorrelated; thus, the $\{v_t\}$ process may be readily interpreted as the 'innovations' for the conditional variance.

The second model used, i.e. the IGARCH model, developed by Engle and Bollerslev (1986), is an extension of the GARCH (1;1), where the IGARCH model may be represented as follows:

$$\phi(L)(1-L)\varepsilon_t^2=\omega+\left[1-\beta(L)\right]v_t \qquad (6.2)$$

$$h_t=\omega\left[1-\beta(L)\right]^{-1}+\left\{1-\left[1-\beta(L)\right]^{-1}\phi(L)(1-L)\right\}\varepsilon_t^2 \qquad (6.3)$$

In Expressions (6.2) and (6.3), above, $\phi(L)\equiv\left[1-\alpha(L)-\beta(L)\right](1-L)^{-1}$ is of order $m-1$. In addition, the autoregressive lag polynomial, $1-\alpha(L)-\beta(L)$, contains a unit root and all the roots of $\phi(L)$ and $\left[1-\beta(L)\right]$ lie outside the unit root circle.

The final extension to the ARCH model examined here is the FIGARCH model, proposed by Baillie, *et al.* (1996a), which is simply obtained by replacing the first difference operator, outlined in Expression (6.3), with the fractional difference operator, denoted *d*, where $0 < d < 1$, hence:

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + \left[1-\beta(L)\right]v_t \qquad (6.4)$$

$$h_t = \omega\left[1-\beta(L)\right]^{-1} + \left\{1-\left[1-\beta(L)\right]^{-1}\phi(L)(1-L)^d\right\}\varepsilon_t^2 \qquad (6.5)$$

In Expressions (6.4) and (6.5), above, the autoregressive lag polynomial, $1-\alpha(L)-\beta(L)$, contains a unit root and all the roots of $\phi(L)$ and $\left[1-\beta(L)\right]$ lie outside the unit root circle. The argument behind the utilisation of this model is that the squared innovations of the current conditional variance would have a slow hyperbolic rate of decay. This would mean that shocks to volatility would persist longer than in the case of the GARCH model, but would eventually decay, unlike in the IGARCH model, hence the use of the term long memory. One should note that no further models, such as those that take into account asymmetry, as this investigation is only concerned with the degree of integration and therefore incorporating any other dynamics into the equation may distort this analysis.

6.3 Description of Data

Having established the methodology to be used in this empirical analysis in the previous section, this section provides a very brief summary of the relevant descriptive statistics for the respective data series, where the complete analysis can be found in Chapter 4 of the thesis. The data set used in this chapter comprises five data series of spot freight rate returns for five different vessel types across the tanker and dry-bulk sectors, namely VLCC, Suezmax (SZMX) and Aframax (AFMX) tankers, and Capesize (CPSZ) and Panamax (PNMX) dry-bulk vessels. The sample extends from 13 January 1989 to 26 June 2009, thus comprising 1,068 observations, where all data was collected from Clarksons Shipping Intelligence Network.

Calculated descriptive statistics indicate that there is are contrasting size effects in the tanker and dry-bulk sectors, where, in the tanker sector, larger vessels are found to exhibit lower returns, which may be as a result of the reduced flexibility of these vessels with respect to smaller vessels. The reverse is found in the dry-bulk sector, however this is attributed to the increased flexibility of dry-bulk vessels, which may enable the vessels to take advantage of economies of scale. Returns, across all vessel types, are found to exhibit excess kurtosis and significant positive skewness, with the exception of the Panamax data series, where returns are negatively skewed. These findings are supported Jarque-Bera statistics, where the null hypothesis of normally distributed returns is rejected in all cases. Results from the Ljung-Box tests on the squared returns indicate the presence of significant ARCH effects, at the 1^{st}, 12^{th} and 24^{th} lags, across all vessel-types. This finding indicates the appropriateness of ARCH-type models for modelling the volatility of the respective returns. Furthermore, the existence of volatility clustering in the freight rate returns gives further indication that there may be GARCH effect, thus indicating the necessity of the GARCH-type modelling.

One should note that these models were also run on the natural logarithms of the respective data series in order to take account of the fact that freight rates can never be negative. One should note that the results from these estimations did not differ significant from those presented here[22].

Having outlined the characteristics of the data series, the next section focuses on the estimation of the respective models, giving insight into the applications and implications of these results.

6.4 Empirical Results

6.4.1 Introduction

Having outlined the methodologies to be used and the characteristics of the data in the sections above, this chapter examines the issue of persistence in and the structure of the second-moment of spot freight rates. In order to achieve this aim, $\mathrm{GARCH}(1;1)$, $\mathrm{IGARCH}(1;1;0)$ and $\mathrm{FIGARCH}(1;d;0)$ models are estimated and the results

[22] The results from these estimations are not presented here due to space constraints and are available from the author upon request.

compared, in order to determine which model, and therefore, which theory of persistence, best fits the data. One should note that the mean equations for all models of tanker freight rate returns follow an $\text{ARFIMA}(1;d;0)$ process, while those for the dry-bulk returns follow an $\text{ARFIMA}(1;d;1)$, where these specifications were obtained Chapter 5 of the thesis. This section presents these estimations and discusses the potential implications for participants both in the shipping market, and in the financial markets in general.

Before examining the results, it is worth defining what is meant by persistence in volatility. The first observation is that the degree of persistence is equal to the sum of the coefficients of the previous volatility, denoted α, and the previous forecasting variance, denoted β, and is generally denoted as $\phi \equiv \alpha + \beta$. As discussed above, the previous volatility reflecting squared "news" about the returns, also known as shocks to, and is denoted ε_t^2; while the forecasted variance reflects past information as to the evolution of volatility, and is denoted h_{t-1}, as discussed in Kang, *et al.* (2009). The second observation to note is that persistence is defined as the rate at which the lagged squared innovations in the conditional variance function decay (Baillie, *et al.* (1996a)), or in simpler terms, the rate at which shocks to volatility decay.

A further point to make in this discussion is with regard to the properties that the volatility models have in terms of their assumptions for persistence. The GARCH model assumes that the conditional variance is "stationary" in that shocks to the conditional variance function will decay exponentially, and therefore there is almost no persistence. In complete contrast, the IGARCH model assumes that the conditional variance is "non-stationary" in that shocks to the conditional variance function will not decay and therefore will persist indefinitely. As mentioned above, the FIGARCH model occupies the middle ground in that shocks to the conditional variance function will decay hyperbolically; hence one has a level of persistence in between that proposed by the GARCH and IGARCH models. It is worth pointing out that the GARCH and IGARCH models are indeed "special" cases of the FIGARCH model, in that the GARCH model is equivalent to a FIGARCH $(p;d;q)$ model, where $d=0$; while the IGARCH model is equivalent to a FIGARCH $(p;d;q)$ model, where $d=1$.

The final point that should be noted here is that, due to computational issues and to ensure consistency across models, each of the models is estimated using the two-step approach. This entails first estimating the mean equation, or ARFIMA component, for each model, where this is generated using the returns on each data series. Following this, the variance equation is estimated and the respective conditional volatility for each data series is calculated.

The following sub-sections present the results of these models and provide a critical evaluation of their meaning and implications for the structure of freight rate returns and the risk-profile of the shipping markets.

6.4.2 Results for the ARFIMA-GARCH Models

This analysis begins with the GARCH model, where results for the $\mathrm{ARFIMA}(1;d;0)$-$\mathrm{GARCH}(1;1)$ models, for the respective tanker series examined here, and $\mathrm{ARFIMA}(1;d;1)$-$\mathrm{GARCH}(1;1)$ models, for the dry-bulk series, are presented in Table 6.1.[23] These results for all the mean equations in the GARCH models across all data series suggest that the long-run average return for the respective freight rates, denoted μ, has no significant impact upon the returns for these series. Additionally, the first-order autoregressive coefficient, denoted $\mathrm{AR}(1)$ here, which measures the impact that previous period's, in this case the previous week's, freight rate returns have on the current returns for that vessel class, is found to be both significant, indicating that these do indeed have an impact on the prevailing returns, and larger for larger vessel types across both sectors. This size effect is logical in that the restrictions, both in terms of the size of the vessel, and thus where it can operate, and the commodities that these vessels can carry, mean these vessels are more likely to operate on specific routes and therefore be more dependent on the returns on that route. Finally, the first-order moving average coefficient, denoted

[23] One should note that due to computational issues the mean and variance equations are estimated separately, however, this should not significantly affect the results.

Table 6.1 – ARFIMA-GARCH Model Results

$$\rho(L)(1-L)^d(y_t-\mu)=u_t$$
$$u_t=z_t\sigma_t \;\;;\;\; z_t \sim i.i.d.N(0\;;\;1)$$
$$\phi(L)(1-L)\varepsilon_t^2=\omega+\left[1-\beta(L)\right]v_t$$

	VLCC	SZMX	AFMX	CPSZ	PNMX
μ	0.004	0.006	0.005	0.002	0.002
	(0.363)	(0.221)	(0.165)	(0.577)	(0.345)
AR(1)	0.194	0.171	0.098	0.410	0.143
	(0.000)	(0.000)	(0.001)	(0.000)	(0.004)
MA(1)	-----	-----	-----	0.013	0.159
	(-----)	(-----)	(-----)	(0.802)	(0.005)
ω	0.000	0.000	0.007	0.000	0.000
	(0.492)	(0.594)	(0.006)	(0.100)	(0.273)
α	0.053	0.020	0.267	0.112	0.120
	(-----)	(-----)	(-----)	(-----)	(-----)
β	0.947	0.980	0.258	0.867	0.840
	(0.000)	(0.000)	(0.268)	(0.000)	(0.000)
ϕ	1.000	1.000	0.526	0.980	0.960
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log L	923.215	849.513	842.936	1674.744	1658.511
AIC	-1.727	-1.589	-1.577	-3.134	-3.104
SBIC	-1.726	-1.588	-1.576	-3.133	-3.103

Note 1: VLCC denotes the weekly spot freight rates for a 270,000 DWT VLCC tanker carrying crude oil from Ras Tanura (Saudi Arabia) to Rotterdam (Netherlands).
SZMX denotes the weekly spot freight rates for a 130,000 DWT Suezmax tanker carrying crude oil from Bonny (Nigeria) to off the coast of Philadelphia (USA).
AFMX denotes the weekly spot freight rates for an 80,000 DWT Aframax tanker carrying crude oil from Sullom Voe (UK) to Bayway (USA).
CPSZ denotes the weekly spot freight rates for a 145,000 DWT Capesize bulk-carrier carrying iron ore from Tubarao (Brazil) to Rotterdam (Netherlands).
PNMX denotes the weekly spot freight rates for a 55,000 DWT Panamax bulk-carrier carrying grain from the Hampton Roads (USA) to Antwerp-Rotterdam-Amsterdam (Benelux).

Note 2: The sample period for the data used for this table extends from 13 January 1989 to 26 June 2009, with a total of 1,068 observations.

Note 3: The data used for this table is all sourced from the Clarkson Shipping Intelligence Network (www.clarskons.net).

Note 4: Figures in parentheses denote the respective *p*-values for the null hypothesis that the coefficients are not significantly different from zero.

Note 5: $\phi \equiv \alpha+\beta$

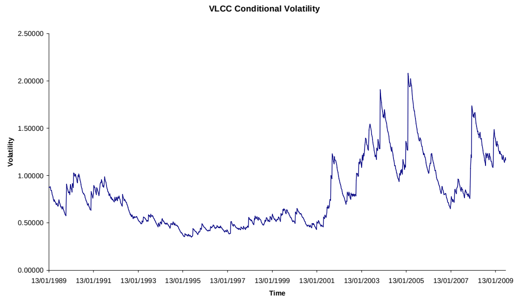
$\mathrm{MA}(1)$, which denotes the impact that past residuals have on the current returns, and is only estimated for the dry-bulk sector, is found to be insignificant for the Capesize data series, and significant for the Panamax data series. The fact that this is significant for the smaller vessel class, for which there are more owners, may be an indication of a lack of transparency in the way in which information regarding market conditions is assimilated between market participants.

Moving onto the variance equations for the models, the long-term average conditional volatility, denoted ω, is only significant for the Aframax data series. This implies that in this case, average volatility over the previous periods does have a significant impact on the volatility currently experienced in the market. In terms of the volatility during the previous period, i.e. shocks in the respective markets, as measured by the α coefficient, these appear to have some impact on the volatility of freight rate returns, although, with the possible exception of the Aframax series, this impact appears to be relatively small. Furthermore, there appears to be a size effect in this respect, as well, where smaller vessels are more exposed to shocks than larger vessels. This may be as a result of the fact that, as larger vessels operate on fewer routes, the frequency of these shocks is diminished. Looking at the impact of past variance on the volatility of the freight rates, denoted β, the impact of these is found to be very large and significant for all data series but the Aframax series, where these are found to be insignificant, hence past shocks are found to play a large role in the risk exposure of market participants.

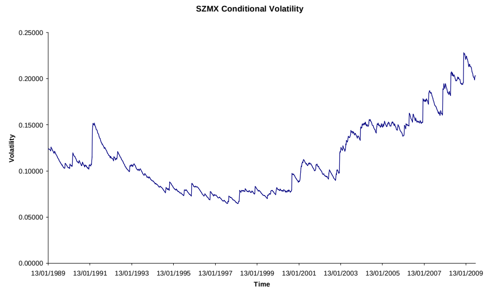
The persistence of volatility in the market, i.e. the time it takes for the impact of shocks, both past and present, to decay, which is measured by the ϕ coefficient, where $\phi = \alpha + \beta$, is found to be very large and significant, where, with the exception of the Aframax data series, this is close to one, indicating a very slow rate of decay for the shocks, not at all consistent with the GARCH specification, thus indicating that either IGARCH or FIGARCH models may be more appropriate in these cases, as these take this slow decay into account. Additionally, the persistence in volatility appears to be greater for larger vessels, which is most probably as a result that, as the trading opportunities are more limited for larger vessels, freight rates, and therefore returns, will take longer to adjust to shocks. One should note that the calculated annualised conditional volatility for each data series is graphically presented in Figure 6.1, where one can see that for the VLCC and Suezmax data series, where the coefficients measuring persistence are highest, the volatility appears to exhibit the "explosive" characteristics generally attributed to non-stationary data series, hence

Figure 6.1 – Conditional Annualised Volatility for the GARCH Models

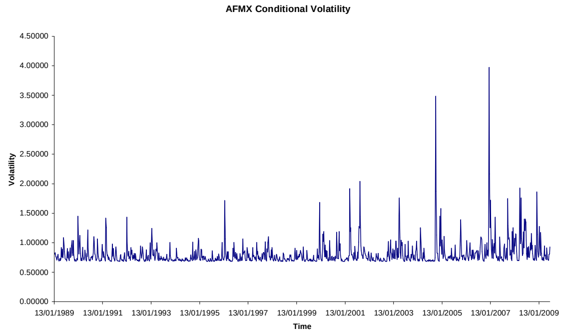
(a) VLCC Data Series



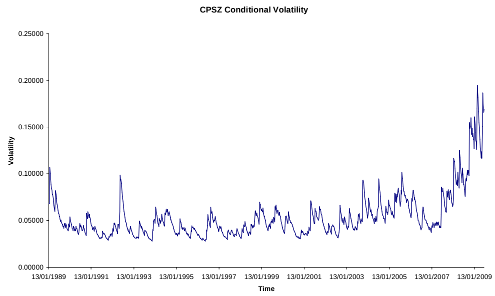
(b) SZMX Data Series



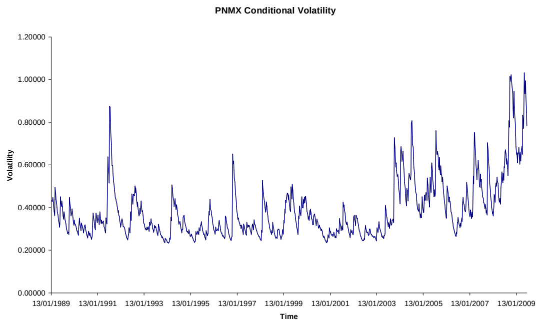
(c) AFMX Data Series



(d) CPSZ Data Series



(e) PNMX Data Series



lending further support to the hypothesis of persistence in the volatility of spot freight levels.

In order to test these results, residual diagnostics were calculated for each of the data series, where these results can be found in Table 6.2. The results of these tests indicate that there appears to be a size effect in that the degree of skewness and kurtosis of the residuals is lower for larger vessels than smaller vessels, as well as a sector effect, in that these measures tend to be larger in the tanker sector than in the dry-bulk sector, where these measures are supported by the results for the Jarque-Bera test, where the null hypothesis that the residuals are normally distributed is rejected in all cases. Looking at the results for the Ljung-Box statistics, one can observe that the

Table 6.2 – ARFIMA-GARCH Standardised Residual Diagnostics

	VLCC	SZMX	AFMX	CPSZ	PNMX
Average	0.006	0.011	0.014	0.007	0.005
Variance	0.997	0.989	0.999	1.008	1.000
Skewness	0.964	1.275	1.888	0.496	0.304
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	6.630	7.056	15.154	5.980	6.740
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	751.826	1021.232	7207.840	438.888	638.736
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Q(1)$	5.972	1.972	1.297	0.001	0.354
	(0.015)	(0.160)	(0.255)	(0.978)	(0.552)
$Q(12)$	27.774	33.900	42.932	15.699	9.140
	(0.006)	(0.001)	(0.000)	(0.205)	(0.691)
$Q(24)$	36.611	46.952	92.677	26.712	23.125
	(0.048)	(0.003)	(0.000)	(0.318)	(0.512)
$Q^2(1)$	0.150	4.522	0.299	1.280	0.015
	(0.699)	(0.033)	(0.585)	(0.258)	(0.902)
$Q^2(12)$	11.401	12.160	8.146	13.640	5.399
	(0.495)	(0.433)	(0.774)	(0.324)	(0.943)
$Q^2(24)$	17.399	18.531	11.618	21.483	10.640
	(0.831)	(0.777)	(0.984)	(0.610)	(0.991)

Note 1: See notes from Table 6.1.

null-hypothesis of no ARCH effects cannot be rejected across all data series and across all lags, thereby indicating that the $\text{ARFIMA}(1;d;0)$-$\text{GARCH}(1;1)$, in the case of the tanker sector, and $\text{ARFIMA}(1;d;1)$-$\text{GARCH}(1;1)$, for the dry-bulk sector, models are well-specified and have removed any ARCH effects present in the data.

Although these models are well specified, the fact that there is such a high degree of persistence, in terms of the volatility, is problematic as this is not characteristic of GARCH models. For this reason, alternate GARCH-type models, i.e. the IGARCH and FIGARCH models, which take this persistence into account, are evaluated in the following sub-sections, where, in the following section, the results for the IGARCH models are examined.

Table 6.3 – ARFIMA-IGARCH Model Results

$$\rho(L)(1-L)^d\left(y_t-\mu\right)=u_t$$
$$u_t=z_t\sigma_t \;\; ; \;\; z_t \sim i.i.d.N(0\;;\;1)$$
$$\phi(L)(1-L)\varepsilon_t^2=\omega+\left[1-\beta(L)\right]v_t$$

	VLCC	SZMX	AFMX	CPSZ	PNMX
μ	0.004	0.006	0.005	0.002	0.002
	(0.363)	(0.221)	(0.165)	(0.577)	(0.345)
AR(1)	0.194	0.171	0.098	0.410	0.143
	(0.000)	(0.000)	(0.001)	(0.000)	(0.004)
MA(1)	-----	-----	-----	0.013	0.159
	(-----)	(-----)	(-----)	(0.802)	(0.005)
ω	0.000	0.000	0.006	0.000	0.000
	(0.390)	(0.197)	(0.002)	(0.178)	(0.459)
β	0.947	0.979	0.167	0.870	0.857
	(0.000)	(0.000)	(0.261)	(0.000)	(0.000)
ϕ	1.000	1.000	1.000	1.000	1.000
	(-----)	(-----)	(-----)	(-----)	(-----)
Log L	923.19	849.45	825.18	1672.746	1653.961
AIC	-1.727	-1.589	-1.543	-3.131	-3.095
SBIC	-1.726	-1.588	-1.542	-3.130	-3.094

Note 1: See notes from Table 6.1.

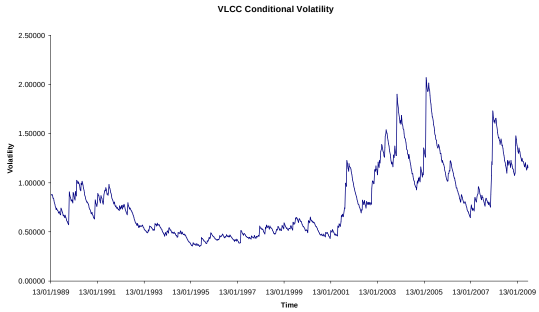
6.4.3 Results for the ARFIMA-IGARCH Models

Having examined the results for the GARCH models in the previous sub-section, the results IGARCH models are now examined, where these results are presented in Table 6.3. It is important to note that the specifications for the models differ, where an $\text{ARFIMA}(1;d;0)$-$\text{IGARCH}(1;1;0)$ specification was used for the tanker series, and an $\text{ARFIMA}(1;d;1)$-$\text{IGARCH}(1;1;0)$ specification, for the dry-bulk series. One should that, for reasons of brevity, the results for the respective mean equations of each data series are not presented here as they are exactly the same as those for the GARCH models discussed above.

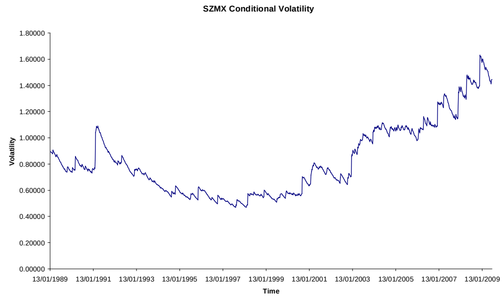
Looking at the variance equation, the long-run average volatility, once again denoted ω, plays no significant role, with the exception of the Aframax data series, in the prevailing volatility of the spot freight returns, a result consistent with that observed

Figure 6.2 – Conditional Annualised Volatility for the IGARCH Models

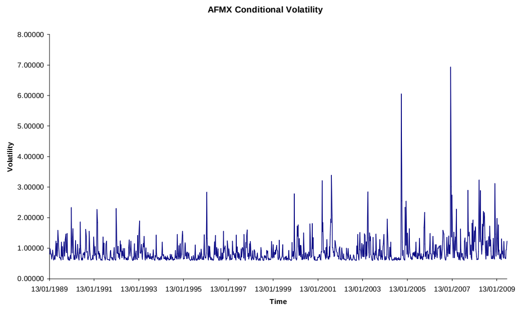
(a) VLCC Data Series



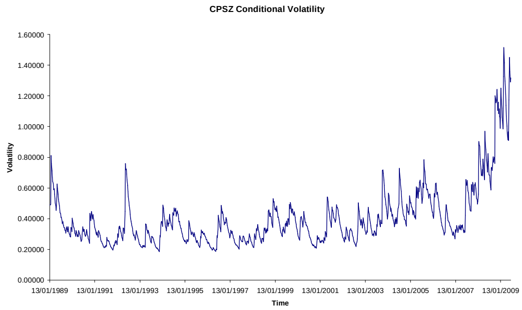
(b) SZMX Data Series



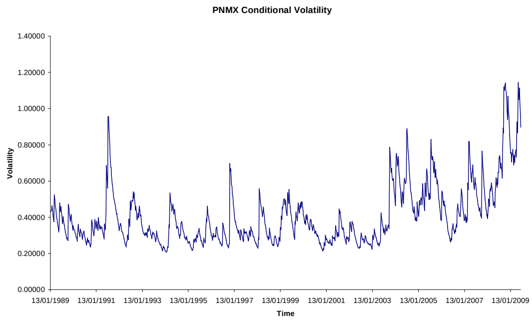
(c) AFMX Data Series



(d) CPSZ Data Series



(e) PNMX Data Series



for the GARCH model. As far as the impact of past variance is concerned, as measured by the β coefficient, the results are very similar to those for the GARCH model in that past shocks are found to play a significant and large role in the current volatility experienced in the market, with the exception of the Aframax data series, where the coefficient is found to be insignificant, however, this may just be an anomaly of the data series used. One should note that the ϕ parameter, which measures the persistence of volatility, is constrained to equal one, hence $\alpha + \beta = 1$, for the IGARCH model, in that this model assumes that volatility persists indefinitely. Once again, the estimated annualised conditional volatility for each data series is graphically presented in Figure 6.2, where one should note that these are remarkably similar to the generated conditional volatility from the GARCH model, exhibited in

Table 6.4 – ARFIMA-IGARCH Standardised Residual Diagnostics

	VLCC	SZMX	AFMX	CPSZ	PNMX
Average	0.006	0.011	0.021	0.005	0.003
Variance	1.003	0.997	0.917	0.961	0.937
Skewness	0.968	1.274	1.660	0.478	0.343
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	6.649	7.040	12.549	6.053	7.288
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	759.263	1015.328	4548.450	455.291	839.132
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Q (1)	5.917	1.930	1.758	0.000	0.490
	(0.015)	(0.165)	(0.185)	(0.998)	(0.484)
Q (12)	27.759	33.792	36.301	15.383	9.706
	(0.006)	(0.001)	(0.000)	(0.221)	(0.642)
Q (24)	36.608	46.869	78.049	25.619	23.451
	(0.048)	(0.003)	(0.000)	(0.373)	(0.493)
Q^2 (1)	0.139	4.546	1.655	0.796	0.080
	(0.710)	(0.033)	(0.198)	(0.372)	(0.778)
Q^2 (12)	11.590	12.129	13.804	14.731	7.838
	(0.479)	(0.435)	(0.313)	(0.256)	(0.798)
Q^2 (24)	17.583	18.500	17.482	20.094	13.987
	(0.823)	(0.778)	(0.827)	(0.691)	(0.947)

Note 1: See notes from Table 6.1.

Figure 6.1. A possible reason for this is that coefficients measuring persistence for the GARCH model are so close to one, that one is effectively merely running an IGARCH model, thus lending support to the persistence hypothesis.

As was the case for the GARCH models, these results are tested by calculating the respective residual diagnostics for each data series, the results of which are presented in Table 6.4. The results of these tests indicate that, as was the case for the GARCH estimations, there appears to significant skewness and excess kurtosis in the residuals, in addition to a size effect in that the degree of skewness and kurtosis of the residuals is lower for larger vessels than smaller vessels, as well as a sector effect, in that these measures tend to be larger in the tanker sector than in the dry-bulk sector, where these measures are supported by the results for the Jarque-Bera test, where the null-hypothesis that the residuals are normally distributed is rejected in all cases. Looking at the results for the Ljung-Box statistics, which were calculated on the standardised residuals, one can once again observe that the null hypothesis of no ARCH effects cannot be rejected across all data series and across all lags, thereby indicating that the $\text{ARFIMA}(1;d;0)$-$\text{IGARCH}(1;1;0)$, in the case of the tanker sector, and $\text{ARFIMA}(1;d;1)$-$\text{IGARCH}(1;1;0)$, for the dry-bulk sector, models are well-specified and have removed any ARCH effects present in the data.

