Managing Risk with Freight Futures from IMAREX:

Testing Hedging Effectiveness and the Unbiasedness Hypothesis

Thomas Samdal Rasmussen
Eivind Tversland

NHH

Managing Risk with Freight Futures

from IMAREX:

Testing Hedging Effectiveness and the Unbiasedness Hypothesis

Master thesis by:

Thomas Samdal Rasmussen and Eivind Tversland

Advisors:

Jonas Andersson and Siri P. Strandenes

Master Profile:

Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS AND

BUSINESS ADMINISTRATION

Abstract

This thesis investigates the hedging effectiveness and unbiasedness hypothesis of the IMAREX PM4TC freight futures contract. First, we present theory of dry bulk shipping and risk management. Then, we study hedging effectiveness of the futures contract. This is performed by using regression models and a VAR model to calculate constant hedge ratios and a VAR-GARCH model to calculate time-varying hedge ratios. We find the hedging effectiveness to range from 29,50% to 31,78%, when hedging one of the four T/C routes underlying the futures contract. Hedging with time-varying hedge ratios is in most cases shown to be superior to hedging with a constant hedge ratio.

Finally, the unbiasedness hypothesis is studied. We find that one month to maturity futures contracts give an unbiased prediction of the spot price at maturity. This implies that a hedger can trade in one month to maturity futures contracts without paying a risk premium. A rolling hedge can thus be executed efficiently. In addition, the futures price can be used to guide in decision making. Unbiasedness is also indicated for two and three months to maturity contracts, but, due to a small data sample and residual diagnostics problems, we have not drawn any conclusions.

Acknowledgements

Contents

1. General Notes

1.1 Introduction

During the last century, the shipping industry has experienced tremendous development: vessel sizes have been increasing continuously, companies have been more and more able to take advantage of economies of scale, and the volume of international trade has grown enormously. The main characteristics of the shipping business have not changed much, however. Still today, the business is defined by its highly volatile freight rates, seasonality, strong business cycles and capital intensiveness.

During the day-to-day operation of his ships, a shipowner is exposed to many risks: bunker prices fluctuate to a great extent, port-congestion might lead to delays and accidents happen. On the longer term, the shipowner must keep an eye on the changes of ship values and prices in the newbuilding and demolition market. As shipping is one of the world's most international industries, a shipowner is also exposed to currency risk, political environments and the world economy as a whole. These factors combined, make due for a high risk industry.

The shipping business is also characterised by its low barriers to entry and exit; all that is needed to enter the industry is a ship and a crew. Crews are available cheap, and the second-hand market for ships is very liquid. In times with booming freight rates, many banks are more than willing to lend money for purchasing a ship. When the rates eventually fall and the shipowner experiences cash-flow problems, new and hopeful investors are waiting to purchase the ship, hoping to make a bargain. Only the most adept shipowners survive in the long-run. Shipping is thus a business with high competition; it is a game that is all about the survival of the fittest, and many fortunes have been made or lost, playing the game.

Being such a high-risk business, it is evident that risk management and analysis of the market conditions are of outmost importance. Modern financial instruments like Forward Freight Agreements (FFAs), freight futures and freight options can be very useful to manage some of the risk in shipping. The instruments can for example be used to hedge future costs and/or revenues. Market participants that actively manage risk with such tools will then be less exposed to short-term volatility in the market than they would otherwise be. They can thereby

be more fitted to deal with the ups and downs of the business than their competitors, and that can be just enough to give them the extra edge to make a fortune.

Since 1985, there has been great progress in risk management in shipping. We have seen the rise and fall of the BIFFEX freight futures contract, and Forward Freight Agreements have been introduced. In recent years, the Norwegian marketplace IMAREX has re-introduced freight futures to the market. IMAREX is also providing bunker fuel oil derivatives and freight options. In this thesis, we will focus on dry-bulk freight futures contracts from IMAREX. We will explain the concept of Forward Freight Agreements, freight futures and the BIFFEX contract in later chapters. IMAREX will be properly introduced in Chapter 1.3.1.

1.2 Objective

As risk management is such an important part of the shipping business, we find it interesting to examine freight futures and find out how they can help shipowners to manage their risk. We choose to study the PM4TC freight futures contract traded at IMAREX, as this is one of their most traded contracts. In the spirit of prior studies of the BIFFEX contract[1], we will investigate the contract's hedging performance. We will also perform tests to see if the unbiasedness hypothesis holds for the PM4TC. Both hedging performance and the unbiasedness hypothesis will be extensively explained in later chapters.

1.3 Involved Parties

In this chapter we will briefly present the different parties with relevance for our thesis. These are: IMAREX, NOS and The Baltic Exchange.

1.3.1 IMAREX – The International Maritime Exchange

The International Maritime Exchange is the first and, at the time being, only authorized and regulated marketplace for trading and clearing of maritime derivatives. IMAREX opened for trading on 2 November 2001 and is publicly listed on Oslo Stock Exchange. IMAREX is

---

[1] Se for example Thuong and Visscher (1990), Kavussanos and Nomikos (1999, 2000a, 2000b, 2000c, 2001) and Kavussanos et al. (2004). The BIFFEX contract will be presented thoroughly in Chapter 3.4 and Chapter 3.5

cleared through the central clearing house NOS (Norsk Oppgjørssentral). It is regulated by the Financial Supervisory Authority of Norway (Kredittilsynet). The underlying indices for the freight futures and options traded on IMAREX are provided by The Baltic Exchange.

Trading on the IMAREX is either done as a direct member or via a member bank. In January 2007, IMAREX has 153 members and counting. In addition to transaction services and trade in bulk-, tanker-, fuel-oil- and power-derivatives, IMAREX offers clearing services in cooperation with NOS and information services such as market pricing and data distribution.

On 1 September 2006, IMAREX merged with the clearing house NOS and created the company IMAREX NOS ASA. In the future, the group will expand into markets where freight and/or energy are in focus.

1.3.2 NOS – Norwegian Futures and Options Clearinghouse (Norsk Oppgjørssentral)

NOS was founded in 1987 and is licensed by the Norwegian Ministry of Finance. Since 1990, the clearing house has offered clearing and settlement services to the derivatives markets within the financial, energy and freight sectors. NOS is the clearing central for all IMAREX derivatives. The merger of the two companies enables them to develop new solutions for clearing of derivatives and thus increases the liquidity for its customers.

1.3.3 The Baltic Exchange

The Baltic Exchange is the world's only independent source of maritime market information for the trading and settlement of physical and derivative contracts. The Baltic Exchange's first freight index was launched in 1985. Today, they publish many different freight indices. The indices are based on daily assessments on the dry- and wet-bulk routes, weekly sale and purchase, demolition and forward prices. The assessments are made by using a panel of international shipbrokers. The Baltic Exchange membership base consists of 550 companies and 2000 individuals (December 2006). They represent the majority of the world's shipping interests.

1.4 Outline

Chapter one has now given a short presentation of the objective of the thesis. In addition, the involved parties have been presented.

In **Chapter two**, a presentation of the bulk shipping industry is given. The four shipping markets are presented, and an examination of what determines freight rates are performed. We also give some explanations as to why the shipping business is as volatile as it is.

In **Chapter three**, risk management in shipping is discussed. We present theory on forward and futures contracts in general and Forward Freight Agreements and freight futures in particular. A review of prior studies of hedging effectiveness and the unbiasedness hypothesis of freight futures is also given. Finally, we present the concept of basis risk and discuss briefly how to hedge optimally with freight futures.

In **Chapter four,** we present different models for calculating optimal hedge ratios. First, models for finding constant hedge ratios are presented. The presented models are a Classical Linear Regression Model (CLRM), a regression model with lags and a VAR model. We then present models for calculating time-varying hedge ratios. The presented models are a VAR-GARCH model, a VECM-GARCH model and a VECM-GARCH-X model. How to calculate the hedging performance is shown for all models.

In **Chapter five**, we test the hedging performance of the IMAREX PM4TC freight futures contract. Tests are performed with both constant and time-varying hedge ratio models. The models are tested for proper specifications and conclusions are drawn.

Chapter six contains tests of the unbiasedness hypothesis for the IMAREX PM4TC freight futures contract. First, the conditions necessary for the unbiasedness hypothesis to hold is discussed. Then, the properties of the data series are presented. Finally, cointegration techniques are used to test the hypothesis.

In **Chapter seven**, the thesis is summarised and conclusions are made.

2. Bulk Shipping

As we will study the dry bulk freight futures contracts traded on IMAREX, we will now give a presentation of the bulk shipping industry. First, we will present bulk shipping in general. Then, we will examine what determines the freight rates in bulk shipping and show how and why they are as volatile as they in fact are.

The modern bulk shipping industry can be traced back to the seventeenth century and the coal trade between the north of England and London. There are two common definitions of bulk cargo. The first one defines bulk cargo as "anything whose physical characteristics allow it to be handled in bulk". The second defines bulk cargo as "any cargo that is transported by sea in large consignments in order to reduce the unit costs" (Stopford, 1997). We see that both definitions emphasize that bulk shipping is built on the minimisation of unit costs. The main principle of bulk shipping is: "one ship, one cargo."