Although the characteristics of persistence with regard to the volatility are consistent with the IGARCH model, the assumption that this persistence is indefinite does not tie in with traditional maritime economic theory. This argues the characteristics of supply and demand imply that there will be periods of fairly stable freight rates, followed by periods in which freights rate may change dramatically, thereby implying mean reversion in the volatility of freight rate returns. In order to reconcile these two issues, the following sub-section examines the results for the FIGARCH model, which argues that, while there may be greater persistence in volatility than that implied by the GARCH model, shocks will eventually decay at a slow rate, unlike the infinite persistence postulated by the IGARCH model.

6.4.4 Results for the ARFIMA-FIGARCH Models

Having examined the results for the GARCH and IGARCH models in the previous sub-section, the results for the FIGARCH model are now examined, where these are presented in Table 6.5. The models differ between sectors of the freight markets, in that an $\text{ARFIMA}(1;d;0)$-$\text{FIGARCH}(1;d;0)$ model is estimated for the tanker sector and an $\text{ARFIMA}(1;d;1)$-$\text{FIGARCH}(1;d;0)$ for the dry-bulk sector. As was the case for the IGARCH model, the results of the respective mean equations are not discussed here as they are identical to those for the GARCH model.

Examining the results for the variance equation, the long-run average conditional volatility, as measured by the ω parameter, is found to be insignificant for all but the Aframax data series, thus suggesting that this does not play a part in the prevailing volatility of freight rate returns. One should note that, as a result of the specifications for the model, no α or ϕ parameters are estimated, however, their implications,

Table 6.5 – ARFIMA-FIGARCH Model Results

$$\rho(L)(1-L)^d\left(y_t-\mu\right)=u_t$$
$$u_t=z_t\sigma_t \;\; ; \;\; z_t\sim i.i.d.N(0\;;\;1)$$
$$\phi(L)(1-L)^d\,\varepsilon_t^2=\omega+\left[1-\beta(L)\right]v_t$$

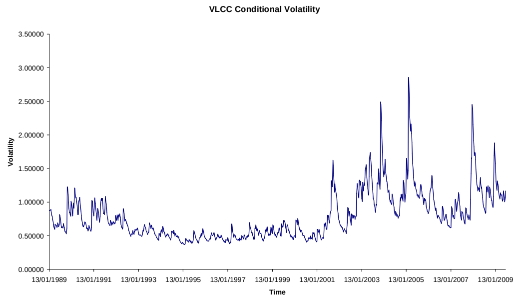
	VLCC	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
μ	0.004	0.006	0.005	0.002	0.002
	(0.363)	(0.221)	(0.165)	(0.577)	(0.345)
AR(1)	0.194	0.171	0.098	0.410	0.143
	(0.000)	(0.000)	(0.001)	(0.000)	(0.004)
MA(1)	-----	-----	-----	0.013	0.159
	(-----)	(-----)	(-----)	(0.802)	(0.005)
ω	0.000	0.001	0.001	0.000	0.000
	(0.208)	(0.355)	(0.000)	(0.303)	(0.156)
β	0.504	0.038	-----	0.203	0.248
	(0.000)	(0.571)	(-----)	(0.015)	(0.001)
d	0.515	0.286	0.136	0.402	0.390
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log L	922.60	844.39	845.99	1677.873	1662.097
AIC	-1.726	-1.579	-1.582	-3.140	-3.111
SBIC	-1.725	-1.578	-1.581	-3.139	-3.110

Note 1: See notes from Table 6.1.

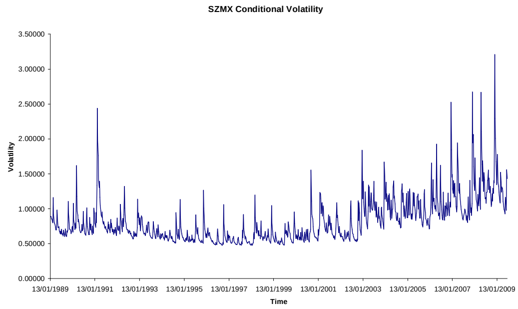
should the specification change, would be as above. For this reason, the only measure of persistence presented here is the impact of past variance on the conditional variance, denoted β, the results for which suggest that past variance, or shocks, has a significant impact on conditional volatility in the case of the VLCC, Capesize and Panamax data series, although the parameter for the Capesize data series is only significant at the 5% and 10% levels of significance. In contrast, this parameter was insignificant for the Suezmax and Aframax data series, where the estimate for the Aframax data series is not presented due to constraint issues. It is interesting to note that the level of persistence, for those series where this is significant, is much lower than for the GARCH and IGARCH models, implying that the FIGARCH model has taken account of and modelled this persistence effectively. One should note that the *d*-parameter measures the order of integration for the conditional volatility series, where $0<d<1$. The results for this parameter indicate that the conditional variance is indeed fractionally integrated, where these values are significant and range between 0.136, in the case of the Aframax data series, and 0.515, for the VLCC series. A size

Figure 6.3 – Conditional Annualised Volatility for the FIGARCH Models

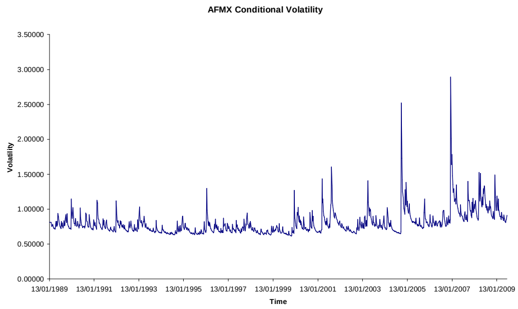
(a) VLCC Data Series



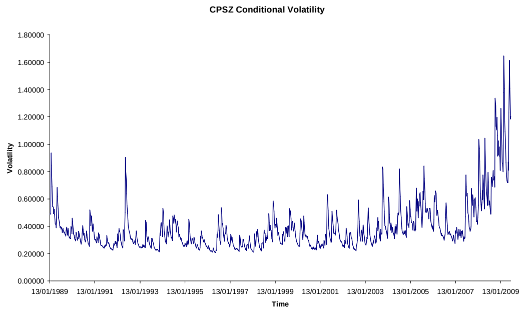
(b) SZMX Data Series



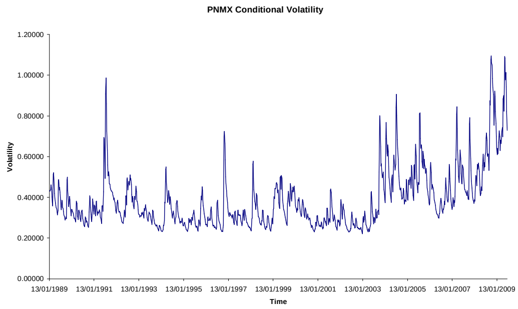
(c) AFMX Data Series



(d) CPSZ Data Series



(e) PNMX Data Series



effect is also observed here in that larger vessels appear to exhibit more long memory characteristics. As was the case for the discussion on the GARCH and IGARCH models, the evolution of the annualised conditional volatility for each data series is graphically presented in Figure 6.3. Interestingly, unlike the generated conditional volatility for the GARCH and IGARCH models, illustrated in Figure 6.1 and Figure 6.2 respectively, the conditional volatility for the FIGARCH model appears to more close resemble a stationary process, hence lending support to the hypothesis that freight rate volatility follows a fractionally integrated process.

Once again the results are tested by calculating the residual diagnostics, summarised in Table 6.6. The results of these tests indicate that, as was the case for the GARCH and IGARCH models, there appears to be a size effect in that the degree of skewness

Table 6.6 – ARFIMA-FIGARCH Standardised Residual Diagnostics

	VLCC	SZMX	AFMX	CPSZ	PNMX
Average	0.005	0.016	0.017	0.009	0.513
Variance	1.021	1.035	0.998	1.019	0.474
Skewness	0.948	1.262	2.209	0.512	0.307
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	6.537	6.605	18.367	6.158	6.667
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	715.359	860.063	11355.765	489.414	614.048
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Q(1)$	6.047	4.156	0.708	0.055	0.513
	(0.014)	(0.041)	(0.400)	(0.814)	(0.474)
$Q(12)$	25.132	35.945	36.767	15.771	9.275
	(0.014)	(0.000)	(0.000)	(0.202)	(0.679)
$Q(24)$	35.379	51.803	84.062	27.130	23.041
	(0.063)	(0.001)	(0.000)	(0.298)	(0.517)
$Q^2(1)$	0.665	0.495	0.035	0.000	0.065
	(0.415)	(0.482)	(0.851)	(0.998)	(0.799)
$Q^2(12)$	9.994	12.317	2.985	12.111	4.694
	(0.617)	(0.421)	(0.996)	(0.437)	(0.967)
$Q^2(24)$	14.771	20.826	4.771	23.576	11.060
	(0.927)	(0.649)	(1.000)	(0.486)	(0.989)

Note 1: See notes from Table 6.1.

and kurtosis of the residuals is lower for larger vessels than smaller vessels, as well as a sector effect, in that these measures tend to be larger in the tanker sector than in the dry-bulk sector. These findings are supported by the results for the Jarque-Bera test, where the null hypothesis that the residuals are normally distributed is rejected in all cases. Looking at the results for the Ljung-Box statistics, one can observe that the null hypothesis of no ARCH effects cannot be rejected across all data series and across all lags, thereby indicating that the $\text{ARFIMA}(1;d;0)$-$\text{FIGARCH}(1;d;0)$, in the case of the tanker sector, and $\text{ARFIMA}(1;d;1)$-$\text{FIGARCH}(1;d;0)$, for the dry-bulk sector, models are well-specified and have removed any ARCH effects present in the data.

In contrast to the findings for the GARCH and IGARCH models, the FIGARCH model appears to have reduced the persistence of the conditional volatility. This has huge implications in that instead of finding that the effects of shocks are felt indefinitely, thereby creating large risk exposure for market participants, as implied by the findings of the GARCH and IGARCH models, the FIGARCH model finds that the effect of these shocks do eventually decay, although it is a slow process. Therefore participants in the shipping markets can adjust the risk expectations and hedging strategies accordingly. Furthermore, this model appears to have reconciled the two issues discussed above, i.e. the fact that traditional maritime economic theory suggests that the volatility process should be mean reverting, whereas the results for the GARCH and IGARCH models suggested that shocks to the volatility process persisted indefinitely across time.

6.5 Comparison of Volatility Models

The previous section discussed the results, and their implications, from the estimation of GARCH, IGARCH and FIGARCH models, however, this focused on the results of each model individual, making no comparison between the models. This section provides a critical comparison between models with the aim of determining which best describes the structure of freight rate volatility.

In terms of the mean equations for each model, these are identical in that the underlying models and data series are identical. In this vain, the results for the residual diagnostics in each model are consistent, in that all models remove any ARCH effects in the data series, thus implying that they are well-specified, while the residuals are found to exhibit both significant positive skewness and excess kurtosis, a finding supported by the results of the Jarque-Bera statistics.

This being said, the analysis becomes more interesting when looking at the parameter estimates for the respective variance equations, as this is where the differences begin to manifest themselves. Although the results for the long-run average conditional volatility, denoted ω, are fairly consistent, in that this parameter is found to be insignificant for all data series but the Aframax series, this rapidly changes when looking a parameters for persistence, i.e. the β and ϕ parameters. In the cases of the GARCH and IGARCH models, these parameters, and in particular the β parameter, which measures the impact of past variance, are found to be very large, and very close to one, thereby suggesting close to infinite persistence in volatility. In contrast, the results for the β parameter in the FIGARCH model suggest that, while shocks to volatility decay slowly, they do actually decay. The question is thus raised as to which of these arguments is correct?

Before addressing this question, one must re-visit what is meant by persistence. Persistence measures the time taken for the impact of shocks to the conditional volatility to decay. In this respect, the GARCH model has the least persistence in that it assumes that the impact of these shocks decays exponentially, hence the volatility series is effectively stationary. At the other end of the scale, the IGARCH model assumes that the impact of shocks to volatility persist indefinitely; hence the volatility series follows a non-stationary process. The middle-ground between these two extremes is held by the FIGARCH model, where the impact of shocks decays in a hyperbolic fashion, thus one can argue that the conditional volatility of freight rate returns is fractionally integrated.

To provide a more substantive evaluation of the best model, Akaike Information Criteria (AIC), attributed to Akaike (1974), and the Schwartz-Bayesian Information Criteria (SBIC), outlined in Schwarz (1978), are calculated for each of the respective models and data series, where these are presented for convenience in Table 6.7. These AIC and SBIC criteria are unanimous in that both are indifferent between the GARCH and IGARCH models for the VLCC and Suezmax data series, and select the FIGARCH model in the case of the Aframax, Capesize and Panamax data series. A possible reason for the indifference between models in the case of the VLCC and Suezmax data series may be that $\phi = 1$ in the GARCH model for these series, therefore one is effectively running an IGARCH model. On this basis, the conclusion is drawn that the IGARCH model provides the best estimate of conditional volatility for these series.

One should note that although information criteria do provide a convenient means of choosing between models, Brooks and Burke (2003) argue that these standard metrics suffer from a lack of ability in that they do not allow for the number of parameters in the models to change, thereby leading to reduced forecasting accuracy. In order to

Table 6.7 – Results for the Model Selection Criteria

Panel A - Akaike Information Criterion	GARCH	IGARCH	FIGARCH
VLCC Data Series	-1.727	-1.727	-1.726
Suezmax Data Series	-1.589	-1.589	-1.579
Aframax Data Series	-1.577	-1.543	-1.582
Capesize Data Series	-3.134	-3.131	-3.140
Panamax Data Series	-3.104	-3.095	-3.111

Panel B - Schwartz-Bayesian Information Criterion	GARCH	IGARCH	FIGARCH
VLCC Data Series	-1.726	-1.726	-1.725
Suezmax Data Series	-1.588	-1.588	-1.578
Aframax Data Series	-1.576	-1.542	-1.581
Capesize Data Series	-3.133	-3.130	-3.139
Panamax Data Series	-3.103	-3.094	-3.110

Note 1: See notes from Table 6.1.

address this issue, the models are also compared in Chapter 8 on the basis of their ability to accurately determine and minimise the respective Values-at-Risk.

A possible rationale for these conclusions relates to the characteristics of the sectors in question. Vessels in the tanker market tend to carry a single commodity, i.e. crude oil, as opposed to those in the dry-bulk market, where vessels are capable of carrying multiple commodities, and generally operate on fewer routes than those in the dry-bulk market. They are therefore heavily exposed to shocks on those routes, and are unable to trade elsewhere should a shock occur, unlike dry-bulk vessels, which can simply swap commodities; hence the effects of these shocks would logically last longer in the tanker market than in the dry-bulk market.

One can therefore make the general conclusion that IGARCH models are more suited to providing insight into the structure of volatility in the tanker sector, while FIGARCH models provide a better fit in this respect for the dry-bulk sector.

6.6 Conclusion

The structure of freight rate volatility, and the degree of persistence in terms of this volatility, is considered to be one of the most crucial issues in the shipping industry, as by being able to accurately model this volatility, one is able to better understand their potential risk exposure and the period for which this exposure will exist. This chapter examined the structure of freight rate volatility, with a particular focus on the level of persistence within the volatility framework, using the GARCH, IGARCH and FIGARCH frameworks. This chapter hypothesised that freight rate volatility would follow a hyperbolic rate of decay in that, although shocks to volatility should persist in the market, they should not persist indefinitely. The rationale behind this hypothesis is similar to those for the freight rate levels discussed in the previous chapter, i.e. that the characteristics of the supply function in the freight rate market imply that although the level of supply is fixed in the short-term, the level of supply can increase in the longer-term, thus causing freight rates to revert to their mean levels. The same can be said for freight rate volatility in the shocks to freight rate returns, such as a sudden increase in freight rates due to a lack of supply, would persist up to the point that freight rates reverted to the mean level, and then begin to slowly discuss as the market entered a more stable period.

One should also note that a correct model for freight rate volatility is vital for all participants in the shipping market, and not only ship-owners and charterers in that freight rates form the basis for all activities in the market, right down to ship-brokers, maritime lawyers and other auxiliary parties. This follows because, by reducing the risk exposure of the ship owners, one is passing that risk reduction down the line to the ancillary parties concerned. In addition, it is worth considering this methodology in other markets as well as shipping freight rate markets (perhaps the only financial market in which the good being provided is a service), as the modelling of freight rate volatility can be readily applied to other markets in which real assets are traded.

In order to evaluate these issues, freight rates for five vessel classes, i.e. VLCC, Suezmax and Aframax tankers, and Capesize and Panamax dry-bulk vessels, were collected for the period between the 13 January 1989 and 26 June 2009. The reasons for specifying the use of the GARCH, IGARCH and FIGARCH models is that each assumes a different rate of decay for shocks to volatility where the GARCH models assume an exponential rate of decay, the IGARCH models an indefinite rate of decay, and the FIGARCH model a hyperbolic rate of decay. Therefore, by determining the best model for the data series, one can draw conclusions as to the persistence of shocks in volatility, and therefore the potential risk exposure of involved parties.

Having run the models, it was found that past variance played a significant role in determining the level of volatility in the shipping freight markets, and that lagged returns are found to have an impact on the current returns in the market. When examining the results for GARCH and IGARCH models, the tentative conclusion was reached that shocks to volatility persisted indefinitely, regardless of vessel type. In contrast, however, the results from the FIGARCH models suggested that shocks with respect to freight rate volatility followed the hyperbolic rate of decay hypothesised.

In order to address the question as to which models were correct, AIC and SBIC measures were calculated and the models compared on this basis. The results for these measures led to the conclusion that, for the dry-bulk sector, the FIGARCH model provides the best fit in terms of the structure volatility, while, for the tanker sector, the IGARCH model is preferred.

Following this, Chapter 7 examines the question of whether models which incorporate conditional skewness and kurtosis, as well as the conditional volatility measured here, may provide a better understanding of the true nature of the risk exposure of market participants in the shipping freight rate markets. An example of a practical application of the findings from both these chapters is given in Chapter 8 of the thesis, where these models are used to forecast volatility. This is done by constructing a series of forecasts of volatility, the accuracy of which are determined by comparing Values-at-Risk calculated using these forecasts with the actual Values-at-Risk incurred in the shipping freight rate markets.

7 The Impact of Higher Moments on Freight Rates

7.1 Introduction

The departure from normality of asset return distributions has been well documented, and has been reported on by Harvey and Siddique (1999), Peiró (1999), Brooks, *et al.* (2005) and Bali, *et al.* (2008), amongst others. To the best of the authors' knowledge, this paper examines, for the first time in the shipping literature, the issues of time-varying skewness and kurtosis in shipping freight rate returns. From this, the aim was to establish the most comprehensive model of determining the dynamics of the returns distribution. These are crucial issues since should the returns distribution be negatively skewed, this would accentuate the left-hand side of the distribution, entailing a higher probability of decreases than increases in freight rate returns. In addition, any excess kurtosis implies that more extreme observations, i.e. extremely high, or low, returns, are more likely to occur than would be the case under the normal distribution, a crucial issue should one wish to minimise their risk exposure.

In the shipping market, freight rates play a pivotal role, and form the basis of almost every function - from the determination of the price of the transport service through to the price of second-hand vessels. Therefore, a correct model for freight rate dynamics is vital for all participants in the shipping market, from the ship-owners and charterers themselves, right down to ship-brokers, maritime lawyers and other auxiliary parties. This follows because, by reducing the risk exposure of the ship owners, one is passing that risk reduction down the line to the ancillary parties concerned.

A number of different methodologies have been proposed to deal with the issue of return dynamics beginning with the Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models introduced by Bollerslev (1986), where, although these allow for time-varying volatility, they assume constant skewness and kurtosis. Harvey and Siddique (1999) extend this model by proposing the Generalised Autoregressive Conditional Heteroscedasticity with Skewness (GARCHS) model in which a time-varying skewness component is introduced which allows for the joint estimation of time-varying conditional variance and skewness, however, this model still assumes constant kurtosis. In a similar vein, Brooks, *et al.* (2005) extend the basic GARCH model in the Generalised Autoregressive Conditional Heteroscedasticity and Kurtosis (GARCHK) model by introducing a time-varying kurtosis component. However, this now assumes constant skewness. The latest extension with respect to these issues was proposed by León, *et al.* (2005), who introduced the GARCH with skewness and kurtosis (GARCHSK) model, which allows for the joint estimation of conditional variance, skewness and kurtosis, thus enabling one to fully explore the dynamics of the data series, and thereby enabling one to obtain a complete picture of the returns distribution.

There are multiple advantages to being able to capture the conditional skewness and kurtosis of the data series, the first of which is that should the distribution of asset returns be skewed, or if there is excess kurtosis, the traditional assumption of normality when estimating the Value-at-Risk will result in an underestimation of the risk. Secondly, it enables one to better describe the distributional properties of financial asset returns, thus enabling one to better understand the performance of assets with these properties. Finally, one could look at the issue of portfolio construction to determine if the risk structure is truly optimal, as well as the fact that examining these properties would enable one to better price options in financial markets where these properties exist.

In order to provide background to the hypotheses, section two of this chapter presents a critical review of the relevant literature, while section three examines the methodologies applicable to the research question. Section four of the document presents the data as well as descriptive statistics. Section five examines the implementation of these models in the shipping markets, while section six makes a comparison with other more traditional models and section seven concludes.

7.2 Methodology

Engle (1982) provided the framework for the variety of ARCH-type models presented here through the development of the original ARCH model. This framework was extended by Bollerslev (1986), who presented the standard GARCH (p; q) model used in this thesis.

This GARCH (*p*; *q*) model is represented as follows:

$$\varepsilon_t = z_t\sqrt{h_t} \;\; ; \;\; z_t \sim N(0;\ 1) \qquad (7.1)$$

$$h_t = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)h_t \qquad (7.2)$$

In this model, L denotes the lag operator, hence $\alpha(L) \equiv \alpha_1 L + \alpha_2 L^2 + \ldots \alpha_q L^q$ and $\beta(L) = \beta_1 L + \beta_2 L^2 + \ldots \beta_p L^p$. In order to ensure the stability and covariance of the $\{\varepsilon_t\}$ process, all the roots of $\left[1-\alpha(L)-\beta(L)\right]$ and $\left[1-\beta(L)\right]$ are constrained to lie outside the unit circle. One should note that in a GARCH (1; 1) model, the sum of α_1 and β_1 reflects the persistence of any shocks to volatility (Baillie, *et al.* (1996a)). This model enables one to generate volatility forecasts which are comprised of the weighted average of the constant long-run, or average, variance, denoted ω, the previous forecasting variance, denoted h_t, and the previous volatility reflecting the squared news about the return, denoted ε_t^2 (Kang, *et al.* (2009)). This model could alternatively be expressed as an ARMA (*m; p*) process in ε_t^2, where:

$$\left[1-\alpha(L)-\beta(L)\right]\varepsilon_t^2 = \omega + \left[1-\beta(L)\right]v_t \qquad (7.3)$$

In the expression above $m \equiv \max\{p;\ q\}$ and $v_t \equiv \varepsilon_1^2 - h_t$ is mean zero serially uncorrelated; thus, the $\{v_t\}$ process may be readily interpreted as the 'innovations' for the conditional variance.

The final extension to the ARCH model examined here is the FIGARCH model proposed by Baillie, *et al.* (1996a). This is simply obtained by including the fractional difference operator, denoted *d*, where $0 < d < 1$, hence:

$$\phi(L)(1-L)^d\,\varepsilon_t^2 = \omega + \left[1-\beta(L)\right]v_t \qquad (7.4)$$

$$h_t = \omega\left[1-\beta(L)\right]^{-1} + \left\{1-\left[1-\beta(L)\right]^{-1}\phi(L)(1-L)^d\right\}\varepsilon_t^2 \qquad (7.5)$$

Once again, the autoregressive lag polynomial, $1-\alpha(L)-\beta(L)$, contains a unit root and all the roots of $\phi(L)$ and $\left[1-\beta(L)\right]$ lie outside the unit root circle. The argument behind the utilisation of this model is that the squared innovations of the current conditional variance would have a slow hyperbolic rate of decay. This would mean that shocks to volatility would persist longer than in the case of the GARCH model, but would eventually decay, unlike in the IGARCH model, hence the term long memory in this case.

The focus of the methodology used here changes from merely modelling conditional volatility to modelling the higher moments as well. Harvey and Siddique (1999) were the first to introduce the concept of autoregressive conditional skewness through the introduction of the GARCHS model. Following this, Brooks, *et al.* (2005) argue that, following the research by Mandelbrot (1963), it is almost universally accepted that asset returns are leptokurtic rather than normally distributed. For this reason, they introduced the GARCHK model, which allows for the kurtosis to develop over time in a manner that is not fixed with respect to the variance, in order to examine the impact that kurtosis has on asset returns. Subsequently, León, *et al.* (2005) extend the work done by Harvey and Siddique and Brooks, *et al.* by jointly estimating the time-variance skewness and kurtosis using their GARCHSK model.

This extends the literature in that it accounts for both time-varying skewness and kurtosis, whereas Harvey and Siddique only account for time-varying skewness and Brooks *et al.* only account for time-varying kurtosis. In addition, the likelihood function, is based on a Gram-Charlier (GC) series expansion of the normal density function, in a manner similar to that suggested by Gallant and Tauchen (1989). The reason for this change is that this easier to estimate the likelihood function based on the non-central *t*-distribution used by Harvey and Siddique. León *et al.*'s GARCHSK model is given by:

$$(1-L)^d \phi(L) r_t = \theta(L)\varepsilon_t \qquad (7.6)$$

$$\varepsilon_t = h^{1/2}\eta_t \ , \ \eta_t \sim (0,1) \ , \ \varepsilon_t \big| I_{t-1} \sim (0\ ,h_{t-1}) \qquad (7.7)$$

$$h_t = \beta_0 + \beta_1\varepsilon_{t-1}^2 + \beta_2 h_{t-1} \qquad (7.8)$$

$$s_t = \gamma_0 + \gamma_1\eta_{t-1}^3 + \gamma_2 s_{t-1} \qquad (7.9)$$

$$k_t = \delta_0 + \delta_1\eta_{t-1}^4 + \delta_2 k_{t-1} \qquad (7.10)$$

In Expressions (7.6) to (7.10), above, $E_{t-1}(\bullet)$ denotes the conditional expectation on an information set till period t – 1 denoted as I_{t-1}. They establish that $E_{t-1}(\eta_t) = 0$, $E_{t-1}(\eta_t)^2 = 1$, $E_{t-1}(\eta_t^3) = s_t$, and $E_{t-1}(\eta_t^4) = k_t$, where both s_t and k_t are driven by a

GARCH (1,1) structure. This means that s_t and k_t represent, respectively, the skewness and kurtosis corresponding to the conditional distribution of the standardised residual $\eta_t = \varepsilon_t h_t^{-1/2}$.

León *et al*. go on to obtain the density function for the standardised residuals, denoted η_t, which is conditional on the information available at time *t* – 1 by using a Gram–Charlier series expansion of the normal density function and truncating at the fourth moment. This density function is as follows:

$$g\left(\eta_t | I_{t-1}\right) = \phi\left(\eta_t\right)\left[1 + \frac{s_t}{3!}\left(\eta_t^3 - 3\eta_t\right) + \frac{k_t - 3}{4!}\left(\eta_t^4 - 6\eta_t^2 + 3\right)\right] = \phi\left(\eta_t\right)\psi\left(\eta_t\right) \qquad (7.11)$$

In Expression (7.11), $\phi(\bullet)$ denotes the probability density function, or pdf, corresponding to the standard normal distribution, while $\psi(\bullet)$ is the polynomial part of fourth order corresponding to the expression between brackets. They go on to state that this is not really a density function in that, for some of the parameter values in Expressions (7.6) to (7.10), the density $g(\bullet)$ might be negative, and, similarly, the integral of $g(\bullet)$ on $\Re$ is not equal to one.

To solve this issue, León *et al*. propose a true pdf, denoted $f(\bullet)$, which is obtained by transforming the density function $g(\bullet)$ using the method proposed in Gallant and Tauchen (1989). Looking at the specifics, to obtain this well defined density everywhere, they squared the polynomial part, i.e. $\psi(\bullet)$, and then divided by the integral of $g(\bullet)$ over $\Re$, where the latter is to ensure that the density function integrates to one. Therefore, the resulting pdf is:

$$f\left(\eta_t | I_{t-1}\right) = \frac{\phi\left(\eta_t\right)\left[1 + \frac{s_t}{3!}\left(\eta_t^3 - 3\eta_t\right) + \frac{k_t - 3}{4!}\left(\eta_t^4 - 6\eta_t^2 + 3\right)\right]^2}{\Gamma_t} = \frac{\phi\left(\eta_t\right)\psi^2\left(\eta_t\right)}{\Gamma_t} \qquad (7.12)$$

In this pdf, $f(\bullet)$, the term Γ_t is defined as follows:

$$\Gamma_t = 1 + \frac{s_t^2}{3!} + \frac{\left(k_t - 3\right)^2}{4!} \qquad (7.13)$$

Following this León *et al.* define, after omitting unessential constants, the log-likelihood function for one observation corresponding to the conditional distribution $\varepsilon_t = h_t^{1/2}\eta_t$, whose pdf is $h_t^{1/2} f\left(\eta_t \middle| I_{t-1}\right)$, as:

$$l_t = -\frac{1}{2}\ln h_t - \frac{1}{2}\eta_t^2 + \ln\left(\psi^2\left(\eta_t\right)\right) - \ln\left(\Gamma_t\right) \qquad (7.14)$$

As one can see, this likelihood function is clearly easier to estimate than that based on a non-central *t* distribution, i.e. the density function proposed by Harvey and Siddique, In addition to this, Expression (7.12) nests the normal density function, which would occur when $s_t = 0$ and $k_t = 3$, while the non-central *t* does not. This implies that the restrictions imposed by the normal density function with respect to the more general density based on a Gram–Charlier series expansion are able to be easily tested, should this be required.

Having outlined the rationale behind the study and the methodology that will be used to test this hypothesis, the following section outlines the characteristics of the data that will be used for this analysis.

7.3 Description of Data

Having outlined the methodology to be used above, this section briefly presents the data used to analyse the higher moments of freight rates in the shipping markets. One should note that a more detailed description of this data may be found in Chapter 4. The data set used in this chapter comprises five data series of spot freight rate returns for five different vessel types across the tanker and dry-bulk sectors, namely VLCC, Suezmax (SZMX) and Aframax (AFMX) tankers, and Capesize (CPSZ) and Panamax (PNMX) dry-bulk vessels. The sample extends from 13 January 1989 to 26 June 2009, thus comprising 1,068 observations, where all data was collected from Clarksons Shipping Intelligence Network.

Calculated descriptive statistics indicate that there are contrasting size effects in the tanker and dry-bulk sectors, where, in the tanker sector, larger vessels are found to exhibit lower returns, which may be as a result of the reduced flexibility of these vessels with respect to smaller vessels. The reverse is found in the dry-bulk sector, however this is attributed to the increased flexibility of dry-bulk vessels, which may enable the vessels to take advantage of economies of scale. Returns, across all vessel types, are found to exhibit excess kurtosis and significant positive skewness, with the exception of the Panamax data series, where returns are negatively skewed. These findings are supported Jarque-Bera statistics, where the null hypothesis of normally distributed returns is rejected in all cases.

Results from the Ljung-Box tests on the squared returns indicate the presence of significant ARCH effects, at the 1st, 12th and 24th lags, across all vessel-types, with the possible exception of the 1st lag of the VLCC data series, where ARCH effects are only significant at the 5% and 10% levels of significance. This finding indicates the appropriateness of ARCH-type models for modelling the volatility of the respective returns. Furthermore, the existence of volatility clustering in the freight rate returns gives further indication that there may be GARCH effect, thus indicating the necessity of the GARCH-type modelling.

Looking at the characteristics of the third and fourth moments, the results from Ljung-Box tests performed on the third moment of the series suggest that there is something of a sector effect, at least at the 1st lag of the series, in that the 1st lag of the tanker series do not exhibit significant conditional skewness, whereas those for the dry-bulk series do. This, however, does not carry across to the 12th and 24th lags, where all data series, regardless of sector, exhibit significant conditional skewness. When examining the results of these tests for the fourth moment, one should note that the Panamax data series does not exhibit significant conditional kurtosis at the 1st and 24th lags, and any conditional kurtosis at the 12th lag is only significant at the 5% and 10% levels of significance. One should also note that any conditional skewness at the 12th lag of the Capesize data series is only significant at the 10% level of significance, but is significant at all conventional levels of significance for the other lags. Furthermore, one should note that the tanker data series do not exhibit any conditional kurtosis at the 1st lag of the series; however, all other lags of these series exhibit significant conditional kurtosis at all conventional levels of significance.

One should note that these models were also run on the natural logarithms of the respective data series in order to take account of the fact that freight rates can never be negative. One should note that the results from these estimations did not differ significant from those presented here[24].

Having outlined the characteristics of the data series, the next section focuses on the estimation of the respective models, giving insight into the applications and implications of these results.

7.4 Empirical Results

7.4.1 Introduction

Having outlined the methodology and characteristics of the data to be used in the analysis, this section presents the results of the empirical work in which a comparison was made between the performances of the standard GARCH model developed by Bollerslev (1986), the FIGARCH model, outlined in Baillie, *et al.* (1996a), and the GARCHSK model developed in the paper by León, *et al.* (2005). Following the initial examination of the Ljung-Box statistics in the descriptive statistics section, the three models were estimated using quasi-maximum likelihood estimation.

7.4.2 Results for the ARFIMA-GARCH Models

So as to avoid repetition, this sub-section provides a brief summary of the results for the ARFIMA-GARCH models, where a more detailed discussion of these findings can be found in Chapter 6 of the thesis. Beginning with actual empirical analysis of the data, the results for which are presented in Table 6.1, the mean equation coefficients suggest that, in all cases, returns display significant autoregressive properties, while, for the Panamax data series, past unforeseen appear to the have a significant effect on the current returns. In terms of the variance equation, current shocks appear to have a significant impact on the conditional volatility for all data series, as does past variance, while the impact of shocks are found to persist almost indefinitely, where this is measured by the sum of the α and β coefficients. For convenience, the evolutions of the conditional volatility for each data series are

[24] The results from these estimations are not presented here due to space constraints and are available from the author upon request.

graphically presented in Figure 6.1, while the results of the residual diagnostics indicated that the model is well-specified.

7.4.3 Results for the ARFIMA-FIGARCH Models

As with the results for the GARCH model, so as to avoid repetition, this sub-section provides a brief synopsis of the results for the ARFIMA-FIGARCH model, where, as before, a more in-depth review of these findings can be found in Chapter 6 of the thesis. Begin with the results of the actual empirical model itself, where these results are summarised in Table 6.5, one should note that the mean equation coefficients are identical to those obtained for the GARCH model above. When looking at the variance equation, past variance is found to have a significant effect on the current volatility, where, in this case, this coefficient measures the persistence of volatility. However, the difference between these models is that in this case, the rate of decay of shocks to volatility is found to be much faster than for the GARCH model. The evolutions of the conditional volatility for each data series are once again graphically presented in Figure 6.3, while the results of the residual diagnostics indicated that the model is well-specified.

7.4.4 Results for the ARFIMA-GARCHSK Models

Having examined the results for the GARCH and FIGARCH models in the sub-sections above, one should note that the GARCHSK model is different from the GARCH and FIGARCH models in that it enables one to model conditional skewness and kurtosis in addition to the conditional variance. This section presents the results from the estimation of the $\text{ARFIMA}(1;d;0)$-$\text{GARCHSK}(1;1)$, in the case of the tanker series, and $\text{ARFIMA}(1;d;1)$-$\text{GARCHSK}(1;1)$, for the dry-bulk series, models, the results for which are presented in Table 7.1.

As was the case for the FIGARCH models, the results of the mean equation are identical to the GARCH models in Chapter 6 and therefore do not need any more discussion. Examining the variance equation, the long-run average variance parameter, β_0, is found to be significant at all conventional levels of significance for

Table 7.1 – Empirical Results for the ARFIMA-GARCHSK Model

$$r_t = \mu_1 + \mu_2 r_{t-1} + \varepsilon_t$$
$$\varepsilon_t = h^{-1/2}\eta_t \ ; \ \eta_t \sim N(0;1) \ ; \ \varepsilon_t \big| I_{t-1} \sim (0;h_t)$$
$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1}$$
$$s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1}$$
$$k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1}$$

Panel A – Results for the Mean Equation

	VLCC	SZMX	AFMX	CPSZ	PNMX
μ	0.004	0.006	0.005	0.002	0.002
	(0.363)	(0.221)	(0.165)	(0.577)	(0.345)
AR(1)	0.194	0.171	0.098	0.410	0.143
	(0.000)	(0.000)	(0.001)	(0.000)	(0.004)
MA(1)	-----	-----	-----	0.013	0.159
	(-----)	(-----)	(-----)	(0.802)	(0.005)

Panel B – Results for the Variance Equation

	VLCC	SZMX	AFMX	CPSZ	PNMX
β_0	0.000	0.000	0.007	0.000	0.000
	(0.008)	(0.019)	(0.000)	(0.000)	(0.525)
β_1	0.075	0.094	0.379	0.107	0.111
	(0.000)	(0.000)	(0.093)	(0.000)	(0.052)
β_2	0.907	0.892	0.079	0.855	0.837
	(0.000)	(0.000)	(0.197)	(0.000)	(0.000)

Panel C – Results for the Skewness Equation

	VLCC	SZMX	AFMX	CPSZ	PNMX
γ_0	0.045	0.176	-----	0.035	0.012
	(0.009)	(0.000)	(-----)	(0.081)	(0.974)
γ_1	0.007	0.010	0.000	0.005	0.006
	(0.083)	(0.017)	(0.558)	(0.067)	(0.875)
γ_2	0.766	0.441	0.998	0.626	-----
	(0.000)	(0.000)	(0.000)	(0.000)	(-----)

Panel D – Results for the Kurtosis Equation

	VLCC	SZMX	AFMX	CPSZ	PNMX
δ_0	6.043	4.056	3.880	6.264	3.366
	(0.000)	(0.000)	(0.041)	(0.000)	(0.000)
δ_1	0.035	0.007	0.001	0.000	0.000
	(0.000)	(0.000)	(0.002)	(1.000)	(0.986)
δ_2	0.157	0.702	0.759	0.167	0.900
	(0.060)	(0.000)	(0.000)	(0.000)	(0.000)
LL	1975.023	1920.805	1911.193	2748.252	2690.728
AIC	-3.697	-3.595	-3.577	-5.145	-5.037
SBIC	-3.696	-3.594	-3.576	-5.144	-5.036

Note 1: See notes from Table 6.1.

the VLCC, Aframax and Capesize data series and at the 5% and 10% levels of significance for the Suezmax data series. However, for the Panamax data series, long-run average variance is found to be insignificant. One can thus state that expected volatility is found to play a significant role in spot freight rate volatility in all but the Panamax market, where it appears that there also may be a size effect in the dry-bulk market, which may be as a consequence of smaller vessels being able to minimise risk by being able to take smaller cargoes.

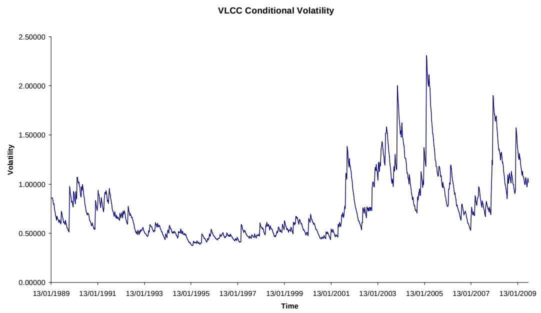
The impact of shocks, as measured by the β_1 coefficient, is found to be significant at all conventional levels for the VLCC, Suezmax and Capesize data series, at the 10% level of significance for the Aframax data series and at the 5% and 10% levels of significance. Once again there appears to be a size effect, this time across both markets, which may be as a result of smaller vessels being able to minimise the impact of shocks as a result of being able to take smaller cargoes. Finally, past variance forecasts, as indicated by β_2, are found to have a significant impact on volatility, at all levels of significance, for all but the Aframax data series, where for this data series the parameter is insignificant at all conventional levels of significance. It should be noted here that the level of persistence in volatility is very high, thus indicating that an IGARCH or FIGARCH model may be more appropriate in terms of modelling the conditional volatility.

Changing focus to the higher moments, the skewness equation provides some interesting results. The impact of the long-run average skewness, denoted γ_0, is found to be significant at all conventional levels of significance for the VLCC and Suezmax data series and at the 10% level of significance for the Capesize data series. In contradistinction, the long-run average skewness is found to be insignificant, at all conventional levels of significance, for the Aframax and Panamax data series. The size effect observed here may be caused by a consistent supply of vessels in the markets for smaller vessels as larger vessels are forced to lay-up in poor market conditions, thus skewing the distribution of freight rate returns. Examining the impact of shocks on the conditional skewness, as measured by the γ_1 coefficient, this is found to be significant at all conventional levels of significance for the Suezmax data series, at the 10% level of significance for the VLCC and Capesize data series. However, shocks are found to have an insignificant effect on skewness for the smaller Aframax and Panamax vessels. The same explanation is given for this size effect as above in that inconsistency in the supply of larger vessels may skew the distribution of freight rate returns. Finally, forecasts of skewness, as measured by γ_2, are found to have a significant impact, at all conventional levels of significance, and therefore past skewness is found to have an impact on current conditional skewness, for all but the Panamax data series.

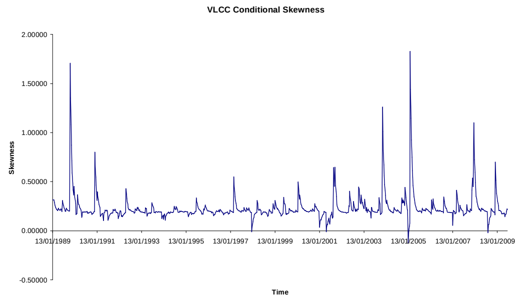
To complete the analysis of the model estimations the determinants of conditional kurtosis are examined. The long-run average kurtosis, denoted δ_0, is found to be significant at all conventional levels of significance, with the exception of the Aframax data series where it is significant at the 5% and 10% levels of significance. The impact of shocks on kurtosis, as illustrated by δ_1, is found to be significant, at all conventional levels of significance, for the VLCC, Suezmax and Aframax data series, but insignificant for the Capesize and Panamax data series. A potential explanation for this market effect is that should market conditions deteriorate, dry-bulk vessels are able to switch cargoes, say from iron ore to bauxite, thus enabling them to be more consistently employed and minimising the probability of extreme returns. To conclude the analysis of the determinants of conditional kurtosis, the impact of past kurtosis forecasts, as measured by δ_2, is examined. The parameter δ_2 is found to significant at all conventional levels of significance for all but the VLCC data series. This indicates that past conditional kurtosis has a significant impact on the current conditional kurtosis for but the VLCC data series. One should note that the conditional variance, skewness and kurtosis series for each vessel type generated by the model are graphed in Figure 7.1 to Figure 7.5, respectively. One can see from these graphs, that where the respective coefficients are significant, there does appear to be significant variation in the skewness and kurtosis of the series across time, therefore indicating the appropriateness of modelling these moments conditionally. A further point to note is that the conditional variance seems to be very close to that obtained from the standard GARCH model, which are graphically illustrated in Figure 6.1 in Chapter 6 of this thesis. The reason for this similarity is that the variance equation in the GARCHSK model is exactly the same as that for the GARCH model, hence the similarity in the conditional variance.

Figure 7.1 – Conditional Moments for GARCHSK Model of the VLCC Series

(a) Conditional Variance



(b) Conditional Skewness



(c) Conditional Kurtosis

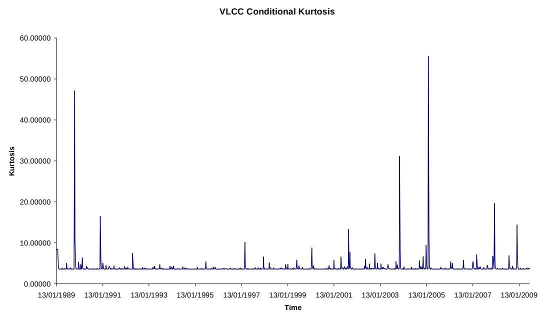
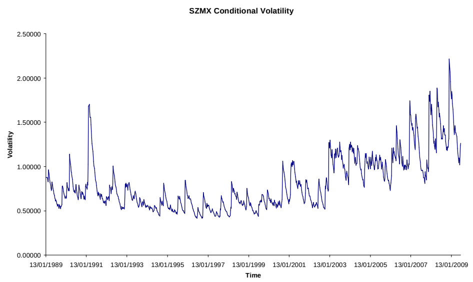
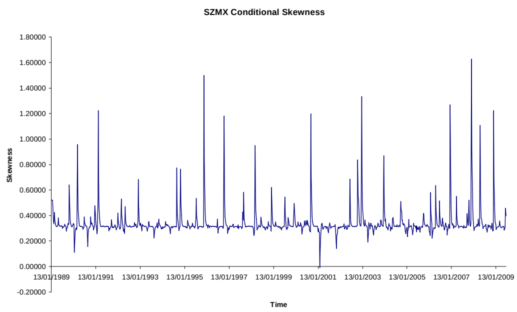


Figure 7.2 – Conditional Moments for GARCHSK Model of the SZMX Series

(a) Conditional Variance



(b) Conditional Skewness



(c) Conditional Kurtosis

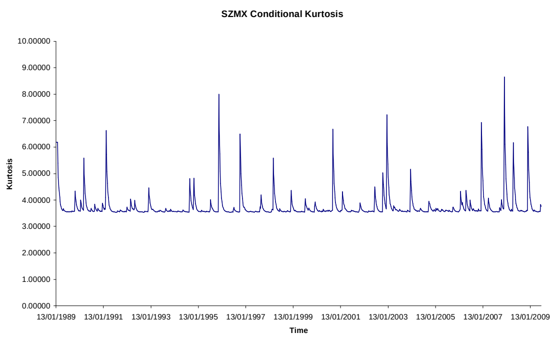
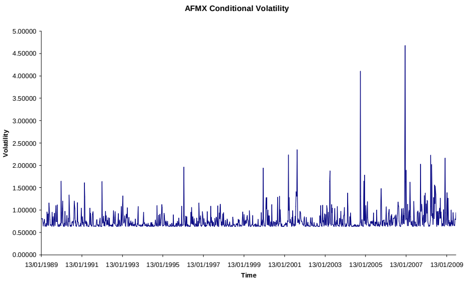
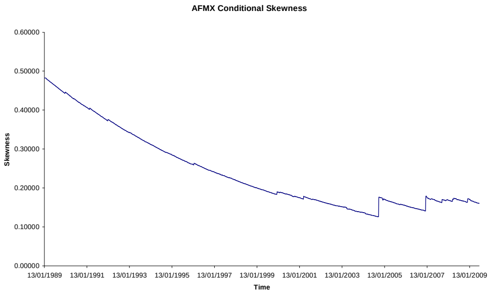


Figure 7.3 – Conditional Moments for GARCHSK Model of the AFMX Series

(a) Conditional Variance



(b) Conditional Skewness



(c) Conditional Kurtosis

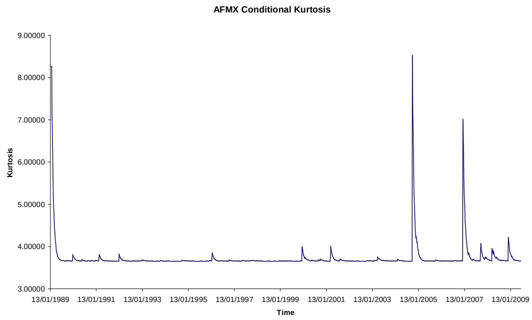
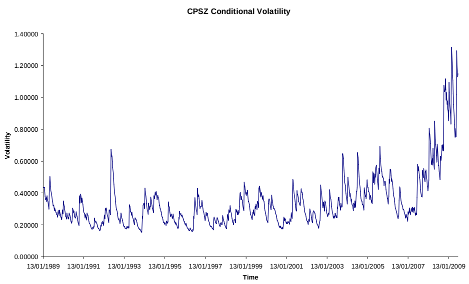
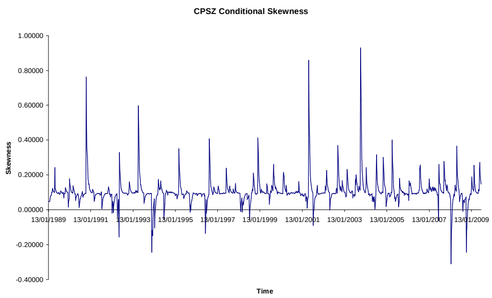


Figure 7.4 – Conditional Moments for GARCHSK Model of the CPSZ Series

(a) Conditional Variance



(b) Conditional Skewness



(c) Conditional Kurtosis

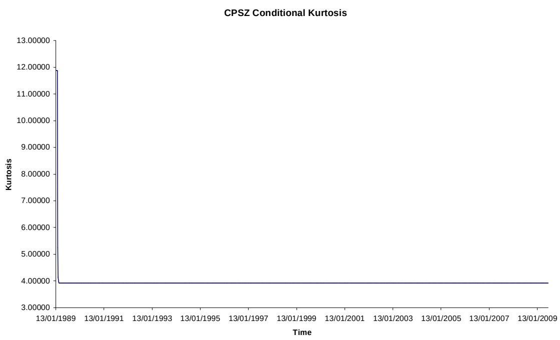
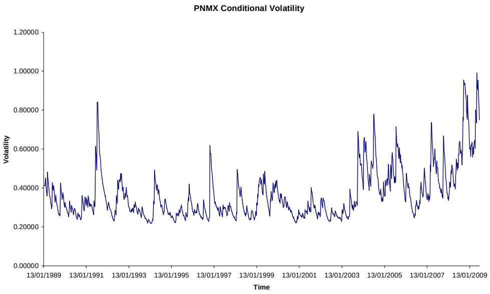
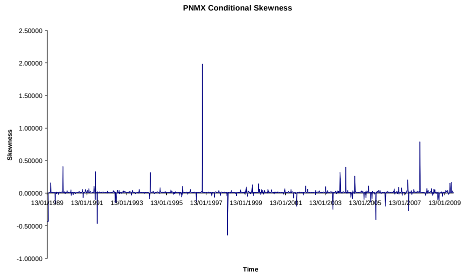


Figure 7.5 – Conditional Moments for GARCHSK Model of the PNMX Series

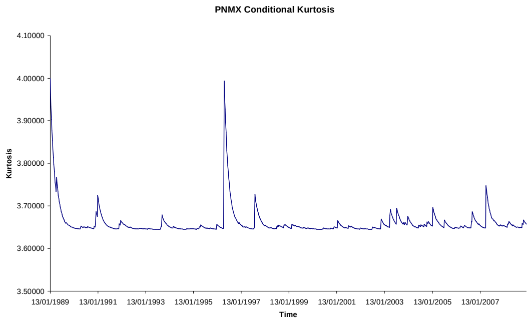
(a) Conditional Variance



(b) Conditional Skewness



(c) Conditional Kurtosis



In order to test the results of the GARCHSK models, residual diagnostics were performed, with the results of these diagnostics presented in Table 7.2. Prior to estimating the models, all data series displayed significant autocorrelations at all conventional levels of significance and across all lags. Examining the post-estimation results for the 1st lag, any autocorrelation at this lag has been removed for all but the Suezmax data series; however, when looking at the 12th and 24th lags, the results become somewhat more mixed. At the 12th lag, the autocorrelation for the Panamax data series is found to have been removed In addition the Capesize data series is found to only display significant autocorrelation at the 10% level of significance, while the VLCC displays significant autocorrelation at the 5% and 10% levels of significance. However, for the Suezmax and Aframax data series, significant autocorrelation is found to persist at all conventional levels of significance. The picture changes once again at the 24th lag in that the VLCC, Capesize and Panamax data series are found to have had any autocorrelation removed, whereas the Suezmax and Aframax data series still display significant autocorrelation at all conventional levels of significance.