As bulk shipping is all about minimising unit costs, efficiency improvements are something that interests a shipowner to a great extent, and he has essentially four ways to go: First, he can use bigger ships to exploit economies of scale. Second, he can reduce the number of times the cargo is handled. Third, the cargo handling procedures can be made more efficient, and finally, the stock size can be reduced.

The volume of seaborne trade has grown considerably since the end of the 19th century. There has also been a substantial increase in the use of bulk shipping to exploit economies of scale. Stopford (1997) writes that the bulk shipping industry has been so successful in minimising unit costs that the nominal price per ton for transporting coal is much the same as it was 125 years ago. This is mainly due to an enormous increase in ship sizes. For example, the size of the biggest ore carriers has increased from 24.000 dwt. in the 1920s to 300.000 dwt. in the 1990s. Today, the bulk fleet consists of over 9.000 vessels and bulk cargo account for more than half of the world's seaborne trade.

There are four main characteristics of bulk cargo that determine if it can be transported as bulk or if it will be transported in liner shipping. These are the volume transported, the value of the cargo, its physical handling and stowage characteristics, and the regularity of the

material flow. If the volume is high and the value relatively low, the cargo will probably be transported in bulk.

Having decided to transport a cargo in bulk, one then has to decide what type of ship and which handling gear to use. This choice is determined by the different characteristics of the bulk cargoes. In general, we divide bulk cargo into five main groups:

First, we have the liquid bulk cargo. The liquid cargoes fall into three main groups: crude oil and products, LNG and LPG, and vegetable oil and liquefied chemicals. Liquid bulk is stored in tanks, handled by pumps and transported in tankers.

Second, we have the homogenous bulk cargo, which is often divided into major and minor bulk. The major bulk cargos, which count five in number, are iron ore, coal, alumina, grain and phosphate rock. The volume transported of the five major bulk trades makes them the driving force behind the dry bulk carrier market. The minor bulk cargos consist of a mass of raw materials and semi-manufactures that are shipped partly or totally in bulk. The minor bulk cargoes are in many ways the most complex of the bulk cargoes. Because of the low volume transported, the minor bulks can sometimes be transported by liner shipping. Common for major and minor bulk cargos, is that they are both shipped in large quantities and are handled with grabs and conveyers.

Third, we have the unit load cargo. This is cargo that must be handled separately, like for example wind mills.

We then have the wheeled cargo, which requires special ships with access ramps and multiple decks. A car carrier is a good example of such ships.

Finally, we have refrigerated cargo. This can for example be fruit or other perishable commodities that need chilling during transport.

2.1 Determination of Freight Rates

In the following section we are going to study the determination of freight rates. In order to do so, we will first present the four shipping markets and explain how they interact. Then, we

will look at the determination of supply and demand of shipping. Finally, we will present figures that show how the market clear and why the shipping business is so volatile.

2.1.1 The Four Shipping Markets

Stopford (1997) divides the shipping market into four main groups: the newbuliding market, the freight market, the sales and purchase market, and the demolition market. All four markets interact and together they determine the freight rates.

2.1.1.1 The Newbuliding Market

The newbuliding market, which mainly consists of shipowners and ship-builders, will ideally reflect the need for capacity of a certain type of ship. This is often not the case. It can take up to four years from the day a ship is ordered until it is delivered. At the time of delivery, the freight rates might have changed considerably from when it was ordered. It is therefore very important for an investor to make a thorough analysis of the expected future spot rates before ordering a ship. When making the expectations, he should consider the whole lifetime of the ship. Stopford (1997) states that the newbuliding prices are as volatile as the second-hand prices and that they sometimes follow the same pattern.

2.1.1.2 The Second-Hand Market

The participants in the second-hand market are shipowners. It is an auxiliary market in the sense that trades in the second-hand market does not affect the number of vessels or the total number of dead-weight tonnes in the shipping business. According to economic theory, the second-hand market makes sure that the vessels are reallocated from the least efficient to the most efficient operators. Thus, the second-hand market facilitates the efficient use of capital and helps to reduce the transport costs in world trade.

There are many participants in the second-hand market. This ensures that the market is liquid and that exit barriers in the shipping business are pushed down. Hence, an efficient second-hand market has a positive effect on the competition between shipowners. The exit barriers will not be entirely eliminated, however. Because the freight rates and ship prices correlate, low freight rates will lead to low ship prices. A shipowner who is not efficient enough to

operate the ship during times with low freight rates might not be satisfied with the corresponding low second-hand prices. This can for example be if the selling price is not high enough to redeem his loan on the vessel. He might then continue to operate the ship despite of the low rates. So an effective second-hand market does not eliminate the exit costs; it only reduces them.

The prices for used ships are very volatile. This makes the second-hand market an important arena for potential gains and losses. The activity of speculating on ship prices is called asset play. The second-hand value of a vessel depends upon many factors; we can for example mention freight rates, inflation, the condition of the ship, age and expectations of future freight rates. Glen and Martin (2002) state that the price of second-hand ships correlates very well with movements of spot and time charter rates.

2.1.1.3 The Freight Market

The participants in the freight market are shipowners, charterers and brokers. Shipowners supply freight and charterers, who need transport for their goods, demand freight. Brokers bring shipowners and charterers together.

The term "charter party" stands for an agreement between a shipowner and a charterer whereby a ship is chartered either for one voyage or for a period of time. There are four main types of contractual arrangements in the freight market: the voyage charter (spot charter), the time charter, the bare boat charter and the contract of affreightment.

Voyage Charter
A voyage charter is a contract for one or more voyages. The agreement states a named vessel, a specific route and the amount of cargo to be transported for a fixed price per ton. The shipowner manages the ship and crew and is responsible for the payment of all expenses including voyage costs. If the contract is made for more than one voyage, it will reduce the shipowner's unemployment risk. A voyage charter that includes one voyage only is commonly called a spot charter agreement.

Time Charter

A time charter is an agreement where a vessel is hired for a specific amount of time and money. The shipowner still manages the ship and crew, but the charterer decides the ports of destination within the trading limits agreed. The length of the charter can be the amount of time to complete a single voyage or up to a period of several years. The shipowner pays the operating costs of the vessel but the charterer pays all voyage and cargo handling costs. A time charter agreement thus reduces the total risk for the shipowner. Payment is most often denoted in USD/day.

Fluctuations in Spot and Time Charter Rates

The degree of seasonal fluctuations of shipping freight rates varies across durations of contracts and vessel sizes. Beenstock and Vergottis (1989) find that, for all types of vessels, the seasonal fluctuations decline as we move from spot rates to one-year and three-year time charters. Time charter rates are formed as the expectation of future spot rates. We therefore expect that time charter rates already have incorporated the expected future fluctuations of the spot rates. Movements in time charter rates are thus smoother than movements in spot rates.

The larger fluctuations in spot rates are also explained by the spot rates being more risky than the time charter rates. This is because a company that operates in the spot market are exposed to many kinds of risk that a company with time charter agreements are not. For instance, a spot market operator faces the possibility that he, for a period of time, might not be able to fix a contract for one of his ships. This is called unemployment risk. The spot market operator might also not be able to get freight revenue for some days because he must relocate a vessel from one port to another in order to get a contract. In addition, the spot market operator is, contrary to the company with the time charter agreement, exposed to voyage cost fluctuations. A shipowner that operates in the time charter market must therefore be prepared to offer a discount to cover the risk reduction compared to the spot market. However, time charter rates might still be higher than spot rates if the spot rate is expected to rise in the future.

Bare Boat Charter

The bare boat charter is a contract that differs significantly from the time charter and the voyage charter contracts. It is an agreement where the charterer takes complete legal responsibility and full charge of the vessel for a period of time – often for years. The charterer

appoints the crew and pays all running costs. The shipowner is thus left only with capital costs and has therefore less risk.

Contract of Affreightment

The contract of affreightment (COA) is an agreement where the operator is to transport an agreed amount of cargo over a period of time at a fixed price. The operator is responsible for all expenses. It is not specified in the contract which vessel to use for the transport. Risk is reduced in such an agreement, as the operator is guaranteed work and payment for a long period of time. If a COA is made for many years, the terms of payment might be renegotiated periodically – for example once per year.

2.1.1.4 The Demolition Market

When a vessel is old, technical obsolete and/or inefficient, it must be scrapped. The most important factor in determining the activities in the scrapping market is the freight rates. When the freight rates are high, very few and mostly technical obsolete ships are scrapped. When the freight rates are low, the least efficient ships are starting to get scrapped. Expectations of future freight rates will also play a role. The demolition market depends on the steel price as well, and a higher price for steel makes it more profitable to scrap ships.

When making scrapping decisions, the shipowners must also consider the international regulatory environment. Requirements regarding ship specifications change from time to time, and ships must often meet these demands within a set date. It has for example been demanded that all ships of a specific type must have double hull. If it is too costly to upgrade the ships to meet this requirement, the ships will be scrapped.

Due to the mentioned facts, the amount of scrapping is time-varying. Scrapping prices will thus fluctuate and add more uncertainties to the otherwise risky business of shipping.