Table 7.2 – Residual Diagnostics for the ARFIMA-GARCHSK Model

	VLCC	SZMX	AFMX	CPSZ	PNMX
Average	0.027	-0.002	0.027	0.019	0.017
Variance	1.090	1.077	1.123	1.330	1.118
Skewness	1.062	1.324	1.689	0.531	0.289
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	4.032	3.883	10.220	2.983	3.709
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	923.024	982.077	5150.756	445.621	626.376
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Q(1)$	0.212	4.869	0.339	0.817	0.005
	(0.645)	(0.027)	(0.560)	(0.366)	(0.944)
$Q(12)$	21.750	36.874	44.332	19.144	4.619
	(0.040)	(0.000)	(0.000)	(0.085)	(0.970)
$Q(24)$	31.348	48.500	93.969	30.060	19.516
	(0.144)	(0.002)	(0.000)	(0.183)	(0.724)
$Q^2(1)$	0.055	0.325	0.783	0.605	0.012
	(0.815)	(0.568)	(0.376)	(0.437)	(0.913)
$Q^2(12)$	10.472	12.821	13.836	15.304	5.331
	(0.575)	(0.382)	(0.311)	(0.225)	(0.946)
$Q^2(24)$	16.751	19.996	18.009	19.564	10.532
	(0.859)	(0.697)	(0.803)	(0.721)	(0.992)
$Q^3(1)$	0.225	0.000	0.004	0.346	0.008
	(0.635)	(0.994)	(0.949)	(0.557)	(0.929)
$Q^3(12)$	11.356	6.511	4.661	2.251	0.769
	(0.499)	(0.888)	(0.968)	(0.999)	(1.000)
$Q^3(24)$	13.402	17.410	5.656	6.890	2.885
	(0.959)	(0.831)	(1.000)	(1.000)	(1.000)
$Q^4(1)$	0.094	0.085	0.018	0.011	0.030
	(0.759)	(0.770)	(0.894)	(0.917)	(0.862)
$Q^4(12)$	2.678	4.621	0.775	4.257	0.714
	(0.997)	(0.969)	(1.000)	(0.978)	(1.000)
$Q^4(24)$	3.915	13.864	0.937	6.489	1.443
	(1.000)	(0.950)	(1.000)	(1.000)	(1.000)

Note 1: See notes from Table 6.1.

Changing focus to the ARCH effects, all series prior to estimating the model, with the exception of the VLCC series, which only exhibited significant ARCH effects at the 5% and 10% levels of significance at the 1st lag, exhibited significant ARCH effects at the 1st, 12th and 24th lags. Having estimated the model, however, any ARCH effects are found to have been removed from all data series at all lags. Looking at the third and fourth moments, the results of the Ljung-Box tests on the standardised residuals indicate that any conditional skewness or kurtosis that existed prior to the estimation of the models has been removed.

One comes to somewhat of a conundrum here in that although the GARCHSK model does not account for the significant persistence of the volatility in the data series, it does allow one to model the structure of the higher moments of the returns series. In contradistinction, the FIGARCH model captures the persistence in volatility, but does not allow for the modelling of the higher moments. A trade-off is therefore necessary in either model. In order to determine the best model, a comparison of the characteristics of each of the three models is made in the following section.

7.5 Comparison of the Estimated Models

Having estimated the GARCH, FIGARCH and GARCHSK models in the section above, this section makes a comparison between the models in order to select the model that best fits the data. The first method used to compare the performance of the models is to compare the characteristics of the conditional variance. In addition to this, likelihood-ratio tests are performed between the various models and the results analysed. This section presents the results of these analyses and provides a recommendation as to which model may best model the conditional variance of freight rate returns.

Table 7.3 presents the descriptive statistics for the conditional variance of the data series, as evaluated using the GARCH model. Average conditional variance was found to range between 0.003 and 0.015, with the average conditional variance in the tanker market being significantly higher than that for the dry-bulk market. This may be as a result of limitations in the cargo that may be carried by tankers as opposed to dry-bulk vessels. The variance of the conditional variance is found to be zero across all data series. Looking at the distribution of the conditional variance, all data series were found to exhibit significant positive skewness, at all conventional levels of significance. In addition to this, the Aframax, Capesize and Panamax data series were found to exhibit significant excess kurtosis, at all conventional levels of significance, and the distribution of the conditional variance was found to be significantly non-normal across all data series.

Table 7.3 – Conditional Variance Statistics for the ARFIMA-GARCH Model

	VLCC	SZMX	AFMX	CPSZ	PNMX
Average	0.015	0.015	0.013	0.004	0.003
Variance	0.000	0.000	0.000	0.000	0.000
Skewness	1.837	1.374	14.049	4.050	2.946
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	3.751	1.377	261.640	19.076	10.901
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	1226.864	420.555	3081391.918	19113.411	6833.432
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note 1: See notes from Table 6.1.

Table 7.4 – Conditional Variance Statistics for the ARFIMA-FIGARCH Model

	VLCC	SZMX	AFMX	CPSZ	PNMX
Average	0.015	0.015	0.014	0.004	0.003
Variance	0.000	0.000	0.000	0.000	0.000
Skewness	3.802	4.672	14.391	4.926	2.984
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	21.617	36.667	274.107	32.224	11.153
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	23367.962	63714.944	3380364.208	50527.365	7120.077
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note 1: See notes from Table 6.1.

Table 7.5 – Conditional Variance Statistics for the ARFIMA-GARCHSK Model

	VLCC	SZMX	AFMX	CPSZ	PNMX
Average	0.013	0.015	0.013	0.003	0.003
Variance	0.000	0.000	0.000	0.000	0.000
Skewness	2.693	2.122	15.519	4.176	2.865
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Kurtosis	9.581	5.800	305.944	20.605	10.233
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Jarque-Bera	5376.168	2298.408	4208146.232	21997.791	6120.978
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note 1: See notes from Table 6.1.

Looking at the descriptive statistics for the conditional variances obtained using the FIGARCH model, as presented in Table 7.4, these results only differ slightly to those obtained from the GARCH model. To this effect, the average conditional variance is found to range between 0.003 and 0.015, while the variance of the conditional variance is found to be zero. A market effect is with respect to the average conditional variance is once again observed. In addition, the conditional variance is found to be significantly positively skewed, at all conventional levels of significance, across all data series and all data series are found to be non-normally distributed. Where the results differ is in the fact that all data series are now found to exhibit significant excess kurtosis. A possible reason for this difference between the two models may be that, as the FIGARCH model captures the persistence of the conditional volatility, it also captures more of the dynamics of the conditional variance of freight rate returns.

The final conditional variance series to be examined was that of the GARCHSK model, as illustrated in Table 7.5. Although the average conditional variance is still found to range between 0.003 and 0.015, the average conditional variances for the VLCC, Aframax and Capesize data series are found to be lower. However, the variance in the conditional variance is still found to be zero. Changing focus to the distribution of the conditional variance, all conditional variance series are found to be significantly positively skewed and illustrate significant excess kurtosis, at all conventional levels of significance. A similar explanation is given for this difference between the results for the GARCH and GARCHSK, as is given above. This is that as the GARCHSK model captures the dynamics of the higher moments of the freight rate returns, it effectively also captures more of the dynamics of the conditional variance of freight rate returns.

Having examined the characteristics of the conditional variance, all that has been established in terms of the selection of the best model is that the FIGARCH and GARCHSK models are both preferred to the GARCH model. In order to provide a definitive answer to this hypothesis, likelihood-ratio tests were performed on the models, the results of which are presented in Table 7.6. The results for the comparison between the GARCH are FIGARCH model are somewhat mixed in that the GARCH model is preferred for the VLCC data series, as well as the Capesize data series at the 1% level of significance, whereas the FIGARCH model is preferred for all other data series. However, the GARCHSK model is preferred over both other models across all data series. Therefore one can conclude from this that the GARCHSK provides the best determinant of the conditional moments in the freight rate returns markets.

Table 7.6 – Likelihood-Ratio Test Results for Comparison of Models

Panel A – Comparison of GARCH and FIGARCH Models					
	VLCC	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
log L (GARCH)	924.412	850.050	844.368	1675.732	1654.546
log L (FIGARCH)	924.328	845.788	856.412	1678.321	1662.948
LR Statistic	0.166	8.525	24.088	5.179	16.805
	(0.684)	(0.004)	(0.000)	(0.023)	(0.000)

Panel B – Comparison of GARCH and GARCHSK Models					
	VLCC	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
log L (GARCH)	924.412	850.050	844.368	1675.732	1654.546
log L (GARCHSK)	1935.598	1902.352	1908.695	2748.177	2695.876
LR Statistic	2022.372	2104.603	2128.655	2144.890	2082.660
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Panel C – Comparison of FIGARCH and GARCHSK Models					
	VLCC	**SZMX**	**AFMX**	**CPSZ**	**PNMX**
log L (FIGARCH)	924.328	845.788	856.412	1678.321	1662.948
log L (GARCHSK)	1935.598	1902.352	1908.695	2748.177	2695.876
LR Statistic	2022.538	2113.128	2104.567	2139.711	2065.855
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Note 1: See notes from Table 6.1.

7.6 Conclusion

This paper examines, for the first time in the shipping literature, to the best of the authors' knowledge, the issues of time-varying skewness and kurtosis in shipping freight rate returns in order to determine the most comprehensive model in order to determine the dynamics of the returns distribution. In order to do this, the GARCHSK model, the FIGARCH model, and the standard GARCH model are estimated, using freight rate returns from five different vessel types over 1,068 observations, and the results compared by looking at the characteristics of the respective conditional variance and using likelihood-ratio tests.

In the quasi-maximum likelihoods estimations of the GARCH and GARCHSK models, the GARCHSK model was found to capture more of the dynamics of the respective data series based on these results, and the conditional skewness and kurtosis parameters were found to be significant across most of the data series. In addition, the results of the likelihood ratio tests uniformly indicated the superiority of the GARCHSK model over the standard GARCH and the FIGARCH models, a fact that was confirmed through a comparison of the characteristics of the conditional variance for the respective data series. Therefore, this paper can conclude quite firmly that the GARCHSK model outperforms the GARCH and FIGARCH models in capturing the dynamics of the data.

There are multiple advantages to being able to capture the conditional skewness and kurtosis of the data series - the first, is that as the distribution of asset returns is skewed, and there is excess kurtosis, and therefore the traditional assumption of normality when estimating Values-at-Risk will result in an underestimation of the risk. Secondly, it enables one to better describe the distributional properties of financial asset returns, thus enabling one to better understand the performance of assets with these properties. Finally, one could look at the issue of portfolio construction to determine if the risk structure is truly optimal, and examining these properties would enable one to better price options in financial markets wherever these properties exist.

The following chapter examines the first of these advantages in that Values-at-Risk are calculated using forecasts of the volatility of the underlying series, where these are generated using the GARCH, IGARCH, FIGARCH and GARCHSK models discussed in the thesis, in addition to some standard Value-at-Risk methodologies. By doing this, one first outlines a potential practical application of these models, but can make actual real-world comparisons of the models.

8 A Practical Application of Freight Rate Modelling

8.1 Introduction

The previous two chapters of the thesis focused on the behaviour of freight rate volatility over the entire sample period, this chapter changes the focus slightly by examining the out-of-sample forecasting performance of the respective models, using standard forecast accuracy measurements and the Value-at-Risk methodology to gauge which of the aforementioned models performs best.

The reason behind the interest in the topic is that shipping provides the primary means of transportation for almost any good traded across the world, therefore the process of evaluating the correct structure of freight rates is essential for any participants in world trade. As supply and demand shifts in the freight market, so freight rates should adjust to the equilibrium price level; however, due to constrains in terms of the structure of supply in the market, this process of adjustment to equilibrium may be delayed. As a result of this, accurate forecasts of future freight rates become essential for both investment decisions for ship-owners and banks that finance shipping activities, as well as for charterers and other auxiliary parties with respect to planning their transportation requirements, where this was addressed in Chapter 5 of the thesis.

Obviously, however, parties in the financial markets are not only interested in the future levels of prices alone but also with the potential fluctuation in these prices about the predicted levels, where these fluctuations, known as the level of volatility in the market, and represent the risk in the market. One commonly used method to discern between models forecasting is to estimate the respective Value-at-Risk (VaR) on a portfolio for the different models available, where, as Christoffersen (2003) points out, the VaR is defined as the dollar, or percentage, loss that will be only be exceeded a given percentage of the time over the forecast horizon. It is worth noting that VaR has become a standard risk management tool and has been adopted by the Basel Committee as a standard method to quantify market risk, however, any chosen VaR methodology must be backtested so as to verify its accuracy, for which a wide range of tools can be applied, as highlighted in Basel (1995a) and Basel (1995b).

An extensive literature exists on the performance of the various VaR methodologies in the conventional equity and bond markets, some of which include Brooks and Persand (2003), Giot and Laurent (2003b), Kuester, *et al.* (2006) and Lehar, *et al.* (2002). In addition to this, a large literature on the issue also exists in the commodity, energy and hedge fund markets, such as Cabedo and Moya (2003), Giot and Laurent (2003a), Krehbiel and Adkins (2005) and Sadorsky (2006), amongst others. In contrast, in the shipping markets there is, to the best of the author's knowledge, a relative dearth of literature on the topic, although recently Kavussanos and Dimitrakopoulos (2007) apply a VaR approach to modelling risk in the tanker freight markets, and Angelidis and Skiadopoulos (2008) apply the VaR methodology to modelling Baltic freight indices in both the dry and wet markets, with mixed results. This chapter uses the VaR approach to distinguish the best model, amongst those discussed in Chapters 6 and 7, to forecast freight rate volatility and seeks to extend the literature by, unlike Kavussanos and Dimitrakopoulos and Angelidis and Skiadopoulos, looking at both the tanker and dry bulk markets, and looking at the actual freight rate series on specific routes as opposed to merely the freight indices. In addition, this chapter introduces the concept of the generalised autoregressive conditional heteroscedasticity with skewness and kurtosis (GARCHSK) model to the VaR literature. One should note, however, that the Value-at-Risk here is not a true Value-at-Risk as one is not trading a portfolio, but effectively measures the potential loss (increase) in profits (costs) incurred by the ship-owner (charterer).

Sections two and three of the chapter introduce the relevant methodologies and describe the data used, respectively. Section 4 looks at the results of the VaR estimation and some preliminary forecast statistics, while section 5 examines the results from backtesting and discusses the best model selection process. To conclude, section 6 gives an overview of the paper and discusses the findings.

8.2 Forecast Accuracy and Value-at-Risk Methodologies

In this section the various methodologies used to measure the accuracy of the volatility forecasts, in addition to the VaR methodologies, both for VaR estimation and backtesting, are discussed. The models used for estimating and forecasting spot freight rate volatilities are exactly those used and discussed in Chapters 5, 6 and 7 of this thesis, however, these have been reviewed in detail and are not discussed further in this chapter. [25]

8.2.1 Measuring the Accuracy of the Volatility Forecasts

Having estimated the initial models over the in-sample period, ex-post forecasts of spot freight rate levels are estimated, and the results evaluated using standard tests, namely the Root Mean Squared Error (RMSE), in addition to the number and percentage of over- and under-predictions. The RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(y_i^a - y_i^f\right)^2} \qquad (8.1)$$

One should note that in Expression (8.1), y_i^a denotes the actual observed volatility at time *i*, y_i^f denotes the forecasted volatility at time *i*, and *M* denotes the forecast horizon.Continuing, the Percentage of Correction Direction Predicted (CDIR) is defined as the percentage of forecasts in which the forecast correctly predicted an increase or decrease in the actual volatility. In order to calculate the CDIR, one must first determine whether the sign of the forecast matches the sign of the actual values, i.e. when an increase (decrease) in the forecast matches an increase (decrease) in the actual values, where if this occurs, it is denoted C_i and M denotes the respective forecast horizon, therefore:

$$CDIR = \frac{1}{M}\sum_{i=1}^{M} C_i \qquad (8.2)$$

One should note that in Expression (8.2), above, *M* denotes the forecast horizon. Having outlined these methodologies, one can now look at how to calculated the Value-at-Risk for the various models.

One should note that the accuracy of standard forecast metrics as applied to forecasts of volatility are called into question. The reason for this is that using squared returns as a proxy for actual volatility is inherently flawed, where Lopez (2001) illustrated that this proxy is over 50% greater or smaller than the actual volatility 75% of the time. This being said, unfortunately no better proxy exists as yet, however, this thesis

[25] A more extensive discussion of the various methodologies can be found in Chapter 3 of the thesis.

addresses this issue by evaluating the accuracy of the respective Values-at-Risk for each model with the actual Value-at-Risk incurred.

8.2.2 A Brief Look at Value-at-Risk

A commonly used tool to quantify market risk exposure is the Value-at-Risk (VaR), where as mentioned above, the VaR is defined as the dollar, or percentage, loss that will only be exceeded a given percentage of the time over the forecast horizon, where:

$$\mathrm{VaR}_{t+i}^{\alpha} = \sigma_{t+i}^{2} \times \Phi^{-1}\left(\alpha\right) \quad (8.3)$$

In Expression (8.3), above, σ_{t+i}^{2} denotes the standard deviation of the series at forecast period *i*, where $i = 1;\ldots;M$ and M is the forecast horizon, and $\Phi^{-1}\left(\alpha\right)$ denotes the inverse of the cumulative density function of the standard normal distribution, where α denotes the coverage rate. In addition to the GARCH, IGARCH, FIGARCH and GARCHSK models discussed above, RiskMetrics$^{\mathrm{TM}}$ (RM), developed by JP Morgan (1996), Historical Simulation (HS) and Filtered Historical Simulation (FHS) methods were applied to replicate the volatility series. This sub-section outlines the VaR methodologies, where these will be used to differentiate between these volatility models, as the best approach will be that which minimises the risk exposure, and hence potential losses.

It is important to note what is meant by the VaR in the context of this thesis. As these measures are not being used to trade a portfolio. In the case of ship-owners, the VaR figure corresponds to the potential loss in terms of reduced profits from operating the vessels; hence one could term it the Profit-at-Risk in this case. On the other side of the market, from the charterer’s point-of view, the VaR figure corresponds to the potential loss in terms of increased costs of transportation, hence one could term this the Cost-at-Risk in this regard.

It is worth highlighting the fact that, as mentioned briefly above, the VaR estimation is based on a standard normal distribution. However, what happens when the series is not normally distributed, as the argument for the use of the GARCHSK model would suggest? In order to address this issue, the Cornish-Fisher (CF) expansion was applied when estimating the GARCHSK model’s variable.

The CF can be written as:

$$\mathrm{CF}^{-1}(\alpha)=\Phi_{\alpha}^{-1}+\frac{\xi_{1;i}}{6}\left[\left(\Phi_{\alpha}^{-1}\right)^{2}-1\right]+\frac{\xi_{2;i}}{24}\left[\left(\Phi_{\alpha}^{-1}\right)^{3}-3\Phi_{\alpha}^{-1}\right] \\ -\frac{\xi_{1;i}^{2}}{36}\left[2\left(\Phi_{\alpha}^{-1}\right)^{3}-5\Phi_{\alpha}^{-1}\right] \qquad (8.4)$$

In the Expression (8.4), above, $\xi_{1;i}$ and $\xi_{2;i}$ denote the conditional skewness and kurtosis at forecast period *i*, respectively, $\Phi^{-1}(\alpha)$ denotes the inverse of the cumulative density function of the standard normal distribution, and α denotes the coverage rate.

Having said this, one could therefore calculate the CF VaR, which would be the equivalent to the VaR for the GARCHSK model, as follows:

$$\mathrm{VaR}_{t+i}^{\alpha}=\sigma_{t+i}^{2}\times\mathrm{CF}_{\alpha}^{-1} \qquad (8.5)$$

In Expression (8.5), above, σ_{t+i}^{2} denotes the standard deviation of the series at forecast period *i*, where $i=1;\ldots;M$ and M is the forecast horizon, and $\mathrm{CF}^{-1}(\alpha)$ denotes the Cornish Fisher expansion, outlined in Expression (8.4), above, where α denotes the coverage rate.

The RM methodology is similar to an IGARCH in that it constrains the sum of the ARCH and GARCH parameters to equal one, however, unlike the IGARCH model, these parameters are fixed, as opposed to being estimated by the model, hence for the RM model, the forecasted variance will be:

$$\sigma_{i}^{2}=\lambda\sigma_{t}^{2}+(1-\lambda)r_{t}^{2} \qquad (8.6)$$

In Expression (8.6), above, σ_{t+i}^{2} denotes the standard deviation of the series at forecast period *i*, where $i=1;\ldots;M$ and *M* is the forecast horizon, σ_{t}^{2} is the standard deviation over the previous 52 weeks at time *t* and r_{t}^{2} is the squared return at time *t*. One should note that, in this chapter, λ is fixed at 0.95, where this was taken as an average of the values given for the monthly and daily frequencies in JP Morgan (1996), as the frequency of the data here is weekly.

The HS methodology assumes that the distribution of tomorrow's returns, denoted r_{t+1}, is well approximated by the empirical distribution of the past N observations, or in other words, the distribution of r_{t+1} is captured by the histogram of $\left\{r_{t+1-\tau}\right\}_{\tau=1}^{N}$. One can thus state that the VaR, assume a coverage rate of α, using the HS technique, will simply be calculated as the 100α th percentile of this sequence of past returns, which means that:

$$\text{VaR}_{t+i}^{\alpha} = Percentile\left[\left\{r_{t+i-\tau}\right\}_{\tau=1}^{N};\ 100\alpha\right] \qquad (8.7)$$

In Expression (8.7), above, $\text{VaR}_{t+i}^{\alpha}$ denotes the VaR at forecast point *i*, $\left\{r_{t+i-\tau}\right\}_{\tau=1}^{N}$ denotes the distribution of returns over the horizon $\tau = 1;\ldots;N$, and α denotes the desired coverage rate.

In contrast, the FHS methodology still uses the same approach of taking the empirical distribution of the past *N* observations, however, instead of taking the distribution of returns, it takes the distribution of the standardised returns, where these are standardised using the forecasted volatility, hence:

$$\hat{z}_{t+i-\tau} = \frac{r_{t+i-\tau}}{\sigma_{t+i-\tau}} \qquad (8.8)$$

In Expression (8.8), above, $\hat{z}_{t+i-\tau}$ denotes the standard return at forecast point *i*, $r_{t+i-\tau}$ denotes the return at forecast point *i*, and $\sigma_{t+i-\tau}$ denotes the standard deviation at forecast point *i* where $\tau = 1;\ldots;N$. One should note that set of standardised returns in denoted $\left\{\hat{z}_{t+i-\tau}\right\}_{\tau=1}^{N}$. The VaR using the FHS technique is thus calculated as follows:

$$\text{VaR}_{t+i}^{\alpha} = \sigma_{t+i} \times Percentile\left[\left\{\hat{z}_{t+i-\tau}\right\}_{\tau=1}^{N};\ 100\alpha\right] \qquad (8.9)$$

In Expression (8.9), $\text{VaR}_{t+i}^{\alpha}$ denotes the VaR at forecast point *i*, $\left\{\hat{z}_{t+i-\tau}\right\}_{\tau=1}^{N}$ denotes the distribution of the standardised returns over the horizon $\tau = 1;\ldots;N$, and α denotes the desired coverage rate.

The HS and FHS methodologies differ from the other techniques described in that they do not make any parametric assumptions regarding the distribution of the standardised returns. In addition to this, the HS methodology is also model-free in that it does not rely on any parametric model to generate the variance of the standardised returns. The advantages of the HS methodology are that it is easy to implement and that it does not need to incorporate any modelling in order to calculate the risk exposure. The fact that it is model-free is a definite advantage if the models of volatility are poor; however, the fact that it does not rely on a well-specified dynamic model mean that there is no theoretically correct way of extrapolating anything beyond the 1-day distribution. In addition, the sample length, denoted *N* above, is randomly chosen. This means that should *N* be too large, then the most recent observations, which theoretically should be important to determine the future distribution, will carry very little weight. On the other hand, should *N* be too small, one may not have incorporated sufficient large losses to be able to calculate the VaR with any accuracy. In contradistinction, the FHS methodology, by standardising the returns, enables one to have the advantages of the model-based approach, while still allowing one to follow what is a somewhat non-parametric approach.

If one is to be confident as to the potential risk exposure different models present, then the accuracy of the VaR estimates is crucial. In the following sub-section, the issue of evaluating the accuracy of these VaR estimates, using both statistical and economic approaches will be discussed.

8.2.3 Testing the Accuracy of Value-at-Risk Estimates

Having calculated the variable, it is of course essential to double check that these estimates are correct. There are two main approaches to this backtesting process, the first of which is the statistical approach, and the second, the economic approach. This sub-section presents the methodologies utilised to perform this backtesting.

Before going any further, one needs to establish what is meant by a violation of the VaR. This occurs when the observed return exceeds the stated VaR for a given observation, within the forecast horizon, and is also known as a hit. Ideally, the fraction of violations, relative to the forecast horizon, should be equal to the proposed coverage rate, i.e. for a 1% VaR, the fraction of violations, or hit ratio, should be equal to 1%. One should note that for a long position, a violation would occur when $\mathrm{VaR}_{t+i}^{\alpha} > R_{t+i}$, while for a short position this would occur when $\mathrm{VaR}_{t+i}^{\alpha} < R_{t+i}$, where $\mathrm{VaR}_{t+i}^{\alpha}$ denotes the VaR at forecast point i and R_{t+i} denotes the return at forecast point *i*. This would mean that one could define the sequence of VaR violations for a long position as:

$$I_{t+i} = \begin{cases} 1 & \text{if } -\mathrm{VaR}_{t+i}^{\alpha} > r_{t+i} \\ 0 & \text{if } -\mathrm{VaR}_{t+i}^{\alpha} < r_{t+i} \end{cases} \qquad (8.10)$$

In contrast, the sequence of VaR violations for a short position would be defined as:

$$I_{t+i} = \begin{cases} 1 & \text{if } \mathrm{VaR}_{t+i}^{\alpha} < r_{t+i} \\ 0 & \text{if } \mathrm{VaR}_{t+i}^{\alpha} > r_{t+i} \end{cases} \qquad (8.11)$$

One should note that in Expressions (8.10) and (8.11), 1 denotes a violation, 0 denotes a non-violation and I_{t+i} denotes the violation sequence at forecast point *i*. Therefore, one can construct a sequence of VaR violations, denoted $\left\{I_{t+i}\right\}_{i=1}^{M}$, where M denotes the forecast horizon, for the entire forecast horizon, thus indicating where past violations occurred. This "hit sequence" will be utilised in the statistical tests that follow.

Moving on to examine the first approach to backtesting, Christoffersen (2003) outlines three main tests to ensure the statistical accuracy of the VaR estimates, namely the Unconditional Coverage (UC), Independence (Ind) and Conditional Coverage (CC) tests. These tests require the construction of three likelihood functions, which will then be used to construct likelihood ratio statistics. The first of these is the likelihood function of the violation sequence, which is assumed to be an *i.i.d.* Bernoulli trial, where:

$$L(\hat{\pi}) = (1-\hat{\pi})^{M_0} \times (\hat{\pi})^{M_1} \qquad (8.12)$$

The second likelihood function is that for the coverage rate, where:

$$L(\alpha) = (1-\alpha)^{T_0} \times (\alpha)^{T_1} \qquad (8.13)$$

In Expressions (8.12) and (8.13), above, M_0 denotes the number of non-violations over the forecast horizon, M_1 denotes the number of violations over the forecast horizon, $\hat{\pi} = M_1 / M$, where M denotes the forecast horizon and $M_1 + M_0 = M$, and α denotes the desired coverage rate. If violations of the VaR are dependent across time, then the violation sequence can be described as a first-order Markov chain with the following transition probability matrix:

$$\hat{\Pi}_1 = \begin{bmatrix} \hat{\pi}_{0;0} & \hat{\pi}_{0;1} \\ \hat{\pi}_{1;0} & \hat{\pi}_{1;1} \end{bmatrix} = \begin{bmatrix} 1-\hat{\pi}_{0;1} & \hat{\pi}_{0;1} \\ 1-\hat{\pi}_{1;1} & \hat{\pi}_{1;1} \end{bmatrix} = \begin{bmatrix} \dfrac{M_{0;0}}{M_{0;0}+M_{0;1}} & \dfrac{M_{0;1}}{M_{0;0}+M_{0;1}} \\ \dfrac{M_{1;0}}{M_{1;0}+M_{1;1}} & \dfrac{M_{1;1}}{M_{1;0}+M_{1;1}} \end{bmatrix} \qquad (8.14)$$

Given the transition probability matrix in Expression (8.14), above, the likelihood function for the first-order Markov chain, which is the final likelihood function to be calculated, is given as follows:

$$L\left(\hat{\Pi}_1\right) = \left(1-\hat{\pi}_{0;1}\right)^{M_{0;0}} \times \left(\hat{\pi}_{0;1}\right)^{M_{0;1}} \times \left(1-\hat{\pi}_{1;0}\right)^{M_{1;0}} \times \left(\hat{\pi}_{1;0}\right)^{M_{1;1}} \qquad (8.15)$$

In Expression (8.15), above, $M_{i;j}$ denotes the number of observations with a j following an i, where $i; j = 0$ and $i; j = 1$ for non-violations and violations of the variable, respectively, and $\hat{\pi}_{i;j}$ is calculated as illustrated in transition probability matrix in Expression (8.14).

The Unconditional Coverage test evaluates whether the fraction of violations for a risk is significantly different from the coverage rate specified, or in other words, whether a model overestimates or underestimates the VaR. In order to do this, the null hypothesis that the proportion of violations, relative to the forecast horizon, is equal to the desired coverage rate is tested using a likelihood ratio statistic, where:

$$LR_{UC} = -2\ln\left[\frac{L(\alpha)}{L(\hat{\pi})}\right] \sim \chi^2_{(1)} \qquad (8.16)$$

In essence, what one is testing here is whether the model in question overestimates or underestimates the “true” but unobservable VaR, and thus the actual risk exposure. The disadvantage of this test is that although it tests for the degree by which the VaR estimate differs statistically from the true value, this estimate could still be dependent over time, thus large losses could follow directly after each other.

The Independence test addresses the issue of dependence in the VaR estimate by testing the null hypothesis that the VaR violation sequence is independently distributed, using the following likelihood ratio statistic:

$$LR_{Ind} = -2\ln\left[\frac{L(\hat{\pi})}{L(\hat{\Pi}_1)}\right] \sim \chi^2_{(1)} \qquad (8.17)$$

What the Independence test essentially establishes is whether losses in excess of the predicted VaR will be followed by other extreme losses, where, should this occur, one's risk exposure would be greatly increased, thus addressing the disadvantage of the Unconditional Coverage test. However, while it does this, it does not enable one to determine whether the estimated VaR overestimates or underestimates the true risk exposure, as would be the case for the Unconditional Coverage test.

The Conditional Coverage, which is the third, and final, statistical test, addresses both these issues by testing the null hypothesis that the violations of the VaR are independently distributed, and that the average number of these violations is correct. This is done using the following likelihood ratio statistic:

$$LR_{CC} = LR_{UC} + LR_{Ind} \sim \chi^2_{(2)} \qquad (8.18)$$

This test basically addresses the disadvantages of the former two tests and enables one to ensure that one's risk exposure is truly estimated.

The second approach to backtesting the VaR estimate focuses on the economic accuracy of the estimate. In order to do this, two techniques are used, the first of which is generating a loss function (LF), where this is in line with previous studies, for example, those by Lopez (1998) and Sarma, *et al.* (2003). The second technique is to perform a Modified Diebold-Mariano (MDM) test, proposed by Harvey, *et al.* (1997). Two main justifications for looking at economic differences between models exist. The first of these is that often more than one model for the VaR will pass the statistical tests described above; therefore it is beneficial to be able to differentiate between them on a different basis. The second is that one of the most important criticisms of the VaR methodologies is that for these, one can only see that a violation has occurred, but one is not certain as to the magnitude of these violations, and therefore the magnitude of the potential loss. The loss function enables one to identify the size of the potential losses, and therefore address the described criticism of the

VaR methodologies. Having calculated the loss function, the MDM test enables one to test whether there is a statistically significant difference between models, and thus choose the model that best minimises the risk exposure of the interested party.

In order to generate the loss function, one first needs to calculate the Expected Shortfall (ES), which is also known as the Conditional VaR (CVaR), where the ES is defined as the average loss incurred for the violations of the VaR. This means that for a long position, the ES will be:

$$\mathrm{ES}(\alpha) = E\left[r_i \middle| r_i \leq \mathrm{VaR}(\alpha)\right] \qquad (8.19)$$

In contrast, for a short position, the ES will be:

$$\mathrm{ES}(\alpha) = E\left[r_i \middle| r_i \geq \mathrm{VaR}(\alpha)\right] \qquad (8.20)$$

In Expressions (8.19) and (8.20), above, r_i denotes the actual return at forecast point *i* and α denotes the desired coverage rate. Having calculated the ES, one can then construct the loss function for the VaR model as follows:

$$LF_q = \frac{1}{M}\sum_{i=1}^{M}\left[r_i - \mathrm{ES}_q(\alpha)\right]^2 \qquad (8.21)$$

In Expression (8.21), above, LF_q and ES_q denotes the loss function and ES for the *q*th VaR model, respectively, r_i denotes the actual return at forecast point *i*, *M* denotes the forecast horizon, and α denotes the desired coverage rate. One should note the following regarding the loss function in Expression (8.21), where for a long position:

$$r_i - \mathrm{ES}_q(\alpha) = \begin{cases} 0, & \text{if } \mathrm{ES}_q(\alpha) \leq r_i \\ r_i - \mathrm{ES}_q(\alpha) & \text{if } r_i < \mathrm{ES}_q(\alpha) \end{cases} \qquad (8.22)$$

While for a short position:

$$r_i - \mathrm{ES}_q(\alpha) = \begin{cases} 0, & \text{if } \mathrm{ES}_q(\alpha) \geq r_i \\ r_i - \mathrm{ES}_q(\alpha) & \text{if } r_i > \mathrm{ES}_q(\alpha) \end{cases} \qquad (8.23)$$

In Expressions (8.22) and (8.23), above, ES_Q denotes the ES for the *q*th VaR model, respectively, r_i denotes the actual return at forecast point *i*, and α denotes the desired coverage rate. One can thus state that the proposed loss function will be equal to the semi-variance of the variable. This means that the loss function will take into account the magnitude of any returns that have exceeded the VaR and are greater than the calculated ES. One will then choose the best model, among the different options proposed above, as that which minimises the loss function, having passed all three of the statistical accuracy tests discussed above.

Having generated the loss function for each model and then selecting the best model based on this function, this process is double-checked using the MDM test outlined in Harvey, *et al.* (1997). This test is an improvement on the previous Diebold-Mariano test, proposed by Diebold and Mariano (1995), in that the latter test has a tendency to commit too many type 1 errors, i.e. reject the null hypothesis when it is in fact true. This test compares forecasts from VaR models by evaluating a second respective loss function, denoted $g\left(e_Q\right)$, where these loss functions are calculated as follows:

$$g\left(e_q\right) = r_{q;i} - \mathrm{ES}_{q;i}\left(\alpha\right) \qquad (8.24)$$

In Expression (8.24), $i = 1;\ldots;M$, where i and M denote the respective forecast point and horizon, respectively, q denotes the qth VaR model, and α denotes the desired coverage rate. Following this, the null hypothesis of equal accuracy in the forecasts of two competing models, i.e. that $E\left(d_i\right) = 0$, is tested, where:

$$d_i = g\left(e_{1;i}\right) - g\left(e_{2;i}\right) \qquad (8.25)$$

In Expression (8.25), above, i denotes the respective forecast point, and $g\left(e_{1;i}\right)$ and $g\left(e_{2;i}\right)$ denote the loss functions for the first and second model, respectively. One must note that in order to perform the test one must first calculate some descriptive statistics for the deviations, where the average can be calculated as follows:

$$\bar{d} = \frac{1}{M}\sum_{i=1}^{M} d_i \qquad (8.26)$$

Therefore, the standard deviation of the deviations will be:

$$\mathrm{Var}\left(d\right) = \frac{1}{M^2}\sum_{i=1}^{M}\left(d_i - \bar{d}\right)^2 \qquad (8.27)$$

In Expressions (8.26) and (8.27), above, i and M denote the respective forecast point and horizon, respectively.

The MDM test will then be given by:

$$\text{MDM} = \left[\frac{M-1}{M}\right]^{1/2} \frac{\bar{d}}{\sqrt{\text{Var}\left(\bar{d}\right)}} \sim t_{(M-1)} \qquad (8.28)$$

One should note that the benchmark model for comparison, as discussed above, will be the model that minimises the loss function, having passed all three of the statistical accuracy tests discussed above.

Having outlined the various methodologies to be used in this chapter in this section, the following section briefly describes the data used in this chapter.

8.3 Description of Data

Having established the methodology to be used in this empirical analysis in the previous section, this section provides a very brief summary of the relevant descriptive statistics for the respective data series, where the complete analysis can be found in Chapter 4 of the thesis. The data set used in this chapter comprises five data series of spot freight rate returns for five different vessel types across the tanker and dry-bulk sectors, namely VLCC, Suezmax (SZMX) and Aframax (AFMX) tankers, and Capesize (CPSZ) and Panamax (PNMX) dry-bulk vessels. The total sample extends from 13 January 1989 to 26 June 2009, thus comprising 1,068 observations, where all data was collected from Clarksons Shipping Intelligence Network. This being said, to enable ex-post forecasts to be made, the sample was then sub-divided into an in-sample period, extending from 13 January 1989 to 26 September 2003, thus comprising 768 observations, and an out-of-sample period, extending from 3 October 2003 to 26 June 2009, thus comprising 299 observations.

Calculated descriptive statistics indicate that there are contrasting size effects in the tanker and dry-bulk sectors, where, in the tanker sector, larger vessels are found to exhibit lower returns, which may be as a result of the reduced flexibility of these vessels with respect to smaller vessels. The reverse is found in the dry-bulk sector, however this is attributed to the increased flexibility of dry-bulk vessels, which may enable the vessels to take advantage of economies of scale. Returns, across all vessel types, are found to exhibit excess kurtosis and significant positive skewness, with the exception of the Panamax data series, where returns are negatively skewed. These findings are supported Jarque-Bera statistics, where the null hypothesis of normally distributed returns is rejected in all cases.

One should note that these models were also run on the natural logarithms of the respective data series in order to take account of the fact that freight rates can never be negative. One should note that the results from these estimations did not differ significant from those presented here[26].

Having outlined the characteristics of the data series, the next section focuses on the estimation of the respective models, giving insight into the applications and implications of these results.

8.4 Empirical Results

Having outlined the methodology and data to be used in this chapter in previous sections, this section outlines the results from the implementation of these and therefore gives an indication of the future direction of the risk exposure encountered by participants in the tanker and dry-bulk shipping markets. These results are divided into the process of generating the VaR estimate and backtesting these results from first the ship-owner's point of view and then from the viewpoint of the charterer.

8.4.1 The Ship-Owner's Point of View

As mentioned above, this sub-section examines the risk exposure of ship-owners in the tanker and dry-bulk market. This is done by examining the 1% and 5% VaRs on a long position, with respect to the VLCC, Suezmax, Aframax, Capesize and Panamax data series, respectively. As briefly mentioned above, one should not interpret these measures as the potential loss on a portfolio, as no portfolio has been constructed here, but instead should consider this to be the potential loss in profits incurred by the ship-owner should market conditions move against them.

[26] The results from these estimations are not presented here due to space constraints and are available from the author upon request.

This research started this model evaluation process by examining the VaR results for a long position on the VLCC data series, the results of which are summarised in the various panels in Table 8.1. Examining the 1% VaR results, presented in Panel A, one is only interested in those models which passed the statistical backtesting process, i.e. the RM, GARCH and GARCHSK models. This would imply that that the VaR estimates for these models neither over or underestimates the true VaR and these estimates are independently distributed. Examining the respective tests of economic accuracy, i.e. the expected shortfall and loss functions, one is indifferent between the models, while the results of the respective MDM tests, outlined in Panel A of Table 8.11 in Appendix 8.A, indicate that one cannot reject the null hypothesis that there is effectively no difference between the respective loss functions. One can thus conclude that in terms of both the statistical and economic accuracy of the VaR figures, one is indifferent between the RM, GARCH and GARCHSK. In order to differentiate between the models, the respective hit ratios and RMSEs are examined, where one is found to be indifferent between the models in terms of their hit ratios; however, on the basis of the lower RMSE, the conclusion is reached that the GARCHSK model performs best in this case.

The picture changes drastically when looking at the results for the 5% VaR results, presented in Panel B of Table 8.1, where none of the VaR estimates for any of the models passes the statistical backtesting process, where although the VaR estimates for the HS (200), FHS (200), FHS (400) and GARCHSK models were found not to significantly under or overestimate the "true" VaR, none of the VaR estimates for any of the models were found to be independently distributed. For this reason, one cannot draw any conclusion as to which model provides the most accurate estimate of the risk exposure incurred by the ship-owner. This means that ship-owners operating VLCC vessels would have to rely on the more stringent 1% VaR measure should they wish to calculate the respective risk exposure in this context.

The analysis of ship-owners risk exposures continues to examine the VaR estimates for a long position on the Suezmax data series, the results of which are presented in the various panels in Table 8.2. Beginning with the 1% VaR results, presented in Panel A, one should first note that the RM, IGARCH and GARCHSK models never violated their respective VaR estimates and hence were excluded on the basis that

Table 8.1 – Value-at-Risk Results on a Long Position for the VLCC Data Series

Panel A - 1% VaR Results on a Long Position for the VLCC Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.315	1.672%	-----	-----	0.287	0.003	0.008	-0.374	0.552
HS (400)	-0.298	1.672%	-----	-----	0.287	0.003	0.008	-0.375	0.542
FHS (200)	-0.392	1.672%	-----	-----	0.287	0.003	0.008	-0.374	0.552
FHS (400)	-0.439	0.669%	-----	-----	0.540	0.001	0.004	-0.450	0.084
RiskMetrics	-0.385	0.669%	77.181%	0.169	0.540	0.841	0.813	-0.433	0.149
GARCH	-0.386	0.669%	78.523%	0.168	0.540	0.841	0.813	-0.433	0.149
IGARCH	-0.385	1.003%	78.523%	0.167	0.995	0.002	0.008	-0.422	0.202
FIGARCH	-0.372	1.003%	68.456%	0.167	0.995	0.002	0.008	-0.422	0.202
GARCHSK	-0.447	0.669%	73.154%	0.162	0.540	0.841	0.813	-0.433	0.149

Panel B - 5% VaR Results on a Long Position for the VLCC Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.200	6.020%	-----	-----	0.432	0.017	0.043	-0.276	2.543
HS (400)	-0.171	8.027%	-----	-----	0.027	0.018	0.005	-0.252	3.576
FHS (200)	-0.225	5.017%	-----	-----	0.989	0.006	0.023	-0.283	2.306
FHS (400)	-0.234	4.013%	-----	-----	0.418	0.008	0.021	-0.310	1.584
RiskMetrics	-0.272	2.341%	77.181%	0.169	0.019	0.005	0.001	-0.303	1.757
GARCH	-0.273	2.341%	78.523%	0.168	0.019	0.005	0.001	-0.303	1.757
IGARCH	-0.272	2.341%	78.523%	0.167	0.019	0.005	0.001	-0.303	1.757
FIGARCH	-0.263	2.341%	68.456%	0.167	0.019	0.005	0.001	-0.308	1.633
GARCHSK	-0.226	3.679%	73.154%	0.162	0.272	0.006	0.013	-0.277	2.519

Note 1: VLCC denotes the weekly spot freight rates for a 270,000 DWT VLCC tanker carrying crude oil from Ras Tanura (Saudi Arabia) to Rotterdam (Netherlands).
SZMX denotes the weekly spot freight rates for a 130,000 DWT Suezmax tanker carrying crude oil from Bonny (Nigeria) to off the coast of Philadelphia (USA).
AFMX denotes the weekly spot freight rates for an 80,000 DWT Aframax tanker carrying crude oil from Sullom Voe (UK) to Bayway (USA).
CPSZ denotes the weekly spot freight rates for a 145,000 DWT Capesize bulk-carrier carrying iron ore from Tubarao (Brazil) to Rotterdam (Netherlands).
PNMX denotes the weekly spot freight rates for a 55,000 DWT Panamax bulk-carrier carrying grain from the Hampton Roads (USA) to Antwerp-Rotterdam-Amsterdam (Benelux).

Note 3: The sample period for the data used for this table extends from 3 October 2003 to 26 June 2009, with a total of 299 observations.

Note 4: The data used for this table is all sourced from the Clarkson Shipping Intelligence Network (www.clarskons.net).

Note 5: Ave VaR, Hit Ratio, % Over and % Under denote the average VaR, percentage of violations of the VaR, percentage of over-predictions and percentage of under-predictions, respectively.

Note 6: The fiigures in green and red denote where one can and cannot reject the null hypothesis for the test of statistical accuracy, respectively.

Note 7: HS (200) and HS (400) denote the Historical Simulation results for the 200 and 400 week horizons, respectively.

Note 8: FHS (200) and FHS (400) denote the Filtered Historical Simulation results for the 200 and 400 week horizons, respectively.

Note 9: For ease of reference, the loss function figures have each been multiplied by 10^4, respectively.

these set too high a reserve for potential losses, where this would mean that too much money was set aside that could have been used for alternative investment opportunities. Of the remaining models, the VaR estimates for the HS (200),

Table 8.2 – Value-at-Risk Results on a Long Position for the SZMX Data Series

Panel A - 1% VaR Results on a Long Position for the SZMX Data Series									
	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.262	2.007%	-----	-----	0.124	0.592	0.265	-0.308	0.203
HS (400)	-0.245	3.010%	-----	-----	0.005	0.430	0.014	-0.279	0.516
FHS (200)	-0.307	1.003%	-----	-----	0.995	0.776	0.960	-0.339	0.054
FHS (400)	-0.330	0.334%	-----	-----	0.179	0.908	0.403	-0.379	0.000
RiskMetrics	-0.400	0.000%	71.812%	0.162	N / A	N / A	N / A	N / A	N / A
GARCH	-0.299	2.676%	60.403%	0.134	0.016	0.481	0.043	-0.282	0.484
IGARCH	-0.376	0.000%	69.128%	0.151	N / A	N / A	N / A	N / A	N / A
FIGARCH	-0.357	0.334%	62.416%	0.150	0.179	0.908	0.403	-0.317	0.147
GARCHSK	-0.482	0.000%	68.792%	0.156	N / A	N / A	N / A	N / A	N / A

Panel B - 5% VaR Results on a Long Position for the SZMX Data Series									
	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.211	7.023%	-----	-----	0.129	0.206	0.142	-0.256	0.925
HS (400)	-0.187	10.702%	-----	-----	0.000	0.009	0.000	-0.232	1.673
FHS (200)	-0.245	4.013%	-----	-----	0.418	0.297	0.418	-0.275	0.576
FHS (400)	-0.254	2.676%	-----	-----	0.044	0.481	0.102	-0.283	0.470
RiskMetrics	-0.283	1.338%	71.812%	0.162	0.001	0.712	0.003	-0.300	0.274
GARCH	-0.212	7.692%	60.403%	0.134	0.047	0.045	0.019	-0.231	1.707
IGARCH	-0.266	1.338%	69.128%	0.151	0.001	0.712	0.003	-0.300	0.274
FIGARCH	-0.252	3.344%	62.416%	0.150	0.163	0.382	0.258	-0.268	0.698
GARCHSK	-0.247	3.679%	68.792%	0.156	0.272	0.384	0.375	-0.264	0.769

Note 1: See the respective notes from Table 8.1

FHS (200), FHS (400) and FIGARCH models were found to pass the statistical backtesting process, where the FHS (400) was found to provide the lowest potential losses, as measured by the loss function. This being said, subsequent MDM tests, the results of which are presented in Panel A of Table 8.12 in Appendix 8.A indicated that one is statistically indifferent between the respective loss functions for each of these models. As one cannot calculate the respective RMSE or CDIR ratios for the non-parametric historical and filtered historical simulation methodologies, the FHS (200) model was selected as providing the best estimate of ship-owners' risk exposure on the basis that its hit ratio was the closest to the desired threshold of 1%.

Changing focus to the 5% VaR, the results for which are displayed in Panel B of Table 8.2, the VaR estimates for the HS (200), FHS (200), FIGARCH and GARCHSK models were found to pass the statistical backtesting process, where, of these, the FHS (200) model was found to provide the smallest loss function. In order to provide a more sound basis on to which to draw a conclusion as to the best model, the results for the further MDM tests, which are summarised in Panel B of Table 8.12

in Appendix 8.A, indicate that the loss function for the FHS (200) model is significantly different from those for the other models, hence the conclusion is reached that the FHS (200) model provides the best indication as to the risk exposure faced by ship-owners in the freight market.

To conclude the analysis of a ship-owner's risk exposure in the tanker sector, the VaR estimates for a long position on the Aframax data series, the results for which are presented in Panel A of Table 8.3, are analysed. Before proceeding any further, these results indicate that the VaR estimate for the FHS (400) model was never exceeded, leading to its exclusion on the basis that the resultant capital held in reserved could be better utilised elsewhere. Having established this, the results indicate that all remaining models, with the exception of the GARCH model, passed the statistical backtesting process, where of these, the GARCHSK model was found to have the best loss function. Subsequent MDM tests on these loss functions, the results for which are displayed in Panel A of Table 8.13 in Appendix 8.A, indicated that one cannot reject the null hypothesis that there is effectively no difference between the loss functions for the HS (200), HS (400), FHS (200), IGARCH and GARCHSK models, while the loss function for the GARCHSK model was found to be significantly different from those for the RM and FIGARCH models, hence the latter two models are excluded from any further analysis. As one cannot calculate the respective RMSEs for the HS (200), HS (400) and FHS (200) models, the conclusion was reached that the IGARCH model provided the best evaluation of a ship-owner's risk exposure as the respective hit ratio was closest to the desired threshold of 1%.

Looking at the results for the 5% VaR, presented in Panel B of Table 8.3, only the HS (200), FHS (200) and FIGARCH models were found to have passed the statistical backtesting process, while, of these, the FHS (200) model was found to minimise the respective loss function. In order to differentiate between these models, MDM tests were performed on each data series, where these results are outlined in Panel B of Table 8.13 in Appendix 8.A. Based on these results, the conclusion is reached that there is a significant difference between the respective loss functions for the HS (200), FHS (200) and FIGARCH models, hence the FHS (200) model is found to provide the best evaluation of the potential risk exposure in the market.

Table 8.3 – Value-at-Risk Results on a Long Position for the AFMX Data Series

Panel A - 1% VaR Results on a Long Position for the AFMX Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.301	1.672%	-----	-----	0.287	0.651	0.512	-0.330	0.282
HS (400)	-0.277	1.338%	-----	-----	0.577	0.712	0.800	-0.347	0.177
FHS (200)	-0.337	1.338%	-----	-----	0.577	0.712	0.800	-0.343	0.195
FHS (400)	-0.437	0.000%	-----	-----	N / A	N / A	N / A	N / A	N / A
RiskMetrics	-0.336	0.334%	23.154%	0.156	0.179	0.908	0.403	-0.278	0.975
GARCH	-0.236	2.676%	42.617%	0.122	0.016	0.481	0.043	-0.288	0.790
IGARCH	-0.308	1.003%	25.168%	0.141	0.995	0.776	0.960	-0.349	0.163
FIGARCH	-0.303	1.338%	38.591%	0.157	0.577	0.712	0.800	-0.299	0.634
GARCHSK	-0.378	0.334%	47.651%	0.146	0.179	0.908	0.403	-0.351	0.155

Panel B - 5% VaR Results on a Long Position for the AFMX Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.169	6.020%	-----	-----	0.432	0.118	0.217	-0.234	2.069
HS (400)	-0.158	7.358%	-----	-----	0.080	0.618	0.190	-0.221	2.513
FHS (200)	-0.190	4.348%	-----	-----	0.597	0.259	0.460	-0.256	1.449
FHS (400)	-0.240	2.676%	-----	-----	0.044	0.481	0.102	-0.289	0.774
RiskMetrics	-0.237	1.672%	23.154%	0.156	0.002	0.651	0.009	-0.314	0.437
GARCH	-0.167	7.692%	42.617%	0.122	0.047	0.045	0.019	-0.216	2.703
IGARCH	-0.218	1.672%	25.168%	0.141	0.002	0.651	0.009	-0.330	0.282
FIGARCH	-0.214	4.682%	38.591%	0.157	0.799	0.224	0.463	-0.243	1.786
GARCHSK	-0.186	7.358%	47.651%	0.146	0.080	0.618	0.190	-0.217	2.664

Note 1: See the respective notes from Table 8.1

To summarise the results for the tanker market, the 1% VaR indicated that for the VLCC data series, the GARCHSK model provided the best evaluation of a ship-owner's risk exposure in the freight market, whereas for the Suezmax and Aframax data series this was provided by the FHS (200) and IGARCH models, respectively. Therefore, it appears that at this threshold in the tanker, parametric models generally outperform non-parametric models. In contradistinction, although no conclusion could be reached as to which model provided the best estimate of a ship-owner's risk exposure for the VLCC data series, as none of the model passed the backtesting process, the results for the Suezmax and Aframax data series were unanimous in their finding that the FHS (200) model outperformed all others, thus indicating that at this threshold, non-parametric specifications outperform parametric specifications.