2.1.1.5 Connecting the Four Shipping Markets

The four shipping markets are tightly connected by cash-flow and transactions. The markets tend to move together, and the correlation between them is a result of the interaction of shipowners, brokers, charterers and shipyards.

To explain the interactions, we can consider an increase in the demand for freight. This leads to higher freight rates and an increase in expected future earnings. As the price of a ship should equal the discounted expected earnings for the rest of its life, prices of both newbuildings and second-hand ships will increase. New ships are ordered from the shipyards, and lack of capacity subsequently drives prices even further up. Fewer ships will be scrapped, as old ships tend to be given a longer life when freight rates are high. This will in turn lead to an increase in scrap prices.

If, on the other hand, demand does not increase as much as expected, there might be an overweight of supply when the newbuildings are delivered. As supply is greater than demand, freight rates will decline. Old and inefficient ships will be laid up or scrapped, pushing scrapping prices down. Second-hand prices then fall as expectations of the future are gloomier. Fewer ships will be ordered, and thus the newbuilding prices falls. The market will then be in recession until demand again increases, expectations of the future improve and the wheels start turning again.

We see then how the freight rates and the cash-flow bind the four markets together and how expectations of the future are of high importance. When later explaining futures contracts, we will see that they can be used to predict the direction of future spot rates, information that is invaluable to a shipowner.

2.1.2 Demand for Shipping

Like any other market, prices in shipping are determined by the clearing of supply and demand. Demand for shipping is very volatile and unpredictable. Stopford (1997) mentions five important variables that determine the demand for shipping; these are the world economy in general, what commodities are traded by sea, average haul, political events and transport costs.

Regarding the total amount of transportation demanded and supplied at sea, it is more appropriate to measure in ton-miles[2] instead of dead-weight tonnes[3]. The reason for this is

---

[2] A ton-mile is defined as one ton being moved one mile (The Baltic Exchange).

that the measure ton-miles includes the distance the commodities are transported in the calculation. At the same time, ton-miles will reflect efficiency changes in vessels. The distance factor in the ton-mile calculation is often referred to as the haul of the trade. The average haul is then an expression that incorporates the total transport distance of all ships world wide.

The commodities that need transportation by sea are produced by the many different industries around the world. This makes the world economy an important – and maybe the most important – factor in determining the demand for shipping. Aspects of the world economy that has an effect on the demand for transport at sea are the business cycle, the trade elasticity[4] and the trade development cycle. Stopford (1997) concludes that the business cycle in the world industry is the most important cause of short-term fluctuations in seaborne trade and ship demand.

It is best to divide the discussion of seaborne commodity trade in two parts: the short- and long-term perspective. Some commodity trades are very seasonal and thus have a great influence on the demand for transport at sea in the short-term. Such commodities are typically agricultural products. It can be very difficult for a shipper to know when a specific commodity is finished (e.g. harvested) and needs transportation. The shippers must therefore rely heavily on the spot charter market in order to meet their tonnage requirements. In the long-term, on the other hand, there are some other factors that market participants must be aware of. These are mainly changes in demand, changes in supply sources, relocation of processing and, finally, the shippers transport policy. For example, the mature economies have been characterised by deindustrialisation throughout the three last decades. Industrial production has been moving to Asian countries, and products often need to be transported over longer distances to get to the market. This has increased average haul and thus the demand for shipping.

---

[3] A dead-weight ton is a common measure of a ship's carrying capacity, i.e. the number of tons (2240 lbs.) of cargo, stores and bunkers that a vessel can transport. Physically, the measure can be explained as the difference between the number of tons of water a vessel displaces "light" and the number of tons it displaces "when submerged to the 'deep load line'" (The Baltic Exchange).

[4] The trade elasticity is the percentage growth in seaborne trade divided by the percentage growth in the world production. This number has been positive during the last few decades.

Political events may have large, sudden and unexpected impact on the demand for shipping. When we use the term political events, we refer to wars, revolutions, strikes, different political decisions, laws, etc. One event that had a great impact on demand was the Six-Day War between Egypt and Israel in 1967. The war resulted in the closure of the Suez Canal, and ships were forced to sail around Africa. As a consequence, the average haul increased dramatically.

During the course of the last century, we have seen dramatically improvements in shipping efficiency. Due to larger ships, improved vessel efficiency and better organization of the shipping companies, the cost for transporting goods at sea has been considerably reduced. In fact, some routes still have the same nominal cost for transporting goods as they did several decades ago. This has made shipping a popular mean of transporting goods, and demand has steadfastly increased.

2.1.3 Supply of Shipping

We will now present the mechanisms of supply of shipping. Stopford (1997) brings up five different determinants; these are the size of the world fleet, fleet productivity, shipbuilding, scrapping and freight rates. Supply is rigid and slow to change, and the nature of supply differs from the short- to the long-term perspective.

In the short-run, the size of the world fleet is given; it is not possible to change the number of vessels or the transport capacity measured in dead-weight tonnes. However, the operators can adjust the operating speed of the vessels or move the ships to and from lay-up or storage. This will affect the number of ton-miles available and therefore also supply. Increasing the vessel speed is, in other words, a way of increasing productivity.

We can measure the fleet productivity in ton-miles. The productivity then depends upon speed, deadweight utilization, lay-up, port time and loaded days at sea. Clarkson (1991) studied how the average VLCC was operating during a year. The result from this survey was that it spent only 137 days carrying cargo. Ballast time accounted for 111 days, and cargo handling accounted for 40 days during the year. The rest of the year was spent on non-trading activities.

In the long-run, a shipowner can buy new ships and scrap old ones. It is then possible to change the total number of dead-weight tonnes in the market. It is expected that, over time, supply of seaborne transport will grow in proportion to demand for seaborne transport.

Freight rates and expectations of future freight rates are the most important regulator of supply. When rates and expectations are high, shipowners might want to build more ships, and supply will increase. When rates and expectations are low, they might want to scrap vessels and supply is reduced.

2.1.4 Freight Rates

We will now show how freight rates are determined by linking supply and demand. The mechanisms for determining freight rates are rather simple and follow ordinary micro-economic theory. If there are too many ships, the freight rates will be low, and if there are too few ships, the freight rates will be high. The market participants try to balance the demand and supply.

2.1.4.1 The Supply Curve for a Single Vessel

First, we will show the supply curve for a single vessel. The supply curve, shown in Figure 2.1, describes the relationship between freight rates and the number of ton-miles transported at sea for a single ship.

Figure 2.1: Supply Function for a Single Ship
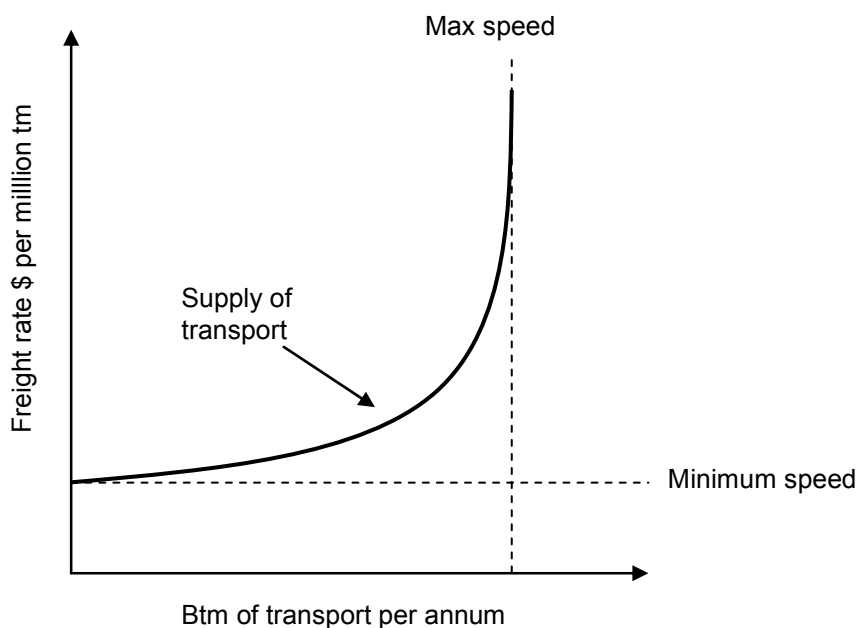
We see that the supply curve is horizontal at first, then grows exponentially and finishes off as a vertical line. If the freight rate is high enough, the vessel will operate at full speed. If the freight rate decreases, the ship will slow down in order to reduce bunker costs (fuel costs). This reduces the total supply of transport in the market. Each specific vessel has got a minimum possible operating speed. The vessel can not be manoeuvred properly unless

holding at least this speed. If the freight rate decreases even further, the shipowner will prefer to put the ship into lay-up. At this point the supply curve is horizontal.

2.1.4.2 The Supply Curve for a Fleet of Vessels

We will now explain the supply curve for a fleet of vessels. This curve is shown in Figure 2.2.

Figure 2.2: Supply Function for a Fleet of Vessels
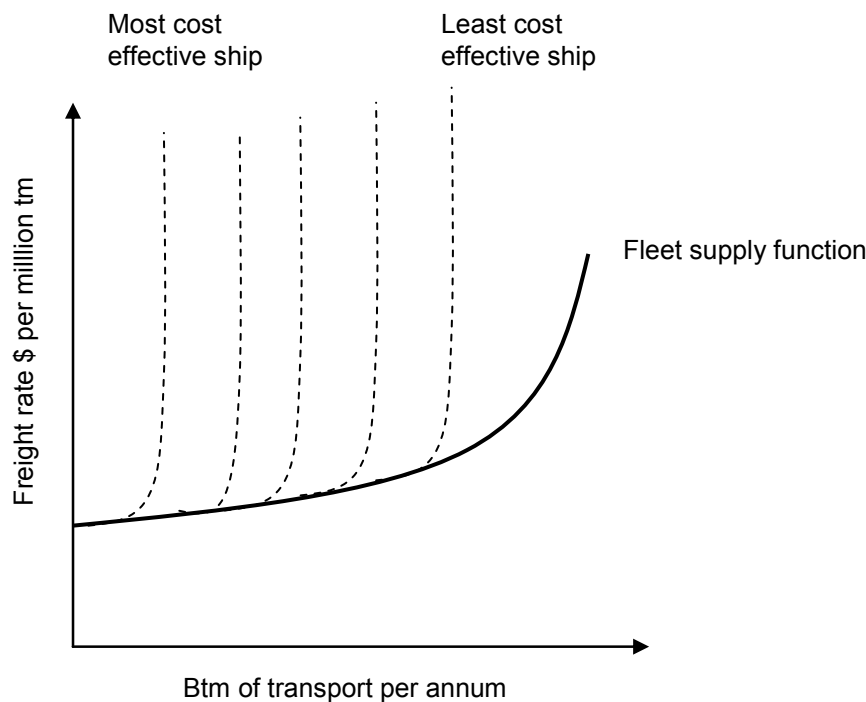
Source: Stopford (1997)

The fleet supply function is the aggregate of the supply functions of the individual ships. In Figure 2.2, each of the vertical hockey shaped lines represents a single vessel. The leftmost line illustrates the supply function of the most efficient ship. As we move to the right, the vessels are getting less and less efficient.

The fleet supply curve depends on three factors. First, we have the operating costs. Old ships tend to have higher operating costs compared to new ships and are therefore moved more quickly into lay-up when the freight rate falls. When the least efficient ship is moved into lay-up, the supply of transport will of course be reduced accordingly. The second least efficient

ship is now operating at break-even, and the other ships earn a small margin. If the freight rate falls even further, more vessels will be moved into lay-up.

The second factor is the size of the ships. Larger ships can normally transport goods at a lower cost per ton compared to smaller ships. Therefore, in times of recession, large ships tend to drive smaller ships into lay-up if they compete for the same cargo. On some occasions, however, the smaller ships will be more advantageous if there is not enough demand to fill up the biggest ships.

Finally, we have the relationship between the operating speed and freight rates. Because each vessel will increase its speed when freight rates rise, the total number of ton-miles for the entire fleet will increase in proportion to the number of vessels (the factor of proportion will of course depend on the average size of the fleet). The opposite is true when the freight rates fall.

We see that, when some or all ships are operating below maximum speed, the supply curve is very elastic. A small raise in the freight rates will then lead to an enormous increase of supply. Near top speed, in the short-run, supply is very inelastic. A small increase in demand will then boost the freight rates tremendously.

The supply curve is considered to lie constant in the short-run. In the long-run, however, there is time for new ships to be built and for old ships to be scrapped. The supply curve will consequently shift. Increases in efficiency and changes in costs will also shift the curve. For example, an increase in bunker prices will shift the entire supply curve upwards.

2.1.4.3 The Demand Curve for Sea Transport

As shown in Figure 2.3, the demand curve is very inelastic. One reason for this is that the transportation cost of goods makes up only a small portion of the total costs. Another reason is that there are few other alternatives of transportation. A shipper, who has a specific consignment of goods to transport, will therefore transport the goods at almost any level of freight rates. On the other hand, the same shipper would not want to make another trip if the freight rates are low (because he has no more goods to transport).

2.1.4.4 Equating Supply with Demand

When analysing the freight rate mechanisms, it is important to take the time factor into consideration. The momentary, short-term and long term equilibrium will now be discussed in turn.

First, we have the momentary equilibrium where deals in the market have to be made right away. Ships are now ready to load, and cargoes are awaiting transport. The freight rates can get very high if there is a surplus of cargo. On the other hand, the freight rates will fall if there is excess supply. Consequently, there can be huge volatilities in the freight rates in the very short-term.

We then have the short-term equilibrium. It is now possible to adjust the supply of transport through lay-up of ships, adjustment of speed or by using the ships to store goods. As combined carriers may choose between operating in different markets, the supply curve in the different markets may be shifted. The fleet supply-curve itself, however, is assumed to lie fixed in the short-run.

Figure 2.3: Supply/Demand Equilibrium
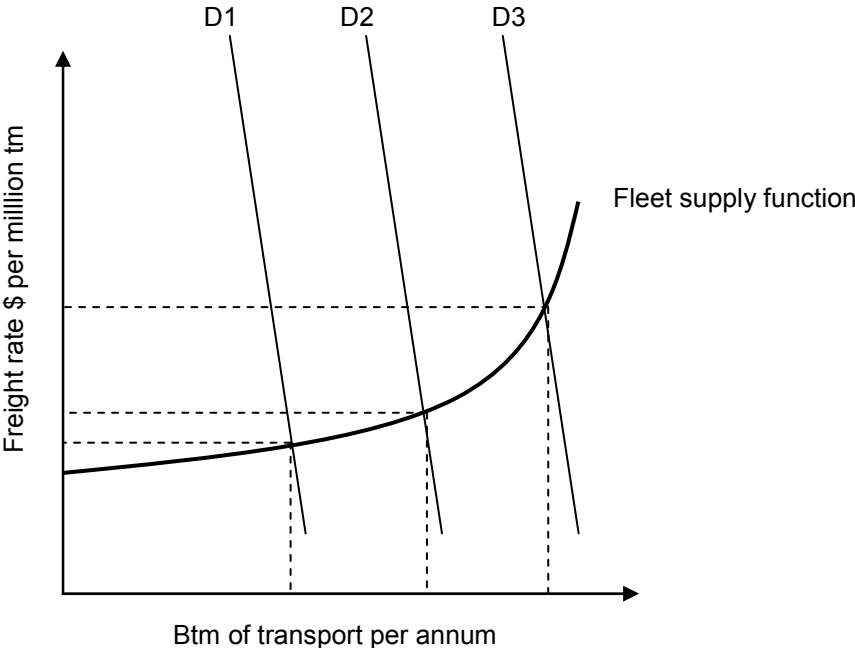
D1, D2 and D3 in Figure 2.3 represent different levels of demand for shipping. We can see that if the demand shift from D1 to D2, the freight rates only increase slightly. The reason for

this is that the operators adjust to the new demand by moving ships out of lay-up and increase the vessel speed. The freight rates will rise dramatically, however, if demand increases further to D3. If demand rises even more, there will eventually be no excess capacity in the market. The charterers will then have to bid against each other in order to get the available capacity. This is an unstable situation where charterers try to find cheaper supply sources and shipowners tend to invest in new ships.

Finally, we have the long-term equilibrium. The shipowners now have time to adjust the supply by scrapping or ordering new ships. In the model, this means that the fleet supply curve is allowed to shift. In the long-run, the shippers can also adjust their production process and rearrange their supply sources. Because of the factors discussed in Chapter 2.1.1.1 and Chapter 2.1.1.4, there is a significant time-lag before supply can adjust to demand. At the same time, demand is very volatile. The combination of these two aspects creates the framework of the cyclical shipping industry.

2.1.4.5 An Example of the Volatility of the Freight Rates

The freight rates are, as mentioned several times, very volatile and make up the greatest risk of the shipping business. We can illustrate the volatility of the business with the following example.

Figure 2.4: Freight Rate Example
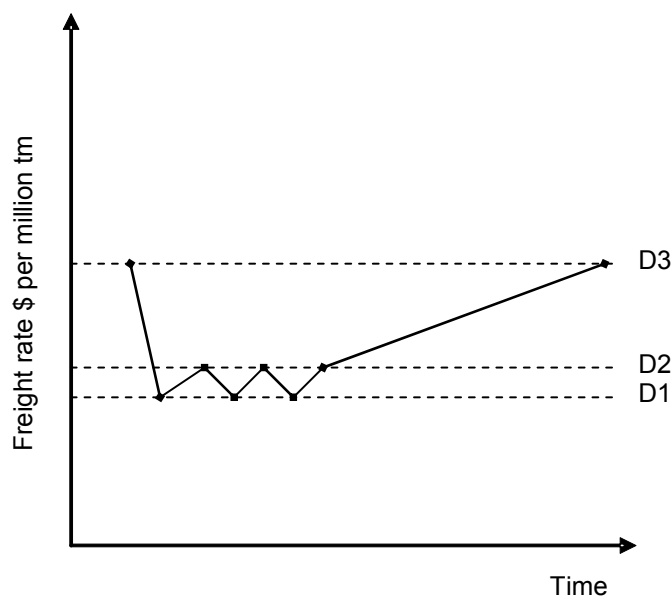
The horizontal lines, D1 to D3 in Figure 2.4, represent the level of freight rate determined by the fleet supply function and the demand functions D1 to D3 in Figure 2.3. The supply curve does not shift in the example. At the beginning of the example time-period, the freight rate is very high and times are good. The economy then suffers an adverse demand shock and demand falls from D3 to D1 (Figure 2.3). This can for example be caused by terrorist attacks, like September 11th, which subsequently affects the business climate in the world economy. We see from Figure 2.4 that the freight rate falls considerably. As time passes, demand shifts between D1 and D2 a few times as the economy stabilises and tries to get back on track. From Figure 2.3, we see that an increase in demand from D1 to D2 is about half as much as an increase from D1 to D3, but has very little impact on the freight rate. This is due to the high elasticity of the fleet supply function in this particular area. The world economy finally stabilises, expectations of the future improve and demand slowly increases. Thus, the freight rate also increases gradually and approaches the high level it had before the terrorist attacks.