Changing focus to the dry-bulk sector, the VaR results on a long position with respect to the Capesize data series are discussed, where these can be found in summarised form in Table 8.4. Panel A of the table presents the results for the 1% threshold,

Table 8.4 – Value-at-Risk Results on a Long Position for the CPSZ Data Series

Panel A - 1% VaR Results on a Long Position for the CPSZ Data Series									
	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.141	5.351%	-----	-----	0.000	0.238	0.000	-0.169	2.153
HS (400)	-0.129	5.351%	-----	-----	0.000	0.238	0.000	-0.176	1.940
FHS (200)	-0.220	0.669%	-----	-----	0.540	0.841	0.813	-0.342	0.018
FHS (400)	-0.284	0.000%	-----	-----	-----	-----	-----	-----	-----
RiskMetrics	-0.185	1.338%	28.523%	0.077	0.577	0.712	0.800	-0.244	0.677
GARCH	-0.126	5.017%	37.248%	0.053	0.000	0.187	0.000	-0.170	2.131
IGARCH	-0.163	2.676%	32.886%	0.065	0.016	0.481	0.043	-0.210	1.197
FIGARCH	-0.170	2.676%	35.570%	0.074	0.016	0.481	0.043	-0.205	1.283
GARCHSK	-0.224	1.338%	29.866%	0.073	0.577	0.712	0.800	-0.207	1.249

Panel B - 5% VaR Results on a Long Position for the CPSZ Data Series									
	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.091	11.371%	-----	-----	0.000	0.000	0.000	-0.141	3.310
HS (400)	-0.080	13.043%	-----	-----	0.000	0.000	0.000	-0.137	3.565
FHS (200)	-0.148	2.676%	-----	-----	0.044	0.481	0.102	-0.219	1.037
FHS (400)	-0.192	0.669%	-----	-----	0.000	0.841	0.000	-0.342	0.018
RiskMetrics	-0.131	5.017%	28.523%	0.077	0.989	0.031	0.097	-0.174	2.021
GARCH	-0.089	10.702%	37.248%	0.053	0.000	0.012	0.000	-0.138	3.495
IGARCH	-0.115	6.689%	32.886%	0.065	0.202	0.006	0.010	-0.173	2.039
FIGARCH	-0.120	6.355%	35.570%	0.074	0.301	0.123	0.178	-0.153	2.782
GARCHSK	-0.111	6.689%	29.866%	0.073	0.202	0.036	0.049	-0.162	2.405

Note 1: See the respective notes from Table 8.1

where the results for the FHS (400) model are ignored on the basis that there are no violations; hence this model may overestimate the required reserves, where this capital may be better utilised for other investments. Of the remaining models, only the FHS (200), RM and GARCHSK models pass the statistical backtesting process, where the results of further MDM tests, presented in Panel A of Table 8.14 in Appendix 8.A, indicated that one cannot reject the null hypothesis that there is effectively no difference between the respective loss functions. Therefore, in order to differentiate between the models, one should look at the respective hit ratios, where the RM and GARCHSK models were found to have the closest hit ratios to the desired threshold of 1%, however the GARCHSK model was found to have the lower RMSE. On this basis, one can conclude that the GARCHSK provides the best approximation of the risk exposure faced by a ship-owner in this context.

Examining the results for the 5% threshold, presented in Panel B of Table 8.4, the only model that passed the statistical backtesting process is the FIGARCH model;

hence one can conclude that this model provides the best evaluation of the risk exposure.

To conclude the analysis of the ship-owner's risk exposure, this research discusses the VaR estimates for a long position on the Panamax data series, the results of which are presented in Table 8.5. The results for the 1% threshold, outlined in Panel A, indicated that only the FHS (200) and RM models passed all three of the tests for statistical accuracy, where the RM model was found to exhibit the lower loss function.

In order to provide a definitive conclusion as to which model provides the best evaluation of a ship-owners respective risk exposure, a further MDM test was performed, the results of which is presented in Panel A of Table 8.15 in Appendix 8.A, where these indicated that there was a statistically significant difference between the two respective loss functions. On this basis, the conclusion was drawn that the RM model provided the best forecasts in this instance. Changing focus to the 5% threshold, the results of which are presented in Panel B of Table 8.5, only the FHS (200), RM and FIGARCH models were found to have passed the statistical backtesting process. Of these models, the RM model was found to minimise the loss function, however, to provide a more definitive answer as to the preferred model, further MDM tests, the results of which are presented in Panel B of Table 8.15 in Appendix 8.A. The results of these further tests indicated that there was a significant difference between the loss functions of the RM and FHS (200) and FIGARCH models, respectively. On this basis, the conclusion was reached that the RM model provided the best evaluation of the potential risk exposure in this context.

To summarise the results for the dry-bulk sector, the results for the Capesize data series indicate that the GARCHSK and FIGARCH models provide the best evaluation of the potential risk exposure at the 1% and 5% thresholds respectively. In contrast, the results for the Panamax data series indicated that for both thresholds the RM model was preferred. There therefore appears to be a size effect here, where for larger vessels in the dry-bulk sector parametric models are found to outperform non-parametric models, and vice versa.

Table 8.5 – Value-at-Risk Results on a Long Position for the PNMX Data Series

Panel A - 1% VaR Results on a Long Position for the PNMX Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.154	3.679%	-----	-----	0.000	0.384	0.001	-0.180	0.647
HS (400)	-0.143	4.682%	-----	-----	0.000	0.605	0.000	-0.180	0.658
FHS (200)	-0.260	0.669%	-----	-----	0.540	0.841	0.813	-0.191	0.453
FHS (400)	-0.309	0.000%	-----	-----	-----	-----	-----	-----	-----
RiskMetrics	-0.179	1.672%	31.879%	0.073	0.287	0.651	0.512	-0.227	0.066
GARCH	-0.139	4.348%	34.228%	0.056	0.000	0.533	0.000	-0.180	0.657
IGARCH	-0.144	4.682%	36.577%	0.057	0.000	0.020	0.000	-0.182	0.602
FIGARCH	-0.174	3.010%	34.564%	0.072	0.005	0.248	0.010	-0.177	0.712
GARCHSK	-0.129	9.030%	35.906%	0.044	0.000	0.000	0.000	-0.090	5.771

Panel B - 5% VaR Results on a Long Position for the PNMX Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	-0.100	10.368%	-----	-----	0.000	0.002	0.000	-0.143	1.785
HS (400)	-0.083	13.043%	-----	-----	0.000	0.005	0.000	-0.132	2.288
FHS (200)	-0.152	4.682%	-----	-----	0.799	0.605	0.847	-0.166	0.992
FHS (400)	-0.178	2.341%	-----	-----	0.019	0.535	0.053	-0.164	1.020
RiskMetrics	-0.126	5.017%	31.879%	0.073	0.989	0.668	0.912	-0.172	0.818
GARCH	-0.098	8.361%	34.228%	0.056	0.015	0.171	0.020	-0.148	1.587
IGARCH	-0.102	8.361%	36.577%	0.057	0.015	0.011	0.002	-0.157	1.250
FIGARCH	-0.123	5.017%	34.564%	0.072	0.989	0.668	0.912	-0.171	0.851
GARCHSK	-0.067	16.388%	35.906%	0.044	0.000	0.002	0.000	-0.115	3.359

Note 1: See the respective notes from Table 8.1

Having evaluated the ship-owner's side of the picture, the following sub-section extends the analysis by examining the charterer's point of view by estimating the VaR on a short position, where, to the best of the author's knowledge, this is done for the first time in the shipping literature.

8.4.2 The Charterer's Point of View

As mentioned above, this sub-section examines the risk exposure of charterer's in the tanker and dry-bulk sectors. This is done by examining the 1% and 5% VaRs on a short position, with respect to the VLCC, Suezmax, Aframax, Capesize and Panamax data series, respectively. As briefly mentioned above, one should not interpret these measures as the potential loss on a portfolio, as no portfolio has been constructed here, but instead should consider this to be the potential increase in transportation costs incurred by the charterer should market conditions move against them.

This analysis begins by examining the VaR estimates for a short position on the VLCC data series, where the results for these are summarised in Table 8.6. Panel A of the table examines the 1% VaR, where the VaR estimates for the HS (200), HS (400), FHS (200), FHS (400) and GARCHSK models are found to pass all three of the tests of statistical accuracy, where the HS (200) and FHS (200) models were found to provide the best lost function. In order to further differentiate between models, MDM tests were performed on the respective models, where these results are outlined in Panel A of Table 8.16 in Appendix 8.A, the results of which indicate that, with the exception of the GARCHSK model, one is unable to reject the null hypothesis that there is effectively no difference between the respective loss functions. Having established this, the conclusion is drawn that, as this model provides the hit ratio closest to the desired threshold of 1%, the FHS (400) model provides the best evaluation of the potential risk exposure faced by charterers.

Changing the threshold, the results for the 5% VaR, which are presented in Panel B of Table 8.6, indicate that the HS (200), HS (400), RM, GARCH and IGARCH model all pass the statistical backtesting process, while the HS (200) model minimised the respective loss function. Given this, further MDM tests were performed, the results of which are outlined in Panel B of Table 8.16 in Appendix 8.A, in order to determine the preferred model to evaluate the risk exposure faced by charterers. These results indicated that the loss function for the HS (200) was significantly different from those for the other four models; hence the conclusion was reached that the HS (200) model provided the best evaluation of the risk exposure at this threshold in this context.

The analysis continues by examining the VaR estimates for a short position on the Suezmax data series, the results for which are summarised in Table 8.7. Beginning with the results for the 1% threshold, which are outlined in Panel A, only the FHS (200), FHS (400) and GARCHSK pass the statistical backtesting process, where the FHS (400) model is found to have the lowest loss function. In order to further differentiate between the models, MDM tests were performed on these models, the results of which is presented in Panel A of Table 8.17 in Appendix 8.A, where these indicated that, while one could not reject the null hypothesis that there is effectively no difference between the respective loss functions for the FHS (400) and GARCHSK models, there was a statistically significant difference between the loss functions for

Table 8.6 – Value-at-Risk Results on a Short Position for the VLCC Data Series

Panel A - 1% VaR Results on a Short Position for the VLCC Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.453	1.672%	-----	-----	0.287	0.651	0.512	0.729	4.026
HS (400)	0.434	2.007%	-----	-----	0.124	0.592	0.265	0.682	5.790
FHS (200)	0.552	1.672%	-----	-----	0.287	0.651	0.512	0.729	4.026
FHS (400)	0.592	1.003%	-----	-----	0.995	0.776	0.960	0.891	0.727
RiskMetrics	0.385	3.679%	77.181%	0.169	0.000	0.052	0.000	0.502	16.664
GARCH	0.386	3.344%	78.523%	0.168	0.001	0.034	0.001	0.530	14.484
IGARCH	0.385	3.344%	78.523%	0.167	0.001	0.034	0.001	0.530	14.484
FIGARCH	0.372	4.348%	68.456%	0.167	0.000	0.106	0.000	0.465	19.881
GARCHSK	0.515	2.007%	73.154%	0.162	0.124	0.592	0.265	0.645	7.487

Panel B - 5% VaR Results on a Short Position for the VLCC Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.306	4.348%	-----	-----	0.597	0.106	0.235	0.497	17.055
HS (400)	0.261	6.355%	-----	-----	0.301	0.426	0.427	0.429	23.388
FHS (200)	0.364	3.010%	-----	-----	0.089	0.021	0.016	0.562	12.260
FHS (400)	0.380	3.010%	-----	-----	0.089	0.248	0.121	0.587	10.690
RiskMetrics	0.272	7.023%	77.181%	0.169	0.129	0.560	0.267	0.405	25.995
GARCH	0.273	6.689%	78.523%	0.168	0.202	0.494	0.350	0.416	24.738
IGARCH	0.272	6.689%	78.523%	0.167	0.202	0.494	0.350	0.416	24.738
FIGARCH	0.263	8.361%	68.456%	0.167	0.015	0.049	0.007	0.373	29.840
GARCHSK	0.252	8.696%	73.154%	0.162	0.008	0.216	0.013	0.365	30.884

Note 1: See the respective notes from Table 8.1

the FHS (400) and FHS (200) models. This being said, having examined the respective hit ratios for the remaining models, and given the fact that the hit ratio for the FHS (400) models was closest to the desired threshold of 1%, the conclusion was drawn that the FHS (400) model provided the best evaluation of the risk exposure incurred by charterers in this context.

Looking at the results for the 5% VaRs, which are summarised in Panel B of Table 8.7, only the FHS (200) and FHS (400) models were found to have passed all three of the tests of statistical accuracy, while the FHS (400) minimised the respective loss function. To further distinguish between these two models, a further MDM tests was performed, the result of which, where this is summarised in Panel B of Table 8.17 in Appendix 8.A, indicated that there is a statistically significant difference between the loss functions of these two models, thus leading to the conclusion that the FHS (400) model provided the best estimate of the risk exposure of charterers in this context.

Table 8.7 – Value-at-Risk Results on a Short Position for the SZMX Data Series

Panel A - 1% VaR Results on a Short Position for the SZMX Data Series									
	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.460	2.676%	-----	-----	0.016	0.481	0.043	0.606	3.609
HS (400)	0.395	2.676%	-----	-----	0.016	0.481	0.043	0.606	3.609
FHS (200)	0.524	2.007%	-----	-----	0.124	0.592	0.265	0.654	2.126
FHS (400)	0.604	1.338%	-----	-----	0.577	0.712	0.800	0.745	0.745
RiskMetrics	0.400	3.344%	71.812%	0.162	0.001	0.382	0.004	0.560	5.591
GARCH	0.299	8.696%	60.403%	0.134	0.000	0.023	0.000	0.392	18.697
IGARCH	0.376	3.010%	69.128%	0.151	0.005	0.430	0.014	0.584	4.477
FIGARCH	0.357	5.351%	62.416%	0.150	0.000	0.165	0.000	0.466	11.625
GARCHSK	0.561	1.672%	68.792%	0.156	0.287	0.651	0.512	0.702	1.249

Panel B - 5% VaR Results on a Short Position for the SZMX Data Series									
	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.278	8.696%	-----	-----	0.008	0.638	0.026	0.400	17.783
HS (400)	0.250	10.368%	-----	-----	0.000	0.508	0.001	0.376	20.697
FHS (200)	0.339	5.017%	-----	-----	0.989	0.668	0.912	0.484	10.247
FHS (400)	0.356	4.013%	-----	-----	0.418	0.297	0.418	0.527	7.378
RiskMetrics	0.283	8.361%	71.812%	0.162	0.015	0.171	0.020	0.400	17.777
GARCH	0.212	13.378%	60.403%	0.134	0.000	0.375	0.000	0.325	28.181
IGARCH	0.266	9.699%	69.128%	0.151	0.001	0.161	0.002	0.382	19.935
FIGARCH	0.252	9.699%	62.416%	0.150	0.001	0.154	0.001	0.375	20.800
GARCHSK	0.277	8.361%	68.792%	0.156	0.015	0.171	0.020	0.399	17.907

Note 1: See the respective notes from Table 8.1

To conclude the analysis of the risk exposure faced by charterers in the tanker market, the VaR estimates for a short position on the Aframax data series are analysed, where the results of this analysis can be found in Table 8.8. Looking at the results for the 1% threshold, outlined in Panel A of the table, the HS (200), HS (400), FHS (200), FHS (400) and GARCHSK models are all found to have passed the statistical backtesting processes, where the FHS (400) model was found to minimise the respective loss functions. Following this, the results of subsequent MDM tests, where these are summarised in Panel A of Table 8.18 in Appendix 8.A, indicate that one cannot reject the null hypothesis that there is effectively no difference between the respective loss functions for these models. In order to distinguish the best model from these, the results for the respective hit ratios indicate that the FHS (400) provided the hit ratio closest to the desired threshold of 1% and therefore is preferred over other models in this context.

Changing focus to the 5% threshold, VaR estimation results, presented in Panel B of Table 8.8, indicated that the HS (200), FHS (200), FHS (400), RM, FIGARCH and

Table 8.8 – Value-at-Risk Results on a Short Position for the AFMX Data Series

Panel A - 1% VaR Results on a Short Position for the AFMX Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.380	2.007%	-----	-----	0.124	0.592	0.265	0.615	9.621
HS (400)	0.395	2.007%	-----	-----	0.124	0.592	0.265	0.633	8.743
FHS (200)	0.435	1.672%	-----	-----	0.287	0.651	0.512	0.680	6.676
FHS (400)	0.655	1.003%	-----	-----	0.995	0.776	0.960	0.820	2.225
RiskMetrics	0.336	2.676%	23.154%	0.156	0.016	0.189	0.023	0.544	13.519
GARCH	0.236	4.682%	42.617%	0.122	0.000	0.224	0.000	0.444	20.234
IGARCH	0.308	3.010%	25.168%	0.141	0.005	0.430	0.014	0.543	13.547
FIGARCH	0.303	3.679%	38.591%	0.157	0.000	0.384	0.001	0.490	16.924
GARCHSK	0.418	1.672%	47.651%	0.146	0.287	0.651	0.512	0.680	6.676

Panel B - 5% VaR Results on a Short Position for the AFMX Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.194	6.689%	-----	-----	0.202	0.161	0.166	0.381	25.961
HS (400)	0.183	7.358%	-----	-----	0.080	0.258	0.113	0.366	27.559
FHS (200)	0.226	6.020%	-----	-----	0.432	0.359	0.483	0.402	23.900
FHS (400)	0.316	4.682%	-----	-----	0.799	0.605	0.847	0.447	20.015
RiskMetrics	0.237	5.351%	23.154%	0.156	0.783	0.238	0.481	0.414	22.770
GARCH	0.167	10.368%	42.617%	0.122	0.000	0.477	0.001	0.306	35.313
IGARCH	0.218	5.686%	25.168%	0.141	0.594	0.066	0.160	0.410	23.081
FIGARCH	0.214	7.023%	38.591%	0.157	0.129	0.206	0.142	0.355	28.778
GARCHSK	0.201	6.355%	47.651%	0.146	0.301	0.675	0.537	0.373	26.804

Note 1: See the respective notes from Table 8.1

GARCHSK models all passed all three of the tests of statistical accuracy. This being said, the results of further MDM tests, where these are summarised in Panel B of Table 8.18 in Appendix 8.A, led to the conclusion that the loss function for the FHS (400) model is significantly different from all the other respective models; hence one can conclude that the FHS (400) model provides the best evaluation of the risk exposure faced by charterers in this context.

One can thus conclude that, with the exception of the 5% threshold for the VLCC data series, the FHS (400) model outperforms all other models in terms of calculating the risk exposure of charterers in the tanker market. This result supports the findings of Angelidis and Skiadopoulos (2008), who found that non-parametric models outperformed parametric models in the FFA market. Interestingly, the results from the ship-owner's perspective, i.e. the long position, correspond with these results at the 5% threshold; however, at the 1% threshold for the long position, these suggested that parametric models outperformed non-parametric models.

Table 8.9 – Value-at-Risk Results on a Short Position for the CPSZ Data Series

Panel A - 1% VaR Results on a Short Position for the CPSZ Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.209	3.679%	-----	-----	0.000	0.052	0.000	0.292	1.727
HS (400)	0.190	4.013%	-----	-----	0.000	0.076	0.000	0.282	2.008
FHS (200)	0.396	1.338%	-----	-----	0.577	0.712	0.800	0.368	0.446
FHS (400)	0.548	0.334%	-----	-----	0.179	0.908	0.403	0.255	2.939
RiskMetrics	0.185	4.682%	28.523%	0.077	0.000	0.143	0.000	0.258	2.819
GARCH	0.126	7.692%	37.248%	0.053	0.000	0.661	0.000	0.215	4.873
IGARCH	0.163	5.017%	32.886%	0.065	0.000	0.187	0.000	0.250	3.112
FIGARCH	0.170	5.017%	35.570%	0.074	0.000	0.187	0.000	0.247	3.277
GARCHSK	0.244	3.344%	29.866%	0.073	0.001	0.034	0.001	0.227	4.228

Panel B - 5% VaR Results on a Short Position for the CPSZ Data Series

	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.119	8.361%	-----	-----	0.015	0.150	0.018	0.214	4.951
HS (400)	0.104	10.033%	-----	-----	0.000	0.062	0.000	0.198	6.046
FHS (200)	0.214	3.344%	-----	-----	0.163	0.034	0.040	0.297	1.590
FHS (400)	0.270	1.672%	-----	-----	0.002	0.060	0.002	0.270	2.399
RiskMetrics	0.131	8.361%	28.523%	0.077	0.015	0.171	0.020	0.203	5.682
GARCH	0.089	12.709%	37.248%	0.053	0.000	0.209	0.000	0.170	8.237
IGARCH	0.115	8.696%	32.886%	0.065	0.008	0.017	0.002	0.204	5.613
FIGARCH	0.120	9.030%	35.570%	0.074	0.004	0.094	0.004	0.198	5.984
GARCHSK	0.119	9.699%	29.866%	0.073	0.001	0.003	0.000	0.185	7.018

Note 1: See the respective notes from Table 8.1

Changing focus to the dry-bulk market, this research now analyses the VaR estimates for a short position on the Capesize data series, the results for which are presented in Table 8.9. The results for the 1% VaR, outlined in Panel A of the table, suggest that only the FHS (200) and FHS (400) models passed the statistical backtesting process, where the FHS (200) model is found to have the lower loss functions. Results from the subsequent MDM test, summarised in Panel A of Table 8.19 in Appendix 8.A, show that there is a significant difference between the two respective loss functions, hence one can conclude that the FHS (200) model outperforms the FHS (400) model in terms of evaluating the risk exposure incurred by charterers in this context.

The picture changes drastically when looking at the results for the 5% VaR results, presented in Panel B of Table 8.9, where none of the VaR estimates for any of the models passed the statistical backtesting process. For this reason, one cannot draw any conclusion as to which model provides the most accurate estimate of the risk exposure faced by the charterer.

Table 8.10 – Value-at-Risk Results on a Short Position for PNMX Data Series

Panel A - 1% VaR Results on a Short Position for the PNMX Data Series									
	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.199	3.010%	-----	-----	0.005	0.003	0.000	0.216	0.496
HS (400)	0.173	3.344%	-----	-----	0.001	0.001	0.000	0.217	0.492
FHS (200)	0.352	1.003%	-----	-----	0.995	0.000	0.000	0.215	0.512
FHS (400)	0.363	1.338%	-----	-----	0.577	0.000	0.001	0.219	0.466
RiskMetrics	0.179	2.676%	31.879%	0.073	0.016	0.002	0.001	0.219	0.454
GARCH	0.139	5.686%	34.228%	0.056	0.000	0.016	0.000	0.181	1.276
IGARCH	0.144	4.013%	36.577%	0.057	0.000	0.002	0.000	0.209	0.613
FIGARCH	0.174	3.010%	34.564%	0.072	0.005	0.003	0.000	0.221	0.437
GARCHSK	0.160	5.351%	35.906%	0.044	0.000	0.000	0.000	0.190	1.027

Panel B - 5% VaR Results on a Short Position for the PNMX Data Series									
	Average VaR	Hit Ratio	CDIR	RMSE	LRUC	LRIn	LRCC	ES	LF
HS (200)	0.106	12.709%	-----	-----	0.000	0.004	0.000	0.150	2.509
HS (400)	0.096	13.378%	-----	-----	0.000	0.003	0.000	0.148	2.577
FHS (200)	0.177	4.013%	-----	-----	0.418	0.000	0.000	0.197	0.863
FHS (400)	0.213	3.344%	-----	-----	0.163	0.000	0.000	0.195	0.907
RiskMetrics	0.126	7.023%	31.879%	0.073	0.129	0.008	0.010	0.173	1.516
GARCH	0.098	12.709%	34.228%	0.056	0.000	0.009	0.000	0.144	2.811
IGARCH	0.102	12.709%	36.577%	0.057	0.000	0.001	0.000	0.151	2.433
FIGARCH	0.123	7.358%	34.564%	0.072	0.080	0.019	0.014	0.166	1.791
GARCHSK	0.079	17.057%	35.906%	0.044	0.000	0.010	0.000	0.132	3.591

Note 1: See the respective notes from Table 8.1

To provide the final piece of the puzzle, the VaR estimates for a short position on the Panamax data series are analysed, where these results are summarised in **Error! Reference source not found.**. The results for both the 1% and 5% thresholds, presented in Panels A and B, respectively, are exactly the same in that none of the models are found to have passed the statistical backtesting process. For this reason, no conclusion can be drawn as towhich model best evaluates the risk exposure faced by the charterer when dealing with the Panamax data series.

To summarise, the results for the dry-bulk market, in this respect, are somewhat disappointing in that one can only reach a conclusion in the case of the 1% threshold for the Capesize data series. A recurring theme does appear in this case, however, in that, as was the case for the tanker sector, non-parametric models were found to have outperformed parametric models, once again lending support to the findings of Angelidis and Skiadopoulos (2008).

8.4.3 Overview of Risk Estimation in the Shipping Markets

Having looked at the risk exposure for both major participants in the shipping market, this section brings the analysis together as a whole. The first thing to note here is that simple non-parametric models are found, as a whole, to outperform the parametric models in the majority of cases. This is interesting as it is provides direct support for the findings in a similar study by Angelidis and Skiadopoulos (2008), where they analysed the VaRs for a long position on freight rate indices and found that the simplest non-parametric models almost always outperformed the more complex parametric models.

The second interesting finding is that, in the tanker and dry bulk sectors, and with a few exceptions, the FIGARCH and GARCHSK models outperformed the other more standard parametric models, regardless of the position taken. A possible explanation for this is that both models take into account the shape of the supply function in the shipping market and the fact that supply is fixed in the short-term, therefore exacerbating this effect. Unfortunately, the results for the dry-bulk market are not as uniform; therefore one cannot draw any conclusions as to the best overall form of model for these series.

In order to check for robustness, the out-of-sample period was further sub-divided into the pre-crisis and post-crisis periods, where these extend from 3 October 2003 to 16 May 2008, and from 23 May 2008 to 16 June 2009, respectively. The results for the pre-crisis period support the findings over the total out-of-sample period in that the long position provides mixed results as to whether parametric or non-parametric models provide the best evaluation of the risk exposure faced by ship-owners in the market. This being said, non-parametric models are found to uniformly outperform parametric models with regards to the short position, and thus provide the best evaluation of the risk exposure incurred by charterers. Interestingly, however, the results for the post-crisis period contrast significantly with the other sample periods in that, for the long position and with the exception of the 1% VaR estimates for the Panamax data series, non-parametric models outperform parametric models in evaluating the potential risk exposure faced by ship-owners in the bulk shipping sectors. The picture changes, however, when looking at the risk exposure faced by charterers, where, with a few exceptions, parametric models are found to outperform non-parametric models.

8.5 Conclusion

The volatility of shipping freight rate returns is notoriously difficult to forecast. This chapter gave an insight into the future risk exposure of participants in the shipping freight markets, with a particular focus on the tanker and dry-bulk sectors of the market. This chapter also introduced the concepts of fractional integration and conditional skewness and kurtosis to forecasting volatility, thereby extending the shipping forecasting literature.