The volatility of freight rates can also be shown with an actual time series of freight rates. Figure 2.5 shows the time charter freight rate assessments of the P1A route[5] from 02.01.2004 to 23.02.2007.

Figure 2.5: Route P1A Time Charter Assessments (02.01.2004 – 23.02.2007)

---
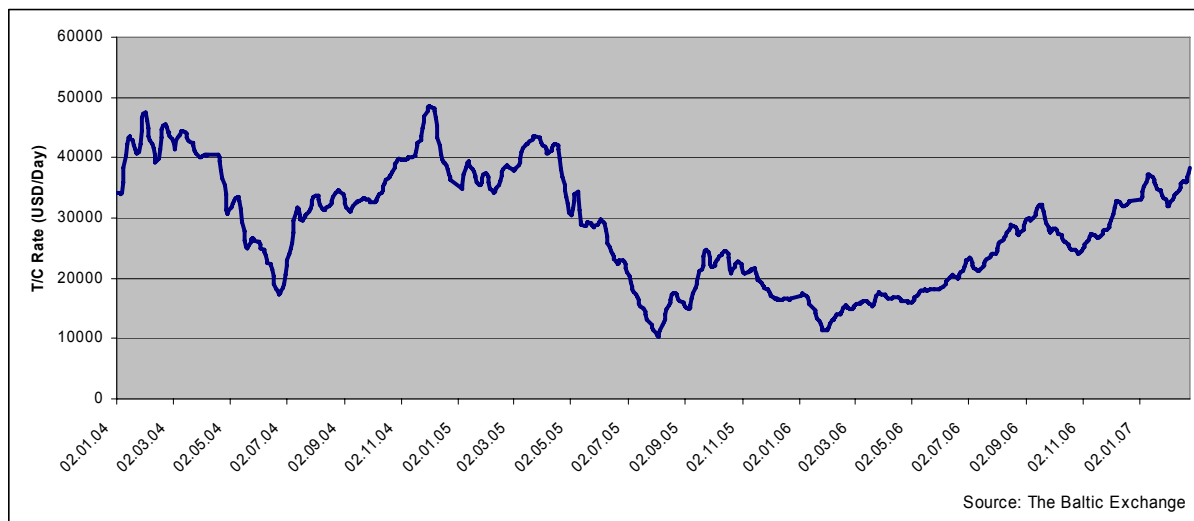
[5] P1A: Transatlantic Round Voyage (Panamax).

We see that the freight rate varies considerably during the time-period in the figure. The highest freight rate is recorded at 48.512 USD/day on 1 December 2004 and the lowest is recorded at 10.383 USD/day on 3 August 2005. The freight rate has thus fallen about 80% during the course of merely eight months, something which is not just a one time event, but is in fact very common in the business of shipping.

2.1.4.6 The Cyclicality of the Shipping Business

The mechanisms that equate the fast moving demand with the rigid supply make the shipping business very cyclical. Stopford (1997) identifies four stages of a shipping-cycle. First, we have *the through* where freight rates are at the operating cost level. Second, we have *the recovery* stage. The world economy is now improving and demand for sea transport rises. A shortage of supply makes the freight rates skyrocket. Shipowners therefore order new ships. After some time, supply catches up with demand and *the peak* stage takes over. Some ships are still in the building process and when they are delivered, supply will exceed demand. Freight rates will fall again and we are in *the collapse* stage. The collapse is often accompanied by an appalling outlook of the world economy. After the collapse, a new cycle can begin.

The shipping cycles are very difficult, if not impossible, to predict. Stopford (1997) shows that the lengths of the different stages vary from cycle to cycle. In addition, they seem to follow no apparent pattern. The volatility of changes in the freight rates are also varying in time.

The cyclicality in shipping is an economically valuable mechanism in the sense that it allows only the most efficient operators to survive. The most efficient operator will prosper during good times and be able to ride out the storm in bad times. The least efficient operator might, on the other hand, lose so much money during the bad times that he has to sell his vessels and exit the market.

3. Risk Management in Shipping

Having established the huge volatility of freight rates and thus the need for risk management and hedging, we now turn to the role of forward and futures contracts. First, we explain the concept of forward and futures contracts in general. We then present a model for pricing of the contracts. Thereafter, we explain the specific contracts available in the shipping business, namely Forward Freight Agreements and freight futures. We continue with a presentation of the unbiasedness hypothesis for forwards and futures in shipping and also explain the price discovery role of such contracts. Finally, we summarize some prior studies on the subject.

3.1 Forward Contracts

A forward contract is an agreement between two parties to buy or sell a specific asset, at a specific future point in time, for a price agreed today. The contracts are made over-the-counter (OTC), which means that they are made directly between two parties. The agreement is often facilitated by a broker.

As an example, we can think of a contract made between a producer of apples and a producer of cider. The apple farmer is anticipating harvest in September, but thinks that the forward prices in May are favorable and is afraid that the prices might fall before harvest comes. He therefore wants to lock prices in May to hedge his price risk. The producer of cider might conversely be afraid that prices could increase before September and also wants to hedge his price risk. The two parties therefore make an agreement in May, where they specify the number of apples to be delivered by the farmer to the cider producer in September and the price the cider producer is to pay for the apples. Thus, they have both locked their positions in May and are therefore free from risk of further price fluctuations.

In general, we define the spot price of an asset at time t as S_t. The forward price for delivery at time t = T is defined as $F_{0,T}$ and is agreed between the parties at time t = 0. For the apple farmer, who is short in the forward contract, the payoff at contract maturity (expiration date) in September is:

$$F_{0,T} - S_T \tag{3.1}$$

i.e. he gets the agreed forward price and has to give up the spot value of the apples. The apple farmer is also long in apples and his payoff of that position is S_T. His profit of the two positions combined is thus:

$$(F_{0,T} - S_T) + S_T = F_{0,T}$$

and we see algebraically that he is guaranteed a profit of $F_{0,T}$ at time t = T.

The cider producer is correspondingly long in the forward contract and short in apples. His contract payoff at time t = T is:

$$- F_{0,T} + S_T \tag{3.2}$$

i.e. he pays $F_{0,T}$ and gets the spot value S_T of the apples. The payoff of his short position in apples is (- S_T) and his combined profit of the two positions is given by:

$$(- F_{0,T} + S_T) - S_T = - F_{0,T}$$

The cider producer is thus guaranteed delivery of the apples and pays $F_{0,T}$ as agreed. This shows that both the apple farmer and the cider producer are perfectly hedged and know exactly their cash-flow at time t = T.

In practice, the two parties can agree to settle the contract at maturity by paying the difference between the spot price and the agreed delivery price in cash, as opposed to going through with the actual exchange of apples against cash.

3.1.1 Pricing of Forward Contracts

When analyzing the pricing of forward contracts, it is usually distinguished between investment assets and consumption assets. Investment assets are assets held for investment purposes, like for example stocks or bonds. Consumption assets are assets held for consumption. This is for example apples, corn, oil etc. Consumption assets can further be divided into storable- and non-storable assets. Examples on storable assets are the

aforementioned apples and corn, whereas non-storable assets are for example electricity and freight, the last one being the focus in our thesis.

3.1.1.1 Investment Assets

Pricing of forward contracts on investment assets are done through non-arbitrage arguments and leads to the following simple formula:[6]

$$F_{0,T} = S_0 * e^{(r - \delta) * T}$$
(3.3)

where S_0 is the current spot price of the asset, r is the continuously compounded interest rate, and δ is the dividend yield on the asset. The formula is also intuitively simple and can be explained as follows: You make an agreement today to receive for example a stock at time $t = T$. The differences between this agreement and purchasing the stock right away are that payment is postponed and that you do not receive dividends on the stock until delivery. It is therefore fair that you pay interest on the postponed payment of the stock and receive a full discount for the forgone dividends. As we see from the Equation (3.3), this is exactly what is done in the pricing of the forward contract. If Equation (3.3) does not hold, it is possible to produce a risk-free profit through arbitrage.