This chapter extends the literature on forecasting volatility by extending the work of Kavussanos and Dimitrakopoulos (2007) and Angelidis and Skiadopoulos (2008). This is because not only does it look at both the tanker and dry-bulk markets, as opposed Kavussanos and Dimitrakopoulos who only consider the tanker market, it also examines the spot freight rate series, as well as examining the risk exposure on both long and short positions, unlike Angelidis and Skiadopoulos who look at freight rate indices and only consider long positions, and introducing the concept of fractional integration to the mix. In addition, this chapter introduces the concept of conditional skewness and kurtosis to the VaR literature, via the use of the GARCHSK model.

Using data from five major ship sizes, across the both the tanker and dry bulk markets, the chapter uses an out-of-sample period of six years to perform ex-post forecasts of freight rate volatility, where these forecasts are then used to calculate the respective VaRs described above. Looking at the estimates of risk exposure, one can see that, in the vast majority of cases, the non-parametric models outperformed the parametric models, where overall, the filtered historical simulation model generally provided the best forecast of the VaR, regardless of the position or sector. This chapter has therefore provided a tool through which participants in the shipping markets can evaluate their potential risk exposure, where both are essential for making investment decisions and enabling ship-owners to plan the positioning of their vessels.

Appendix 8.A – Modified Diebold-Mariano Test Results

Table 8.11 – MDM Tests on a Long Position for the VLCC Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
RiskMetrics vs. GARCH	---	---	(---)
RiskMetrics vs. GARCHSK	---	---	(---)

Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
None	---	---	(---)

Note 1: VLCC denotes the weekly spot freight rates for a 270,000 DWT VLCC tanker carrying crude oil from Ras Tanura (Saudi Arabia) to Rotterdam (Netherlands).
SZMX denotes the weekly spot freight rates for a 130,000 DWT Suezmax tanker carrying crude oil from Bonny (Nigeria) to off the coast of Philadelphia (USA).
AFMX denotes the weekly spot freight rates for an 80,000 DWT Aframax tanker carrying crude oil from Sullom Voe (UK) to Bayway (USA).
CPSZ denotes the weekly spot freight rates for a 145,000 DWT Capesize bulk-carrier carrying iron ore from Tubarao (Brazil) to Rotterdam (Netherlands).
PNMX denotes the weekly spot freight rates for a 55,000 DWT Panamax bulk-carrier carrying grain from the Hampton Roads (USA) to Antwerp-Rotterdam-Amsterdam (Benelux).

Note 3: The sample period for the data used for this table extends from 3 October 2003 to 19 June 2009, with a total of 299 observations.

Note 4: The data used for this table is all sourced from the Clarkson Shipping Intelligence Network (www.clarskons.net).

Note 5: Tests where the results have been left blank imply that the resultant loss functions were identical and, therefore, no test needed to be performed.

Table 8.12 – MDM Tests on a Long Position for the SZMX Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Sim. (400) vs. Filtered Historical Sim. (200)	1.015	2.592	0.311
Filtered Historical Sim. (400) vs. FIGARCH	1.282	2.592	0.201
Filtered Historical Sim. (400) vs. Historical Sim. (200)	1.424	2.592	0.156

Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Sim. (200) vs. FIGARCH	2.010	1.968	0.045
Filtered Historical Sim. (200) vs. GARCHSK	2.041	1.968	0.042
Filtered Historical Sim. (200) vs. Historical Sim. (200)	2.530	1.968	0.012

Note 1: See notes from Table 8.11.

Table 8.13 – MDM Tests on a Long Position for the AFMX Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
GARCHSK vs. IGARCH	1.417	2.592	0.158
GARCHSK vs. Historical Sim. (400)	1.417	2.592	0.158
GARCHSK vs. Filtered Historical Sim. (200)	1.417	2.592	0.158
GARCHSK vs. Historical Sim. (200)	1.632	2.592	0.104
GARCHSK vs. FIGARCH	1.724	2.592	0.086
GARCHSK vs. RiskMetrics	1.731	2.592	0.084
Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Sim. (200) vs. FIGARCH	2.251	1.968	0.025
Filtered Historical Sim. (200) vs. Historical Sim. (200)	2.251	1.968	0.025

Note 1: See notes from Table 8.11.

Table 8.14 – MDM Tests on a Long Position for the CPSZ Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Simulation (200) vs. RiskMetrics	1.404	2.592	0.161
Filtered Historical Simulation (200) vs. GARCHSK	1.410	2.592	0.159
Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
None	---	---	(---)

Note 1: See notes from Table 8.11.

Table 8.15 – MDM Tests on a Long Position for the PNMX Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
RiskMetrics vs. Filtered Historical Sim. (200)	2.236	2.592	0.026
Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
RiskMetrics vs. FIGARCH	3.041	1.968	0.003
RiskMetrics vs. Filtered Historical Sim. (200)	3.041	1.968	0.003

Note 1: See notes from Table 8.11.

Table 8.16 – MDM Tests on a Short Position for the VLCC Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Historical Sim. (200) vs. Filtered Historical Sim. (200)	---	---	(---)
Historical Sim. (200) vs. Historical Sim. (400)	-1.738	2.592	0.083
Historical Sim. (200) vs. Filtered Historical Sim. (400)	1.504	2.592	0.134
Historical Sim. (200) vs. GARCHSK	-1.738	2.592	0.083

Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Historical Sim. (200) vs. Historical Sim. (400)	-2.191	1.968	0.029
Historical Sim. (200) vs. GARCH	-2.213	1.968	0.028
Historical Sim. (200) vs. IGARCH	-2.213	1.968	0.028
Historical Sim. (200) vs. RiskMetrics	-2.314	1.968	0.021

Note 1: See notes from Table 8.11.

Table 8.17 – MDM Tests on a Short Position for the SZMX Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Sim. (400) vs. GARCHSK	-1.140	2.592	0.255
Filtered Historical Sim. (400) vs. Filtered Historical Sim. (200)	-1.842	2.592	0.066

Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Sim. (400) vs. Filtered Historical Sim. (200)	-2.316	1.968	0.021

Note 1: See notes from Table 8.11.

Table 8.18 – MDM Tests on a Short Position for the AFMX Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Sim. (400) vs. Filtered Historical Sim. (200)	-1.417	2.592	0.158
Filtered Historical Sim. (400) vs. GARCHSK	-1.417	2.592	0.158
Filtered Historical Sim. (400) vs. Historical Sim. (400)	-1.417	2.592	0.158
Filtered Historical Sim. (400) vs. Historical Sim. (200)	-1.417	2.592	0.158

Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Sim. (400) vs. RiskMetrics	-2.408	1.968	0.017
Filtered Historical Sim. (400) vs. Filtered Historical Sim. (200)	-2.438	1.968	0.015
Filtered Historical Sim. (400) vs. Historical Sim. (200)	-2.563	1.968	0.011
Filtered Historical Sim. (400) vs. GARCHSK	-2.587	1.968	0.010
Filtered Historical Sim. (400) vs. FIGARCH	-2.682	1.968	0.008

Note 1: See notes from Table 8.11.

Table 8.19 – MDM Tests on a Short Position for the CPSZ Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
Filtered Historical Sim. (400) vs. Filtered Historical Sim. (200)	-2.140	2.592	0.033

Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
None	---	---	(---)

Note 1: See notes from Table 8.11.

Table 8.20 – MDM Tests on a Short Position for the PNMX Data Series

Panel A - Modified Diebold-Mariano Tests on the 1% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
None	---	---	(---)

Panel B - Modified Diebold-Mariano Tests on the 5% Value-at-Risk Estimates			
	MDM Stat	***t*-stat**	***p*-value**
None	---	---	(---)

Note 1: See notes from Table 8.11.

9 Conclusion

9.1 Introduction

This thesis aimed to provide an alternate look at the structure of freight rates in the bulk shipping sectors, which have been a constant source of academic interest since they were first discussed by Koopmans (1939). This being said, understanding the nature of these prices is not merely an academic exercise but does have definite practical applications in that freight rates form the price for transporting goods by sea, where over 7,500 million tonnes of goods were transported by sea in 2007. The correct structure of freight rates is of great interest in that freight rates play a pivotal role, and form the basis of almost every function in the shipping markets, from the determination of the price of the transport service through to the price of second-hand vessels. Therefore, a correct model for freight rates is vital for all participants in the shipping market, from the ship-owners and charterers themselves, right on down through the market to ship-brokers, maritime lawyers and other auxiliary parties.

The aim of this research was to expand on the traditional models of the structure of freight rates through the use of Autoregressive Fractionally Integrated Moving Average (ARFIMA), Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity (FIGARCH) and Generalised Autoregressive Conditional Heteroscedasticity with Skewness and Kurtosis (GARCHSK) models. Once this was done, these models were used to forecast spot freight rate levels and freight rate volatility, and then evaluate the underlying risk through the use of the Value-at-Risk methodology. By doing this, one should have gained a better understanding of the behaviour of the different moments of freight rates, thereby enabling participants in the shipping markets to have a better understanding of both the direction of spot freight rate levels, and the underlying risk.

This research is of interest to a number of different parties, both in terms of participants in the shipping market, as well as the financial markets as a whole. One reason for this is that it adds another dimension to the debate as to the exact structure of freight rates, as well as the degree of stationarity of these, as well as providing insight as to how the higher moments would affect freight rate risk as a whole. As mentioned above, this is of interest to participants in the shipping markets as this kind of understand is crucial for investment timing and planning decisions, as well as for indirectly linked parties to be able to quantify their exposure to the market. Interestingly, the fact that the shipping freight market is perhaps the only financial market in which the underlying asset is a service, as well as the fact that it may be used as a proxy for world trade, means that this market can be of interest to participants in other financial market.

Following this, the chapter continues by outlining the rationale behind, as well as highlighting the relevant empirical findings for, each of the hypotheses outlined in Chapter 1 of the thesis, where the first section deals with the hypothesis that freight rate levels follow a fractionally integrated process. Following this, the second sections examines the hypothesis that freight rate volatility also follows a fractionally integrated process, while the third section examines the hypothesis that incorporating conditional third and fourth moments of freight rates may give market participants a better understanding of their potential risk exposure. The fourth section summarised the findings as to the performance of the various risk models, while the fifth section summarises the overall findings and outlines proposals for further research.

9.2 Hypothesis 1 – The Dynamics of the First Moment

This thesis began its analysis, in Chapter 5, by investigating the proposal that freight rates follow a fractionally integrated process. The rationale behind this statement was that in the short-term, the supply function for shipping services is fixed, while demand is relatively price inelastic; however, in the longer-term, as new vessels are delivered, the supply function will expand accordingly. This means that in the short-term, freight rates will exhibit non-stationary behaviour in that, due to the fixed nature of supply, as demand increases, so will freight rates, but up to the point where freight rates make other, more expensive, alternative means of transportation viable, as was illustrated in Figure 1.3 in Chapter 1. However, as high freight rates induce ship-owners to order new vessels, and these vessels are delivered, usually after between 18 and 36 months, but this can be extended to over five years, the supply function will shift to the right, as illustrated in Figure 1.2 in Chapter 1, and freight rates will revert to their mean level. Thus, as has been shown, freight rates are mean reverting; however, this mean reversion process will occur with a lag, where this is exactly how one would expect a fractionally integrated process to behave.

In order to test this hypothesis fully, and given the inconclusive findings of standard unit root tests, Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA) and Autoregressive Fractionally Integrated Moving Average (ARFIMA) models were estimated and the results compared using standard forecast metrics, both in-sample and out-of-sample. One should note that prior to estimating the full ARFIMA models, and in order to determine whether the data series were merely fractionally integrated white noise, ARFIMA $(0;d;0)$ models were estimated, and the residuals tested using the Nielsen test for fractional integration, outlined in Nielsen (2005). The results of these tests indicated that the residuals of these models were non-stationary, hence it was concluded that the data was not white noise and therefore had some information content within it.

It was observed that across all five data series and the ARMA and ARIMA models, one should never need more than an ARMA$(1;1)$ or ARIMA$(1;1;1)$; therefore, when running the ARFIMA $(p;d;q)$ models, the conclusion was drawn that this type of lag dynamic should be sufficient. The results for these ARFIMA models were unanimous in determining significant *d*-parameter, which measure the level of fractional integration, where $-0.5 < d < 1.0$, thus indicating that the data series exhibit long-memory. In order to test the stationarity of these residuals, and therefore double-check that these results were not spurious, Nielsen tests were performed on the respective residuals, unanimously indicating that the residuals were stationary, and, as a result, that the ARFIMA model results were not spurious.

The results for in-sample comparisons between the three model types were somewhat inconclusive, however it is proposed that both the ARMA and ARFIMA models outperformed the ARMA models, where it is postulated that limitations in the size of the sample may have contributed to a lack of a conclusive results. In terms of the forecasting performance of the models, the ARMA models were found to outperform the ARIMA and ARFIMA models for the VLCC and Suezmax data, however, as these results are based on the assumption that freight rates are stationary, these results may be somewhat flawed. For the Aframax, Capesize and Panamax data series, the ARIMA models were found to outperform the ARMA and ARFIMA models.

This research thus concludes that there are some grounds for the hypothesis regarding the long-memory nature of freight rates thereby providing an alternative dimension to debate as to the true nature of the structure of the first moment of freight rates. As mentioned in the introduction to this chapter, above, this has a profound impact not only on the primary users of ships, i.e. ship-owners and charterers, but also on the wide number of auxiliary parties in the shipping markets. In addition to this, it may also provide an insight into other markets, for instance, the real estate market, in which the underlying asset in the market is also a real asset, or other such service-based industries.

9.3 Hypothesis 2 – The Dynamics of the Second Moment

Having established the dynamics of the first moment of the underlying freight rates, the obvious next question was whether a similar structure applies to the second moment, or volatility, of these freight rates. In this respect, the concepts of stationarity, non-stationarity and fractionally integration could be extended from the spot freight rate levels to the volatility of freight rates, as illustrated in Baillie, *et al.* (1996a). This research proposes that, should shocks to the volatility decay in a hyperbolic manner, as illustrated by news impact curve B in Figure 1.7, where $\lambda_k\left[1-\beta_1-\left(1-d\right)k^{-1}\right]\cdot\Gamma\left(k+d-1\right)\Gamma\left(k\right)^{-1}\Gamma\left(d\right)^{-1}\ ;k>1$, then the volatility series could be argued to follow a "fractionally integrated" process. The rationale behind this hypothesis is the same as for the spot freight rate levels. Imbalances in supply and demand in the short-term cause freight rate levels to "explode". Consequently, the volatility, or standard deviation, of these freights will also increase dramatically, until such a time as the level of spot freight rates stabilises. As new vessels are delivered, spot freight rates revert to the mean spot freight level, and volatility stabilises, however, this process of stabilisation occurs with a lag, due to the fixed nature of supply in the short-term.

One should note that the reasons for specifying the use of the GARCH, IGARCH and FIGARCH models is that each assumes a different rate of decay for shocks to volatility where the GARCH models assume an exponential rate of decay, the IGARCH models an indefinite rate of decay, and the FIGARCH model a hyperbolic rate of decay. Therefore, by determining the best model for the data series, one can draw conclusions as to the persistence of shocks in volatility, and therefore the potential risk exposure of involved parties.

Having run the models, it was found that past variance played a significant role in determining the level of volatility in the shipping freight markets, and that lagged returns are found to have an impact on the current returns in the market. When examining the results for GARCH and IGARCH models, the tentative conclusion was reached that shocks to volatility persisted indefinitely, regardless of vessel type. In contrast, however, the results from the FIGARCH models suggested that shocks with respect to freight rate volatility followed the hyperbolic rate of decay hypothesised.

In order to address the question as to which models were correct, AIC and SBIC measures were calculated and the models compared on this basis. The results for these measures led to the conclusion that, for the dry-bulk sector, the FIGARCH model provided the best fit in terms of the structure volatility, while, for the tanker sector, the IGARCH model was preferred. Although information criteria do provide a convenient means of choosing between models, Brooks and Burke (2003) argue that these standard metric suffer from a lack of ability in that they do not allow for the number of parameters in the models to change, thereby leading to reduced forecasting accuracy. In order to address this issue, the models were also compared on the basis of their ability to accurately determine and minimise the respective Value-at-Risk in Chapter 8.

9.4 Hypothesis 3 – Conditional Third and Fourth Moments

Moving on with the analysis of the moments, this research examined the higher, i.e. third and fourth, moments. Incorporating skewness and kurtosis into models of price series is well established, however, a relatively new introduction to the financial markets literature is the concept of conditional skewness and kurtosis. This thesis adopted these concepts, and tested the hypothesis that conditional skewness and kurtosis also plays a significant role in the structure of the underlying freight rates, where, to the best of the author's knowledge, this is done for the first time in the shipping literature and in the literature for markets in which real assets are traded. The rationale behind this hypothesis is that the shape of the supply function in the freight markets is such that when one is positioned at a relatively price elastic portion of the supply curve, the degree of skewness and excess kurtosis will be relatively low; however, as the price elasticity decreases, as short-term supply reaches its maximum level, and freight rates shoot up, so will the degree of skewness and excess kurtosis, resulting in an extremely fat-tailed, positively skewed distribution.

Chapter 7 of the thesis presents the results from the process of testing this hypothesis, where, in order to do this, the GARCHSK model, first introduced by León, *et al.* (2005), the FIGARCH model, developed by Baillie, *et al.* (1996a), and Bollerslev (1986)'s standard GARCH model were estimated, using freight rate returns from five different vessel types over 1,068 observations, and the results compared by looking at the characteristics of the respective conditional variance and using likelihood-ratio tests.

In the quasi-maximum likelihoods estimations of the GARCH and GARCHSK models, the GARCHSK model was found to capture more of the dynamics of the respective data series based on these results, and the conditional skewness and kurtosis parameters were found to be significant across most of the data series. In addition, the results of the likelihood ratio tests uniformly indicated the superiority of the GARCHSK model over the standard GARCH and the FIGARCH models, a fact that was confirmed through a comparison of the characteristics of the conditional variance for the respective data series. Therefore, this paper can conclude quite firmly that the GARCHSK model outperforms the GARCH and FIGARCH models in capturing the dynamics of the data.

There are multiple advantages to being able to capture the conditional skewness and kurtosis of the data series - the first, is that as the distribution of asset returns is skewed, and there is excess kurtosis, and therefore the traditional assumption of normality when estimating Values-at-Risk will result in an underestimation of the risk. Secondly, it enables one to better describe the distributional properties of financial asset returns, thus enabling one to better understand the performance of assets with these properties. Finally, one could look at the issue of portfolio construction to determine if the risk structure is truly optimal, and examining these properties would enable one to better price options in financial markets wherever these properties exist.

9.5 A Look at the Risk Exposure of Market Participants

Having estimated the risk models in Chapters 6 and 7, Chapter 8 provided a practical extension to the hypotheses outlined, by giving an insight into the future risk exposure of participants in the shipping freight markets, with a particular focus on the tanker and dry-bulk sectors of the market. This chapter also introduced the concepts of fractional integration and conditional skewness and kurtosis to forecasting volatility, thereby extending the shipping forecasting literature.

This analysis extended the literature on forecasting volatility by extending the work of Kavussanos and Dimitrakopoulos (2007) and Angelidis and Skiadopoulos (2008). This is because not only did it look at both the tanker and dry-bulk markets, as opposed Kavussanos and Dimitrakopoulos who only consider the tanker market, it also examines the spot freight rate series, as well as examining the risk exposure on both long and short positions, unlike Angelidis and Skiadopoulos who look at freight rate indices and only consider long positions, and introducing the concept of fractional integration to the mix.

Using data from five major ship sizes, across the both the tanker and dry bulk markets, the chapter uses an out-of-sample period of six years to perform ex-post forecasts of freight rate volatility, where these forecasts are then used to calculate the respective VaRs described above. Looking at the estimates of risk exposure, one can see that, in the vast majority of cases, the non-parametric models outperformed the parametric models, where overall, the filtered historical simulation model generally provided the best forecast of the VaR, regardless of the position or sector.

This analysis therefore provided a tool through which participants in the shipping markets can evaluate their potential risk exposure, where both are essential for making investment decisions and enabling ship-owners to plan the positioning of their vessels.

9.6 Summary and Proposals for Further Research

This thesis examined the structure of freight rates in the shipping freight market, where, in particular, the concepts of fractional integration, in terms of the first and second moments, as well as conditional skewness and kurtosis, were introduced for the first time in the shipping literature. The results of this empirical analysis suggest that, while shipping freight rate levels do appear to follow a fractionally integrated process, forecasts of the spot freight rate levels indicated that ARMA and ARIMA specifications were found to outperform the fractionally integrated specifications, although arguments do exist that simpler specifications outperform more complicated models in terms of forecasting ability. When modelling freight rate volatility, FIGARCH models were found to outperform other specifications in the dry-bulk sector, while the non-stationary IGARCH models were found to provide a better evaluation of volatility in the tanker sector. This being said, when incorporating conditional skewness and kurtosis into the picture, models which account for this outperform other specifications in this context. To conclude, when looking at calculating the risk exposure faced by market participants, the risk exposure incurred by ship-owners was found to be better evaluated using non-parametric models in the tanker sector, and parametric models in the dry-bulk sector. This being said, when evaluating the risk exposure faced by charterers, non-parametric specifications were found to outperform parametric models in both sectors.

As with any research, there are, however, some limitations in terms of the analysis and findings presented above. One possible drawback of this research is that it does not take into account the inflationary behaviour of prices in the dry-bulk sector, although Worldscale rates are adjusted each January in order to take account of this inflationary tendency. By modelling real instead of nominal freight rates, one could remove this non-stationary and time-varying trend, although this is left as an area for further research. One should also note that this inflationary trend could also pose a possible reason for the level of persistence seen in freight rate levels and returns. Furthermore, this research utilise raw spot freight rates levels themselves, as opposed to using the natural logarithm of these series, potentially ignoring the fact that freight rates can never be negative. This being said, when estimating the respective models using these log series, the results for which are not presented in this thesis due to space limitations, the results were found not to differ significantly to the results presented here.

Another possible limitation is the sample length, due to a lack of availability of data, and the fact that there is only one data source available to researchers in the shipping field. As with any single data source, this brings up questions as to the accuracy of the data concerned, as there is no means of verifying the freight rates reported. One should also note that the returns series used in Chapters 6, 7 and 8, may be biased upwards as a result of using arithmetic, as opposed to the more standard log, returns. The reason this approach is used is that, should one wish to incorporate an ARFIMA model into the mean equation for the volatility model, differencing the logs of the respective series to generate standard returns would render the series stationary. This should, however, not significantly influence the choice of model for the volatility of freight rates in that all returns series were treated identically. Finally, the accuracy of standard forecast metrics as applied to forecasts of volatility are called into question. The reason for this is that using squared returns as proxy for actual volatility is inherently flawed, where Lopez (2001) illustrated that this proxy is over 50% greater or smaller than the actual volatility 75% of the time, however, unfortunately, no better proxy exists as yet. This thesis addresses this issue by evaluating the accuracy of the respective Values-at-Risk for each model with the actual Value-at-Risk incurred.

To conclude, and given the fact that any research is an ongoing process, there is still room for further analysis. A possible extension to this thesis would be to determine whether the characteristics found with respect to spot freights could also apply to modelling second-hand prices in the second-hand ship vessel market, as well as the structure of Freight Forward Agreement (FFA) prices. One should note that, currently, the relatively recent launch of the FFA contracts means that one is currently limited in terms of the data set available, which could in turn cause the problems discussed above, and outlined by Schwert (1989); however, as time passes and more data becomes available, this problem should solve itself. This thesis has, for reasons of brevity and also so as to ensure complete clarity in terms of the findings, also not explored fully the problem of asymmetry in the volatility of freight rates. A possible extension in this respect could be to examine the application of Fractionally Integrated Exponential Generalised Autoregressive Conditional Heteroscedasticity (FIEGARCH) model to model freight rate volatility in the shipping freight market, thereby addresses this gap.

References and Bibliography

Adelman, I., 1965, Long Cycles: Fact or Artifact?, *The American Economic Review* 55, 444-463.

Adland, R., 2003, The Stochastic Behavior of Spot Freight Rates and the Risk Premium in Bulk Shipping, (Massachusetts Institute of Technology, Cambridge, MA).

Adland, R. and K. Cullinane, 2006, The Non-Linear Dynamics of Spot Freight Rates in Tanker Markets, *Transportation Research Part E: Logistics and Transportation Review* 42, 211-224.

Adland, R. and S. P. Strandenes, 2004, A Discrete-Time Stochastic Partial Equilibrium Model of the Spot Freight Market (SNF Report No. 04/04), (Centre for International Economics and Shipping, Institute for Research in Economics and Business Administration, Bergen).

Ait-Sahalia, Y., 1996, Nonparametric Pricing of Interest Rate Derivative Securities, *Econometrica* 64, 527-560.

Akaike, H., 1974, A New Look at the Statistical Model Identification, *IEEE transactions on Automatic Control* 19, 716-723.

Angelidis, T. and G. Skiadopoulos, 2008, Measuring the Market Risk of Freight Rates: A Value-at-Risk Approach, *International Journal of Theoretical & Applied Finance* 11, 447-469.

Aydogan, K. and G. G. Booth, 1988, Are There Long Cycles in Common Stock Returns?, *Southern Economic Journal* 55, 141.

Baba, Y., R. F. Engle, D. F. Kraft and K. F. Kroner, 1990, Multivariate Simultaneous Generalised ARCH, (Department of Economics, University of California, San Diego).

Backus, D. K. and S. E. Zin, 1993, Long-Memory Inflation Uncertainty: Evidence from the Term Structure of Interest Rates, *Journal of Money, Credit and Banking* 25, 681-700.

Baillie, R. T., 1989, Econometric Tests of Rationality and Market Efficiency, *Econometric Reviews* 8, 151 - 186.

Baillie, R. T., 1996, Long Memory Processes and Fractional Integration in Econometrics, *Journal of Econometrics* 73, 5-59.

Baillie, R. T. and T. Bollerslev, 1989, Common Stochastic Trends in a System of Exchange Rates, *The Journal of Finance* 44, 167-181.

Baillie, R. T. and T. Bollerslev, 1994a, Cointegration, Fractional Cointegration, and Exchange Rate Dynamics, *The Journal of Finance* 49, 737-745.

Baillie, R. T. and T. Bollerslev, 1994b, The Long Memory of the Forward Premium, *Journal of International Money and Finance* 13, 565-571.

Baillie, R. T., T. Bollerslev and H. O. Mikkelsen, 1996a, Fractionally Integrated Generalised Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 74, 3-30.

Baillie, R. T., C.-F. Chung and M. A. Tieslau, 1996b, Analysing Inflation by the Fractionally Integrated ARFIMA-GARCH Model, *Journal of Applied Econometrics* 11, 23-40.

Bali, T. G., H. Mo and Y. Tang, 2008, The Role of Autoregressive Conditional Skewness and Kurtosis in the Estimation of Conditional VaR, *Journal of Banking & Finance* 32, 269-282.