3.1.1.2 Consumption Assets

Pricing of consumption assets or commodities are more difficult and depends on the properties of the specific asset in question. Differences in storability, storage costs, production and demand will lead to different pricing formulas. In general, however, Equation (3.3) will hold also for storable commodities, with δ here being the return that makes an investor willing to buy and then lend a commodity. δ is then called the commodity lease rate. Spot and forward prices are thus linked by the so called cost-of-carry-relationship.

For non-storable commodities like freight, the non-arbitrage and cost-of-carry arguments are violated. The forward rates are then free to be determined by supply and demand, and

---

[6] For a more thorough presentation of the pricing of forward contracts, see for example Robert L. McDonald (2003) or John C. Hull (2003).

speculative activities (Batchelor et al., 2007). When presenting Forward Freight Agreements (FFAs), we will revisit and expand our presentation of the pricing of such forward contracts.

3.2 Futures

A futures contract is an exchange traded forward contract. The futures contracts are standardized regarding delivery dates, delivery locations and procedures. As opposed to forward contracts, where each contract is settled at maturity, the futures are settled daily. This procedure is known as marking-to-market. As long as the market is liquid, it is therefore possible to close one's position at any time.

Whether one is long or short in a futures contract, a clearinghouse connected to the exchange is acting as counterparty in the agreement. This effectively removes all counterparty risk, which one is exposed to in a regular forward contract. As a guarantee for the clearinghouse, each trader is required to deposit a certain amount of money to cover the daily settlements of the futures. This deposit is called a margin. The clearinghouse requires the margin to exceed a minimum level, which is called the maintenance margin. If the daily settlements reduce the margin below this level, the clearinghouse makes a margin call, asking for the deposit of additional funds to cover the margin.

Margrabe (1976) demonstrates that if the interest rates were not random, forward and futures prices would be exactly the same. Because of the marking-to-market and randomly varying interest rates, forward and futures prices differ. French (1983) shows empirically that the prices are still very similar, and that the difference between forward and futures contracts is increasing with the length of the contracts.

3.3 Forward Freight Agreements (FFAs)

A Forward Freight Agreement is essentially a forward contract on freight. Adam Sonin (2005) defines an FFA as an agreement to pay the difference between a price agreed today and the future price of moving a product from one location to another, or for the future price of hiring a ship over a period of time. The FFAs are purely financial agreements and do not involve any actual freight or ships. FFAs and related instruments serve two important purposes. First and foremost, they are instruments to hedge exposure to freight market risk. Second, if the market

is efficient and transparent, they reveal additional information on the future direction of the spot freight rates. This is called the price discovery function.

FFAs were pioneered by Clarksons Securities Limited in 1991 and are traded over-the-counter. They are tailor-made and flexible regarding cargo size, settlement dates and as to which index or assessment the position is closed out against. In addition, one can enter an FFA very quickly, and the cost is relatively low compared to operating in the spot market with actual vessels. An FFA is a so called Contract-for-Differences (CFD), meaning that settlement is made in cash on the difference between the contract price and a settlement price. The settlement price is, for spot routes, usually the average of the route over the last seven days of a month. For time charter routes the settlement price is most often the average hire-rate over a month. Dry bulk FFAs are mostly traded in units of USD per day to reflect time charter hire of dry cargo ships. One lot is then one day of time charter income. Contracts are traded as blocks of time as months, quarters or calendar years.

Users of FFAs include shipowners who sell contracts to hedge against falling freight rates, charterers who buy contracts to fix shipping costs, arbitrageurs and volatility speculators (Sonin, 2005). All users benefit from transparent pricing and from the price discovery function of the forward rates (Kavussanos and Visvikis, 2006a).

3.3.1 Hybrid FFAs

A hybrid FFA is an FFA that is cleared through a clearinghouse. As of today, only IMAREX/NOS, LCH.Clearnet, NYMEX Clearport and SGX offer this service – probably with others to come.

3.4 Freight Futures

Freight futures are standardized and exchange traded FFAs. Today, freight futures offer on-screen trading with instant straight-through clearing. This again leads to transparent and more efficient pricing. Traders of freight futures can be anonymous. In addition, there are low barriers to entry compared to the FFA market (Bøe, 2005).

The first attempt at introducing freight futures was made by The Baltic International Freight Futures Exchange in 1985 when they opened trading of the BIFFEX. The BIFFEX was a dry bulk freight futures contract and was initially settled on The Baltic Freight Index (BFI). Due to high basis risk, the hedging effectiveness of the BIFFEX was low compared to evidence from other commodity and financial futures markets. In a study by Kavussanos and Nomikos (2000b), the hedging effectiveness of the BIFFEX was calculated to be in the 4,0% to 19,2% range. This resulted in low trading volumes of the contract. The composition of the index underlying the BIFFEX was changed several times to reduce basis risk, but this did not improve trading volumes by much. With the introduction of the more popular FFA contracts, trading volumes declined even further, and trading of the BIFFEX ceased in April 2002. In Figure 3.1, we show the yearly number of BIFFEX contracts traded from the beginning in 1985 until the end in 2002. In 1988, the year with the highest trading volume of the BIFFEX, about 400 contracts were traded each trading day.

Figure 3.1: Yearly Number of Traded BIFFEX Contracts (1985:05-2002:04)

Source: LIFFE
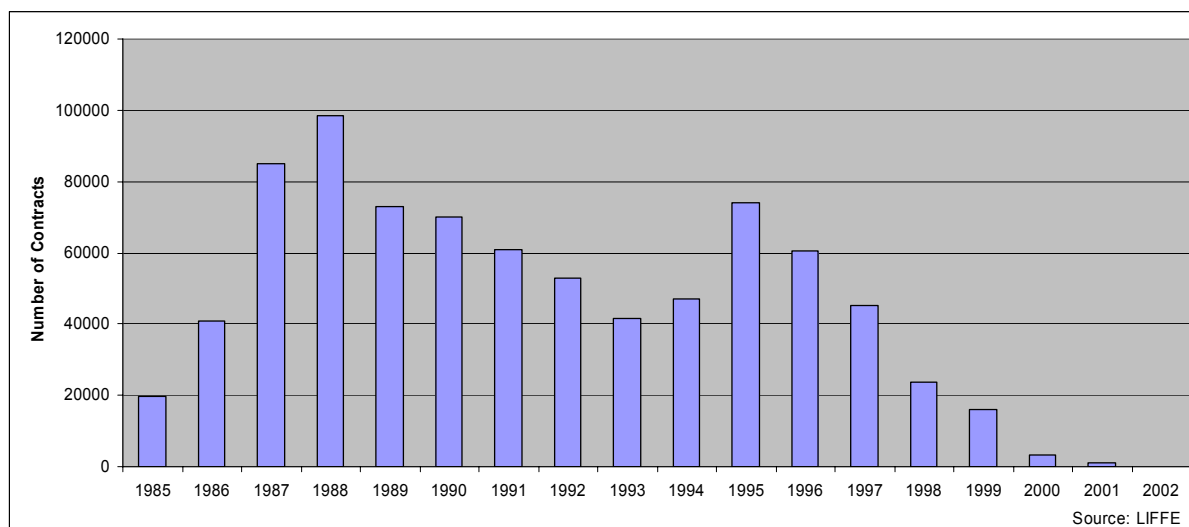
IMAREX has now continued the mission of providing freight futures and other shipping derivatives to the shipping business.[7] Most important for our study are the dry bulk freight futures. In dry bulk, IMAREX offers freight futures contracts on several categories of ships and routes, with the PM4TC contract being the most liquid. The PM4TC freight futures contract is based on the average of the four Panamax time charter routes: P1A, P2A, P3A and

---

[7] The only other supplier of freight futures is NYMEX, which at the present time only offers trading in tanker freight futures.

P4.[8] One lot of the PM4TC is equivalent to one day of time charter income. In Figure 3.2, we show the quarterly number of lots traded of IMAREX dry bulk freight futures.[9] The total number of lots traded was 95 771 in 2006. This is just below the trading volume the BIFFEX had in its best year.[10] In March 2007, 10 708 dry bulk freight futures were traded on IMAREX. 51% of these were PM4TC contracts (5 428 lots). On average, 247 PM4TC contracts were traded each trading day in March 2007.

Figure 3.2: Quarterly Number of Lots Traded of IMAREX Dry Bulk Freight Futures

Source: IMAREX
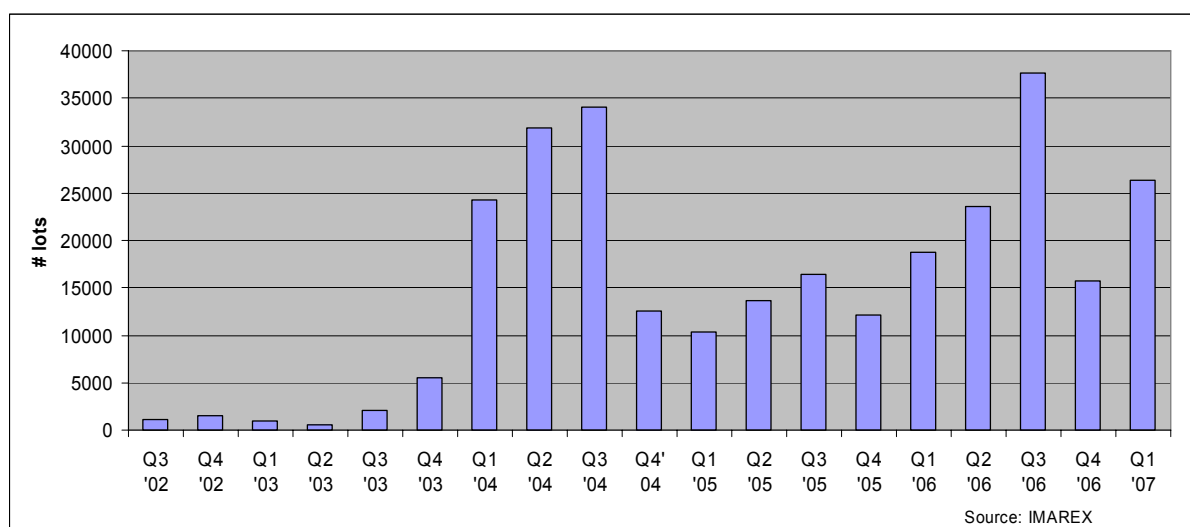
In practice, the difference between freight futures, FFA contracts and especially hybrid FFAs might not be particularly large. For example, OTC brokers trade FFA contracts with very similar specifications to the IMAREX PM4TC contract. This is because the four T/C routes are the most popular and because standardisation might lead to better liquidity and more efficient pricing. If each contract was to be tailor-made, as they indeed could be, it would be very time consuming and complex negotiations would be needed. This standardisation is positive for both FFA contracts and the freight futures as high liquidity facilitates market efficiency and correct pricing. The price of the IMAREX PM4TC freight futures contract and an FFA contract with similar or equal specifications should therefore be almost identical on a given day; if they were not, it would be possible to make riskless profit through arbitrage. A small difference in price might still occur due to the marking-to-market procedure for the

---

[8] See Table 5.1 for a description of the composition of the PM4TC.
[9] See Appendix 1 for information on the volume and number of trades of the IMAREX dry bulk freight futures.
[10] One lot of IMAREX PM4TC is equivalent to one BIFFEX contract.

freight futures and because the clearinghouse demands a risk premium for taking the counterparty risk.

As mentioned earlier, the BIFFEX failed mainly because of high basis risk and low liquidity. We saw that the daily average of IMAREX PM4TC lots traded in March 2007 was lower than the daily average for the BIFFEX in 1988. One might therefore argue that low liquidity could be a problem for IMAREX as well. However, considering the discussion in the last paragraph, IMAREX will free-ride on the liquidity of similar contracts traded in the OTC market.[11] We therefore believe that low liquidity is not going to be a problem for IMAREX.

3.4.1 Pricing of FFAs and Freight Futures – The Unbiasedness Hypothesis

As mentioned earlier, non-arbitrage arguments can not be used to price non-storable commodities. In studies by Kavussanos (2002) and Kavussanos and Visvikis (2004 and 2006a) it is shown that prices of FFAs and freight futures are given by the following relationship:

$$F_{t,T} = E_t(S_T) \tag{3.4}$$

where $F_{t,T}$ is the delivery price at time $t = T$ determined at time t. $E_t(S_T)$ is the market's expectation of the spot rate at time $t = T$ formed at time t. It is assumed that there is no risk premium and that market agents are completely rational and do not make any systematic mistakes (Fama, 1991).

The pricing relationship is called *the unbiasedness hypothesis* and implies that the market is efficient and that forward prices are unbiased estimators of future spot prices. Changes in forward prices for a given date are purely random and are reflecting the release of news. This does not mean that we expect the forward price and the future spot price to equate, but rather that the forward price mirrors the expectations of the future spot price of the average of the market agents.

---

[11] IMAREX had a 7-10% market share of dry bulk derivatives market in 2004 (IMAREX).

In practice, the hypothesis might not always hold. If there is a mismatch between how many hedgers who want long positions in futures contracts and how many hedgers who want short positions, the difference has to be covered by speculators. The speculators will then demand a risk premium to take the shortage position. Hedgers are on the other hand willing to pay the risk premium to reduce their risk. The presence of a risk premium will lead to biased futures prices.

3.4.2 The Price Discovery Role of FFAs and Freight Futures

Because theory states that forward prices serve as unbiased estimators of future spot prices, forward prices should comprise of more information than current spot prices alone. Speculators analyze the futures prices and trade contracts based on their analyses. If they think futures prices are too low, they buy, and if they think prices are too high, they sell. This activity drives the prices in the right direction. Market agents, whether they are actively trading forwards or futures or not, can then use forward prices as information of which direction the spot price will move in the future. This is the price discovery role of the forward and futures contracts.

A factor that facilitates price discovery is the fact that trading in futures contracts is much faster and easier than trading in the spot market by owning an actual vessel. Investors will therefore prefer to speculate on price movements of the freight rates by trading in the futures market. If the futures prices are unbiased – and therefore trading without a risk premium – trading futures instead of trading spot will be even more preferable. Futures prices should as a result process new information more rapidly than spot prices; empirical tests indicate that this is in fact so (Kavussanos and Nomikos, 2001).

3.5 A Review of Prior Studies

In this chapter, we review and refer to studies of the unbiasedness hypothesis and studies of the hedging performance of freight futures and FFAs. The referred studies are mostly on the BIFFEX contract and serve as a reference point when we later are to study these two important functions of the IMAREX PM4TC freight futures contract.

3.5.1 Studies of the Unbiasedness Hypothesis of FFAs and Freight Futures

As information on the future is invaluable in decision-making, several studies have been made to test the unbiasedness hypothesis and the forecasting performance of forward and futures prices on freight. In the following few paragraphs, we briefly summarize some of the findings.

Kavussanos and Nomikos (1999) show that one and two months to maturity BIFFEX contracts are unbiased predictors of future spot rates. They also test three months to maturity contracts, but find only marginal evidence of unbiasedness. Future prices for all maturities are still found to provide better forecasts of future spot prices than forecasts generated from time series models and random walk models.

Haigh (2000) also tests the unbiasedness expectations in the BIFFEX freight futures market. The study uses more observations than Kavussanos and Nomikos (1999), and the tests results should therefore be more powerful. Haigh's results indicate that the BIFFEX futures market is unbiased for current, one, two and quarterly contract horizons.

In 2001, Kavussanos and Nomikos study further the price discovery role of the freight futures market. In line with other studies, they find that some price discovery is accomplished by the BIFFEX. They still argue that because the contract is based on a composite index, liquidity and price discovery suffers. In 1999, The Baltic Panamax Index (BPI) replaced the BFI as the basis for the BIFFEX. Kavussanos and Nomikos then examine the causal relationship between the BIFFEX and the BFI (BPI) over different time periods to determine if the change from BFI to BPI as the underlying index improved the price discovery function of the BIFFEX. The results came back positive.

In Kavussanos et al. (2004), cointegration techniques are used to test for unbiasedness in the FFA market. The study shows that FFA prices of contracts one and two months to maturity are unbiased estimators of future spot rates. For contracts three months to maturity the evidence is mixed; contracts on some routes are unbiased estimators, while contracts on other routes are not.

Batchelor et al. (2007) study different models for forecasting spot and forward prices. They argue that because the FFA market is relatively new and illiquid compared to the financial forward markets, it might not be efficient enough to serve as an unbiased estimator on future spot prices. Their results, however, show that forward prices do help forecast future spot rates and that some degree of speculative efficiency therefore is present.

3.5.2 Hedging Performance of the BIFFEX Contract

In the following paragraphs we will summarize two studies of the hedging performance of the deceased BIFFEX contract. This is done so that we later can compare these findings to the hedging performance of the IMAREX PM4TC contract calculated in this thesis.

Hedging is the most important function of forward and futures contracts. In Kavussanos and Nomikos (2000b) it is referred to studies that find the hedging performance of futures contracts on financial assets and storable commodities to be between 80% and 99%. As we will see, the BIFFEX did not nearly achieve the same extent of variance reduction. It is believed that – due to the extensive composition of the underlying Baltic Freight Index – the poor hedging performance was mainly caused by a large basis risk[12].

Thoung and Visscher were in 1990 the first researchers to study the hedging effectiveness of the BIFFEX contract. Their data sample covers a 29-month period from August 1986 to December 1988. The calculation of hedging effectiveness is done by using a regression model in line with Ederington's Framework from 1979.[13] As we see from Table 3.1, a maximum hedging performance of 33,68% is achieved for Route 2. The poorest performance is achieved for Route 12 with 0,74%.

---

[12] Basis risk is the risk that the two opposite positions in a specific hedge does not perfectly correlate and thus that all price movements are not offset. See Chapter 3.6.2.
[13] See Chapter 4.1 for a presentation of the model.

Table 3.1: Hedging Performance of the BIFFEX

	Route 1	Route 2	Route 3	Route 4	Route 5	Route 6	Route 7	Route 8	Route 9	Route 10	Route 11	Route 12	Route 13
CLRM	29,37 %	33,68 %	17,76 %	3,60 %	9,84 %	4,78 %	17,80 %	4,07 %	21,02 %	14,11 %	10,11 %	0,74 %	2,88 %

Routes are as specified by The Baltic Exchange at the time of study (1990). See the original article for more information.