Basel, C. o. B. S., 1995a, An Internal Model-Based Approach to Market Risk Capital Requirements, in *Basel Commitee on Bank Supervision* (Basel Commitee on Bank Supervision, Basel, Switzerland).

Basel, C. o. B. S., 1995b, Planned Supplement to the Capital Accord to Incorporate Market Risks, in *Basel Commitee on Bank Supervision* (Basel, Commitee on Bank Supervision, Basel, Switzerland).

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

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

Berg-Andreassen, J. A., 1996, Some Properties of International Maritime Statistics, *Maritime Policy & Management* 23, 381-395.

Berg-Andreassen, J. A., 1997, The Relationship between Period and Spot Rates in International Maritime Markets, *Maritime Policy & Management* 24, 335 - 350.

Birkeland, S., 1998, Bulk Shipping Freight Rates, (Norwegian School of Economics and Business Administration, Bergen).

Bjerksund, P. and S. Ekern, 1995, Contingent Claims Evaluation of Mean-Reverting Cash Flows in Shipping, in L. Trigeorgis, ed.: *Real Options in Capital Investment: Models, Strategies, and Applications* (Praeger, Westport).

Black, F., 1976, The Picing of Commodity Contracts, *Journal of Financial Economics* 3, 167-179.

Black, F. and M. Scholes, 1973, The Pricing of Options and Corporate Liabilities, *The Journal of Political Economy* 81, 637-654.

Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31, 307-327.

Bollerslev, T., R. Y. Chou and K. F. Kroner, 1992, ARCH Modeling in Finance : A Review of the Theory and Empirical Evidence, *Journal of Econometrics* 52, 5-59.

Bollerslev, T. and H. O. Mikkelsen, 1996, Modeling and Pricing Long Memory in Stock Market Volatility, *Journal of Econometrics* 73, 151-184.

Bollerslev, T. and J. M. Wooldridge, 1992, Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances, *Econometric Reviews* 11, 143 - 172.

Box, G. E. P. and G. Jenkins, 1970. *Time Series Analysis, Forecasting and Control* (Holden-Day, San Fransisco).

Breidt, F. J., N. Crato and P. de Lima, 1993, Modeling Long-Memory Stochastic Volatility, (Johns Hopkins University, Baltimore).

Breitung, J. and U. Hassler, 2002, Inference on the Cointegration Rank in Fractionally Integrated Processes, *Journal of Econometrics* 110, 167-185.

Brennan, M. J. and E. S. Schwartz, 1979, A Continuous Time Approach to the Pricing of Bonds, *Journal of Banking & Finance* 3, 133-155.

Brooks, C., 2002. *Introductory Econometrics for Finance* (Cambridge University Press, Cambridge).

Brooks, C. and S. P. Burke, 2003, Information Criteria for GARCH Model Selection, *The European Journal of Finance* 9, 557-580.

Brooks, C., S. P. Burke, S. Heravi and G. Persand, 2005, Autoregressive Conditional Kurtosis, *Journal of Financial Econometrics* 3, 399-421.

Brooks, C. and G. Persand, 2003, Volatility Forecasting for Risk Management, *Journal of Forecasting* 22, 1-22.

Cabedo, D. J. and I. Moya, 2003, Estimating Oil Price 'Value at Risk' Using the Historical Simulation Approach, *Energy Economics* 25, 239-253.

Campbell, J. Y. and R. J. Shiller, 1987, Cointegration and Tests of Present Value Models, *The Journal of Political Economy* 95, 1062-1088.

Caner, M. and L. Kilian, 2001, Size Distortions of Tests of the Null Hypothesis of Stationarity: Evidence and Implications for the PPP Debate, *Journal of International Money and Finance* 20, 639-657.

Chan, K., K. C. Chan and G. A. Karolyi, 1991, Intraday Volatility in the Stock Index and Stock Index Futures Markets, *The Review of Financial Studies* 4, 657-684.

Chang, Y.-T. and H. B. Chang, 1996, Predictability of the dry bulk shipping market by BIFFEX, *Maritime Policy & Management* 23, 103-114.

Cheung, Y.-W., 1993, Long Memory in Foreign-Exchange Rates, *Journal of Business & Economic Statistics* 11, 93-101.

Cheung, Y.-W. and K. S. Lai, 1993, A Fractional Cointegration Analysis of Purchasing Power Parity, *Journal of Business & Economic Statistics* 11, 103-112.

Chowdhury, A. R., 1991, Futures Market Efficiency: Evidence from Cointegration Tests, *Journal of Futures Markets* 11, 577-589.

Christoffersen, P. F., 1998, Evaluating Interval Forecasts, *International Economic Review* 39, 841-862.

Christoffersen, P. F., 2003. *Elements of Financial Risk Management* (Academic Press, San Diego).

Clarksons Research Services Ltd., 2010, Clarksons Shipping Intelligence Network.

Conrad, C. and M. Karanasos, 2005, On the Inflation-Uncertainty Hypothesis in the USA, Japan and the UK: A Dual Long Memory Approach, *Japan and the World Economy* 17, 327-343.

Coulson, N. E. and R. P. Robins, 1985, Aggregate Economic Activity and the Variance of Inflation : Another Look, *Economics Letters* 17, 71-75.

Crato, N. and P. J. F. de Lima, 1994, Long-Range Dependence in the Conditional Variance of Stock Returns, *Economics Letters* 45, 281-285.

Crato, N. and P. Rothman, 1994a, Fractional Integration Analysis of Long-Run Behavior for US Macroeconomic Time Series, *Economics Letters* 45, 287-291.

Crato, N. and P. Rothman, 1994b, A Reappraisal of Parity Reversion for UK Real Exchange Rates, *Applied Economics Letters* 1, 139-141.

Crowder, W. J. and A. Hamed, 1993, A Cointegration Test for Oil Futures Market Efficiency, *Journal of Futures Markets* 13, 933-941.

Cullinane, K., 1992, A Short-Term Adaptive Forecasting Model for BIFFEX Speculation: A Box-Jenkins Approach, *Maritime Policy & Management* 19, 91 - 114.

Deaves, R. and I. Krinsky, 1995, Do Futures Prices for Commodities Embody Risk Premiums?, *Journal of Futures Markets* 15, 637-648.

Devanney, J. W., 1973, A Model of the Tanker Charter and a Related Dynamic Program, in P. Lorange, and V. D. Norman, eds.: *Shipping Management* (Institute for Shipping Research, Bergen).

Dickey, D. A. and W. A. Fuller, 1979, Distribution of the Estimators for Autoregressive Time Series With a Unit Root, *Journal of the American Statistical Association* 74, 427-431.

Dickey, D. A. and W. A. Fuller, 1981, Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica* 49, 1057-1072.

Diebold, F. X., J. Gardeazabal and K. Yilmaz, 1994, On Cointegration and Exchange Rate Dynamics, *The Journal of Finance* 49, 727-735.

Diebold, F. X., S. Husted and M. Rush, 1991, Real Exchange Rates under the Gold Standard, *The Journal of Political Economy* 99, 1252-1271.

Diebold, F. X. and R. S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 253-263.

Ding, Z., C. W. J. Granger and R. F. Engle, 1993, A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance* 1, 83-106.

Dixit, A., 1992, Investment and Hysteresis, *The Journal of Economic Perspectives* 6, 107-132.

Dixit, A. K. and R. S. Pindyck, 1994. *Investment under Uncertainty* (Princeton University Press, Princeton).

Domowitz, I. and C. S. Hakkio, 1985, Conditional Variance and the Risk Premium in the Foreign Exchange Market, *Journal of International Economics* 19, 47-66.

Engle, R. F., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50, 987-1007.

Engle, R. F. and T. Bollerslev, 1986, Modelling the Persistence of Conditional Variances, *Econometric Reviews* 5, 1 - 50.

Engle, R. F. and C. W. J. Granger, 1987, Co-Integration and Error Correction: Representation, Estimation, and Testing, *Econometrica* 55, 251-276.

Engle, R. F., D. M. Lilien and R. P. Robins, 1987, Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model, *Econometrica* 55, 391-407.

Eriksen, I. E. and V. D. Norman, 1976, Ecotank: Modelllanalyse for Tankmarkedenes Virkemate (Ecotank: A Model Analysis of the Tanker Markets), (Institute for Shipping Research, Norwegian School of Economics and Business Administration, Bergen).

Evans, J. J., 1994, An Analysis of Efficiency of the Bulk Shipping Markets, *Maritime Policy & Management* 21, 311-329.

Fox, R. and M. S. Taqqu, 1986, Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series, *The Annals of Statistics* 14, 517-532.

Gallant, A. R. and G. Tauchen, 1989, Semi-Non-Parametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications, *Econometrica* 57, 1091-1120.

Geweke, J. and S. Porter-Hudak, 1983, The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series Analysis* 4, 221-238.

Gilbert, C. L., 1986, Testing the Efficient Markets Hypothesis on Averaged Data, *Applied Economics* 18, 1149-1166.

Giot, P. and S. Laurent, 2003a, Market Risk in Commodity Markets: A VaR Approach, *Energy Economics* 25, 435-457.

Giot, P. and S. Laurent, 2003b, Value-at-risk for long and short trading positions, *Journal of Applied Econometrics* 18, 641-663.

Glen, D. R., 1997, The Market for Second-Hand Ships: Further Results on Efficiency Using Cointegration Analysis, *Maritime Policy & Management* 24, 245 - 260.

Glen, D. R. and P. Rogers, 1997, Does Weight matter? A Statistical Analysis of the SSY Capesize Index, *Maritime Policy & Management* 24, 351 - 364.

Granger, C. W. J., 1980, Long Memory Relationships and the Aggregation of Dynamic Models, *Journal of Econometrics* 14, 227-238.

Granger, C. W. J., 1981, Some Properties of Time Series Data and their Use in Econometric Model Specification, *Journal of Econometrics* 16, 121-130.
Granger, C. W. J., 1986, Developments in the Study of Cointegrated Economic Variables, *Oxford Bulletin of Economics & Statistics* 48, 213-228.
Granger, C. W. J., 1988, Some Recent Developments in a Concept of Causality, *Journal of Econometrics* 39, 199-211.
Granger, C. W. J. and R. Joyeux, 1980, An Introduction to Long-Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis* 1, 15-29.
Gray, J. W., 1990. *Shipping Futures* (Lloyd's of London Press, London).
Greene, M. T. and B. D. Fielitz, 1977, Long-Term Dependence in Common Stock Returns, *Journal of Financial Economics* 4, 339-349.
Greene, W. H., 2003. *Econometric Analysis* (Pearson Education Inc., Upper Saddle River, NJ).
Hafer, R. W., S. E. Hein and S. S. MacDonald, 1992, Market and Survey Forecasts of the Three-Month Treasury-Bill Rate, *The Journal of Business* 65, 123-138.
Haigh, M. S., 2000, Cointegration, Unbiased Expectations, and Forecasting in the BIFFEX Freight Futures Market, *Journal of Futures Markets* 20, 545-571.
Haigh, M. S. and M. T. Holt, 2002, Hedging Foreign Currency, Freight, and Commodity Futures Portfolios: A Note, *Journal of Futures Markets* 22, 1205-1221.
Hakkio, C. S. and M. Rush, 1989, Market Efficiency and Cointegration: An Application to the Sterling and Deutschemark Exchange Markets, *Journal of International Money and Finance* 8, 75-88.
Hakkio, C. S. and M. Rush, 1991, Cointegration: How Short is the Long run?, *Journal of International Money and Finance* 10, 571-581.
Hale, C. and A. Vanags, 1989, Spot and Period Rates in the Dry Bulk Market, *Journal of Transport Economics and Policy* 29, 281-291.
Harris, R. I. D., 1995. *Using Cointegration Analysis in Econometric Model* (Prentice Hall, Upper Saddle River).
Harvey, C. R. and A. Siddique, 1999, Autoregressive Conditional Skewness, *Journal of Financial & Quantitative Analysis* 34, 465-487.
Harvey, D., S. Leybourne and P. Newbold, 1997, Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting* 13, 281-291.
Hawdon, D., 1978, Tanker Freight Rates in the Short and Long Run, *Applied Economics* 10, 203-217.
Holt, C., 1957, Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages, (Carnegie Institute of Technology, Pittsburgh).
Hosking, J. R. M., 1981, Fractional Differencing, *Biometrika* 68, 165-176.
Hung, M.-W. and H. Zhang, 1995, Price Movements and Price Discovery in the Municipal Bond Index and the Index Futures Markets, *Journal of Futures Markets* 15, 489-506.
Hurst, H. E., 1951, Long-Term Storage Capacity of Reservoirs, *Transactions of the American Society of Civil Engineers* 116, 770-799.
Hylleberg, S., R. F. Engle, C. W. J. Granger and B. S. Yoo, 1990, Seasonal Integration and Cointegration, *Journal of Econometrics* 44, 215-238.
ISL Bremen, 2008. *Shipping Statistics Yearbook* (Institute of Shipping and Logistics, Bremen).
Jarque, C. M. and A. K. Bera, 1980, Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals, *Economics Letters* 6, 255-259.

Jiang, G. J., 1998, Nonparametric Modeling of U.S. Interest Rate Term Structure Dynamics and Implications on the Prices of Derivative Securities, *The Journal of Financial and Quantitative Analysis* 33, 465-497.

Johansen, S., 1988, Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control* 12, 231-254.

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

Johansen, S. and K. Juselius, 1990, Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand for Money, *Oxford Bulletin of Economics & Statistics* 52, 169-210.

JP Morgan, 1996, RiskMetrics - Technical Document, (New York).

Kang, S. H., S.-M. Kang and S.-M. Yoon, 2009, Forecasting Volatility of Crude Oil Markets, *Energy Economics* 31, 119-125.

Kapetanios, G., 2006, Nonlinear Autoregressive Models and Long Memory, *Economics Letters* 91, 360-368.

Kapetanios, G. and Y. Shin, 2008, GLS Detrending-Based Unit Root Tests in Nonlinear STAR and SETAR Models, *Economics Letters* 100, 377-380.

Kapetanios, G., Y. Shin and A. Snell, 2003, Testing for a Unit Root in the Nonlinear STAR Framework, *Journal of Econometrics* 112, 359-379.

Kavussanos, M. G., 1996, Comparison of Volatility in the Dry-Cargo Ship Sector: Spot versus Time Charters, and Smaller vesus Larger Vessels, *Journal of Transport Economics and Policy* 30, 67-82.

Kavussanos, M. G., 1997, The Dynamics of Time-Varying Volatilities in Different Size Second-Hand Ship Prices of the Dry-Cargo Sector, *Applied Economics* 29, 433 - 443.

Kavussanos, M. G. and A. H. Alizadeh, 2002, Seasonality Patterns in Tanker Spot Freight Rate Markets, *Economic Modelling* 19, 747-782.

Kavussanos, M. G. and D. N. Dimitrakopoulos, 2007, Measuring Freight Risk in the Tanker Shipping Sector, (Athens University of Economics and Business).

Kavussanos, M. G. and N. K. Nomikos, 1999, The Forward Pricing Function of the Shipping Freight Futures Market, *Journal of Futures Markets* 19, 353-376.

Kavussanos, M. G. and N. K. Nomikos, 2000a, Constant vs. Time-Varying Hedge Ratios and Hedging Efficiency in the BIFFEX Market, *Transportation Research Part E: Logistics and Transportation Review* 36, 229-248.

Kavussanos, M. G. and N. K. Nomikos, 2000b, Hedging in the Freight Futures Market, *Journal of Derivatives* 8, 41-58.

Kavussanos, M. G. and N. K. Nomikos, 2003, Price Discovery, Causality and Forecasting in the Freight Futures Market, *Review of Derivatives Research* 6, 203-230.

Kavussanos, M. G. and I. D. Visvikis, 2004, Market Interactions in Returns and Volatilities between Spot and Forward Shipping Freight Markets, *Journal of Banking & Finance* 28, 2015-2049.

Kawaller, I. G., P. D. Koch and T. W. Koch, 1990, Intraday Relationships Between Volatility in S&P 500 Futures Prices and Volatility in the S&P 500 Index, *Journal of Banking & Finance* 14, 373-397.

Kendall, M. G., A. Stuart and K. Ord, 1991. *Kendall's Advanced Theory of Statistics* (Oxford University Press, New York, USA).

Kilic, R., 2007, Conditional Volatility and Distribution of Exchange Rates: GARCH and FIGARCH Models with NIG Distribution, *Studies in Nonlinear Dynamics & Econometrics* 11, 1-31.

Kim, Y., 1990, Purchasing Power Parity in the Long Run: A Cointegration Approach, *Journal of Money, Credit and Banking* 22, 491-503.
Koekebakker, S., R. Adland and S. Sodal, 2006, Are Spot Freight Rates Stationary?, *Journal of Transport Economics and Policy* 40, 449-472.
Kolb, R. W., 1992, Is Normal Backwardation Normal?, *Journal of Futures Markets* 12, 75-91.
Koopmans, T. C., 1939. *Tanker Freight Rates and Tankship Building* (Erven F. Bohn, Haarlem).
Koutmos, G. and M. Tucker, 1996, Temporal Relationships and Dynamic Interactions Between Spot and Futures Stock Markets, *Journal of Futures Markets* 16, 55-69.
Koyck, L. M., 1954. *Distributed Lags and Investment Analysis* (North-Holland Publishing Company, Amsterdam).
Krehbiel, T. and L. C. Adkins, 1994, Interest Rate Futures: Evidence on Forecast Power, Expected Premiums, and the Unbiased Expectations Hypothesis, *Journal of Futures Markets* 14, 531-543.
Krehbiel, T. and L. C. Adkins, 2005, Price Risk in the NYMEX Energy Complex: An Extreme Value Approach, *Journal of Futures Markets* 25, 309-337.
Kuester, K., S. Mittnik and M. S. Paolella, 2006, Value-at-Risk Prediction: A Comparison of Alternative Strategies, *Journal of Financial Econometrics* 4, 53-89.
Kumar, M. S., 1991, Forecasting Accuracy of Crude Oil Futures Prices, *IMF Staff Papers* 39, 432-461.
Kuo, B.-S. and C.-C. Tsong, 2005, Bootstrap Inference for Stationarity, (Helsinki Center of Economic Research).
Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin, 1992, Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root : How Sure are We that Economic Time Series have a Unit Root?, *Journal of Econometrics* 54, 159-178.
Lai, K. S. and M. Lai, 1991, A Cointegration Test for Market Efficiency, *Journal of Futures Markets* 11, 567-575.
Lehar, A., M. Scheicher and C. Schittenkopf, 2002, GARCH vs. Stochastic Volatility: Option Pricing and Risk Management, *Journal of Banking & Finance* 26, 323-345.
Lensberg, T. and H. Rasmussen, 1992, A Stochastic Intermodal Model of the Tanker Market, (Foundation for Research in Economics and Business Administration, Bergen).
León, Á., G. Rubio and G. Serna, 2005, Autoregresive Conditional Volatility, Skewness and Kurtosis, *The Quarterly Review of Economics and Finance* 45, 599-618.
Lo, A. W., 1991, Long-Term Memory in Stock Market Prices, *Econometrica* 59, 1279-1313.
Lopez, J. A., 1998, Methods for Evaluating Value-at-Risk Estimates, *Federal Reserve Bank of New York: Economic Policy Review* 4, 119-124.
Lopez, J. A., 2001, Evaluating the Predictive Accuracy of Volatility Models, *Journal of Forecasting* 20, 87-109.
Ma, C. W., 1989, Forecasting Efficiency of Energy Futures Prices, *Journal of Futures Markets* 9, 393-419.
Maddala, G. S. and I.-M. Kim, 1998. *Unit Roots, Cointegration, and Structural Change* (Cambridge University Press, Cambridge).

Mandelbrot, B., 1963, The Variation of Certain Speculative Prices, *The Journal of Business* 36, 394-419.
Martinussen, E. A., 1993, The Importance of Scrapping and Lay-Up for the Valuation of VLCCs, (Norwegian School of Economics and Business Administration, Bergen).
McLeod, A. I. and K. W. Hipel, 1978, Preservation of the Rescaled Adjusted Range, 1: A Reassessment of the Hurst Phenomenon, *Water Resources Research* 14, 491-508.
Meese, R. A. and K. J. Singleton, 1982, On Unit Roots and the Empirical Modeling of Exchange Rates, *The Journal of Finance* 37, 1029-1035.
Mossin, J., 1968, An Optimal Policy for Lay-Up Decisions, *Swedish Journal of Economics* 70, 170.
Ng, V. K. and S. C. Pirrong, 1996, Price Dynamics in Refined Petroleum Spot and Futures Markets, *Journal of Empirical Finance* 2, 359-388.
Nielsen, M. Ø., 2005, Multivariate Lagrange Multiplier Tests for Fractional Integration, *Journal of Financial Econometrics* 3, 372-398.
Norman, V. D. and T. Wergeland, 1981, Nortank - A Simulation Model of the Freight Market for Large Tanker (Report 4/81), (Centre for Applied Research, Norwegian School of Economics and Business Administration, Bergen).
Peiró, A., 1999, Skewness in Financial Returns, *Journal of Banking & Finance* 23, 847-862.
Phillips, P. C. B. and P. Perron, 1988, Testing for a Unit Root in Time Series Regression, *Biometrika* 75, 335-346.
Pindyck, R. S. and D. L. Rubinfeld, 1991. *Econometric Models and Economic Forecasts* (McGraw-Hill, Singapore).
Robinson, P. M., 1990, Time Series with Strong Dependence, in C. A. Simms, ed.: *Advances in Econometrics: Volume 1, 6th World Conference* (Cambridge University Press, Cambridge).
Robinson, P. M., 1994, Efficient Tests of Nonstationary Hypotheses, *Journal of the American Statistical Association* 89, 1420-1437.
Rosenblatt, M., 1956, A Central Limit Theorem and a Strong Mixing Condition, *Proceedings of the National Academy of Sciences* 42, 43-47.
Sadorsky, P., 2006, Modeling and Forecasting Petroleum Futures Volatility, *Energy Economics* 28, 467-488.
Sarma, M., S. Thomas and A. Shah, 2003, Selection of Value-at-Risk models, *Journal of Forecasting* 22, 337-358.
Schwarz, G., 1978, Estimating the Dimension of a Model, *The Annals of Statistics* 6, 461-464.
Schwert, G. W., 1989, Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business and Economic Statistics* 7, 147-159.
Sephton, P. S. and H. K. Larsen, 1991, Tests of Exchange Market Efficiency: Fragile Evidence from Cointegration Tests, *Journal of International Money and Finance* 10, 561-570.
Shea, G. S., 1991, Uncertainty and Implied Variance Bounds in Long-Memory Models of the Interest Rate Term Structure, *Empirical Economics* 16, 287-312.
Sims, C. A., 1980, Macroeconomics and Reality, *Econometrica* 48, 1-48.
Sowell, F., 1986, Fractionally Integrated Vector Time Series, Department of Economics (Duke University, Durham, NC).

Sowell, F., 1992, Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, *Journal of Econometrics* 53, 165-188.
Steigerwald, D. G., 1996, Purchasing Power Parity, Unit Roots, and Dynamic Structure, *Journal of Empirical Finance* 2, 343-357.
Stoll, H. R. and R. E. Whaley, 1990, The Dynamics of Stock Index and Stock Index Futures Returns, *The Journal of Financial and Quantitative Analysis* 25, 441-468.
Stopford, M., 2009. *Maritime Economics* (Routledge, London).
Strandenes, S. P., 1984, Price Determination in the Time Charter and Second Hand Markers, (Norwegian School of Economics and Business Administration).
Strandenes, S. P., 1986, NORSHIP - A Simulation Model for Bulk Shipping Markets, (Centre for Applied Research, Norwegian School of Economics and Business Administration, Bergen).
Tanaka, K., 1999, The Nonstationary Fractional Unit Root, *Econometric Theory* 15, 549-582.
Taqqu, M. S., 1975, Weak Convergence to Fractional Brownian Motion and to the Rosenblatt Process, *Zeitschtift fiir Wahrseheinlichkeitstheorie und Verwandte Gebiete* 31, 287-302.
Theil, H., 1958. *Economic Forecasts and Policy* (North Holland, Amsterdam).
Theodossiou, P., 1998, Financial Data and the Skewed Generalised T Distribution, *Management Science* 44, 1650-1661.
Tinbergen, J., 1931, Ein Schiffbauzclus?, *Weltwirtschaftliches Archiv* 34, 152-164.
Tinbergen, J., 1934, Scheepsruimte en Vrachten, *De Nederlandshe* March.
Tong, H., 1990. *Nonlinear Time Series: A Dynamical Systems Approach* (Oxford University Press, Oxford).
Tse, Y. K., 1995, Lead-Lag Relationship Between Spot Index and Futures Price of the Nikkei Stock Average, *Journal of Forecasting* 14, 553-563.
Tvedt, J., 1996, Market Structures, Freight Rates and Assets in Bulk Shipping, (Norwegian School of Economics and Business Administration, Bergen).
Tvedt, J., 1997, Valuation of VLCCs Under Income Uncertainty, *Maritime Policy & Management* 24, 159 - 174.
Tvedt, J., 1998, Valuation of a European Futures Option in the BIFFEX Market, *Journal of Futures Markets* 18, 167-175.
Tvedt, J., 2003, A New Perspective on Price Dynamics of the Dry Bulk Market, *Maritime Policy & Management* 30, 221 - 230.
Vasicek, O., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* 5, 177-188.
Veenstra, A. W. and P. H. Franses, 1997, A Cointegration Approach to Forecasting Freight Rates in the Dry Bulk Shipping Sector, *Transportation Research Part A: Policy and Practice* 31, 447-458.
Wahab, M. and M. Lashgari, 1993, Price dynamics and error correction in stock index and stock index futures markets: A cointegration approach, *Journal of Futures Markets* 13, 711-742.
Weiss, A. A., 1986, Asymptotic Theory for ARCH Models: Estimation and Testing, *Econometric Theory* 2, 107-131.
Wergeland, T., 1981, Norbulk: A Simulation Model of Bulk Freight Rates, (Norwegian School of Economics and Business Administration, Bergen).
Whittle, P., 1951. *Hypothesis Testing in Time Series Analysis* (Almqvist and Wiksells, Uppsala).

Winters, P. R., 1960, Forecasting Sales by Exponentially Weighted Moving Averages, *Management Science* 6, 324-342.

Witte, J. G., Jr., 1963, The Microfoundations of the Social Investment Function, *The Journal of Political Economy* 71, 441-456.

Working, H., 1970, Economic Functions of Futures Markets, in H. H. Bakken, ed.: *Futures Trading in Livestock: Origins and Concepts.* (Chicago Mercantile Exchange, Chicago).

Zannetos, Z. S., 1966. *The Theory of Oil Tankship Rates* (MIT Press, Cambridge, MA).