A two-week hedge period and three-month contract is used.

Green colouring denotes the route with the highest hedge effectiveness.

Red colouring denotes the route with the lowest hedge effectiveness.

Source: Thuong and Visscher (1990)

In 2000(b), Kavussanos and Nomikos review the hedging effectiveness of the BIFFEX contract. The data sample is now considerably larger, covering the period from 23 September 1992 to 31 October 1997. In this period the underlying index changed several times. The details of the index adjustments are covered in the original article. Kavussanos and Nomikos argue that the Classical Linear Regression Model is too simple, due to the possible existence of autocorrelation, time-varying standard deviation and a cointegrating relationship between futures and spot prices.[14] In addition to calculate the hedging effectiveness with a regression model, they therefore employ more advanced models as well. We will explain the used models in detail in Chapter 4. At this point, however, it is sufficient for the reader to observe the hedging performance for the different routes. As we see from Table 3.2, the best hedge is again achieved for Route 2 with 19,20% variance reduction. The poorest hedge is achieved for Route 7 with only 4,00% reduction in variance.

Table 3.2: Hedging Performance of the BIFFEX

	Route 1	Route 1A	Route 2	Route 2A	Route 3	Route 3A	Route 6	Route 7	Route 8	Route 9	Route 10
Naive hedge (1:1)	-22,06 %	-4,47 %	-25,22 %	-11,02 %	-112,40 %	-19,16 %	-304,10 %	-124,20 %	-221,30 %	-12,39 %	-163,30 %
CLRM	17,56 %	15,70 %	19,20 %	16,08 %	13,15 %	15,66 %	4,02 %	3,80 %	4,08 %	15,03 %	3,83 %
VECM	17,12 %	15,38 %	19,04 %	15,94 %	13,12 %	15,64 %	4,00 %	3,79 %	4,08 %	15,00 %	3,77 %
VECM-GARCH	17,52 %	15,39 %	-	-	11,47 %	15,32 %	5,07 %	3,92 %	5,02 %	16,34 %	3,01 %
VECM-GARCH-X	18,96 %	16,85 %	-	-	12,31 %	16,92 %	4,81 %	4,00 %	4,64 %	-	4,21 %

Routes are as specified by The Baltic Exchange. See the article for an overview of the index composition at the time of study.

Yellow colouring denotes the model with greatest variance reduction for the individual route.

Green colouring denotes the route with the highest hedge effectiveness.

Red colouring denotes the route with the lowest hedge effectiveness.

Route 9 has later been renamed to P4.

Routes 1A, 2A, 3A and 9 combined are equivalent to the routes in IMAREX's PM4TC contract.

Source: Kavussanos and Nomikos (2000b)

The index underlying the BIFFEX has, as mentioned, changed several times. In November 1999, the BFI as the underlying index was changed to the BPI. Kavussanos and Nomikos (2000b) then use data from 3 November 1999 to 28 June 2000 to calculate the hedging

---

[14] The concept of cointegration is explained in Chapter 4.2.2.2.

effectiveness after the change. Due to few data points, they only use regression and thus also a constant hedge ratio. Variance reduction and change from the pre-BPI period are shown in Table 3.3.

Table 3.3: Hedging Performance of the BIFFEX with BPI as the Underlying Index

	Route 1	Route 1A	Route 2	Route 2A	Route 3	Route 3A	Route 9	Average
CLRM	18,46 %	24,55 %	30,85 %	35,21 %	35,64 %	39,95 %	33,97 %	31,15 %
Change from Pre BPI Period	-0,50 %	7,70 %	11,65 %	19,13 %	22,49 %	23,03 %	17,03 %	14,36 %

Routes are as specified by The Baltic Exchange. See the article for an overview of the index composition at the time of study.

Green colouring denotes the route with the highest hedge effectiveness.

Red colouring denotes the route with the lowest hedge effectiveness.

Route 9 has later been renamed to P4.

Routes 1A, 2A, 3A and 9 combined are equivalent to the routes in IMAREX's PM4TC contract.

Source: Kavussanos and Nomikos (2000b)

We see from the table that the hedging effectiveness has been greatly improved. For Route 3A, the hedging effectiveness is now 39,95%. This is an improvement of 23,03 percentage points. The average hedging effectiveness for the Panamax routes are now 31,15%.

The most liquid IMAREX dry bulk freight futures contract is the PM4TC. The PM4TC is based on a basket with the routes P1A, P2A, P3A and P4 included (P4 is equivalent to Route 9). As this is even fewer routes than the BPI-basket mentioned above, we would expect even greater correlation between the individual routes and the basket. When later studying the hedging effectiveness of the PM4TC, we therefore expect at least equally good and perhaps better hedging performance than shown in Table 3.3.

3.6 Hedging with Freight Futures

The purpose of hedging is to minimise the price variance of one's portfolio. In this chapter we will first introduce some considerations to have in mind when choosing the best futures contract to hedge a specific spot position. We will then explain the ideas of basis and basis risk.

3.6.1 Finding the Optimal Futures Contract

Consider a shipowner who anticipates a drop in the freight rates. He is long in freight and thus wants to take the opposite position in freight futures to hedge his risk. His challenge is to find

the futures contract that correlates the most with his spot freight rate. Correlation is of essence because an increase in freight rates will then lead to an increase in freight futures prices. As the shipowner is long in freight and short in futures, price fluctuations will offset each other and portfolio variance will be reduced. If the correlation is 100%, the shipowner can establish a perfect hedge and all price movements will cancel out by the two opposite positions.

Finding a futures contract with a perfect correlation might pose some problems. For financial assets, like exchange rates, and homogenous commodities, like gold, futures and spot prices will often have a high correlation. For freight, which is a service or a non-storable commodity, it is more difficult to find a highly correlating futures contract. This is because the underlying index or route(s) are not necessarily identical to the shipowner's actual spot position. He might for example have a different ship than the standard vessel that the index is based upon, and/or his actual route might differ more or less from the route specified in the futures contract.

In addition to finding the futures contract with the most suitable underlying index or route(s), one has to choose the optimal expiration date of the specific contract. The best case scenario is when the delivery of the freight service and expiration of the freight futures contract are on the exact same day. This is most often not the case. In general, the basis (defined in Chapter 3.6.2) increases as the time difference between hedge expiration and delivery month increases. In addition, futures with a short time to maturity are more correlated with the spot. This is because unexpected changes to the spot have a greater effect on one's expectations in the near future than in the more distant future. The best approach is therefore to identify the futures contract with expiration closest to, but later than, delivery of the freight service. When time of freight service delivery comes, the shipowner then reverses his freight futures position. Challenges to finding the optimal expiration date might arise if one has not yet decided when to sell or purchase the asset to be hedged.

3.6.2 The Basis

The problems mentioned in the prior few paragraphs give rise to what is called basis risk. Hull (2003) defines the basis as:

Basis = Spot price of asset to be hedged – Futures price of contract used

Using the notation from Chapter 3.1, we have that:

$$b_t = S_t - F_{t,T}$$ (3.5)

where b_t is the basis at time t. The basis can initially be positive or negative. For storable commodities, this is due to carrying costs. If the asset and the underlying asset are the same, the basis should be zero at expiration, or else there would be an arbitrage opportunity. If there is at least some correlation, basis should decline when approaching expiration in time. For freight futures with perfect correlation with the spot freight rate, a positive basis means that the spot price is expected to fall, whereas a negative basis means that the spot price is expected to rise.

The basis can further be divided into two components as follows:

Basis = Time basis + Cross hedging basis

or

$$b_t = (S_t^* - F_{t,T}) + (S_t - S_t^*)$$ (3.6)

where S_t is the spot price of the asset underlying the futures contract and S_t^* is the spot price of the asset being hedged. When these two differ ($S_t \neq S_t^*$), we have a so called cross hedge. Time basis, $(S_t^* - F_{t,T})$, is the basis that would exist if the asset being hedged were exactly the same as the asset underlying the futures contract. The cross hedging basis, $(S_t - S_t^*)$, is then the extra basis that would arise if the two assets were not exactly the same. When cross-hedging, the basis does not normally equal zero at expiration.

A simple example on a cross hedge can be a coffee producer from Kenya who uses coffee futures specified on Brazilian coffee to hedge his future spot price. Coffee from the two countries are not perfect substitutes for each other, but a considerable price correlation is expected. Thus, the Brazilian contract can also be used to hedge fluctuations in Kenyan coffee prices. We would have an even more obvious cross hedge if an apple farmer were to use futures contracts on oranges to hedge his price risk. The correlation might not be expected to

be as high in this case as in the coffee example, but, as long as there is some correlation, a cross hedge can be made.

3.6.2.1 Basis Risk

We then come to the concept of basis risk. To achieve a perfect hedge, it is required that the portfolio basis at expiration is zero. The risk that this is not the case is called basis risk. As mentioned earlier, the basis could be different from zero at expiration of the hedge if, for example, the asset that is to be hedged and the underlying asset are not exactly the same, the hedger does not know exactly when the asset in question is to be purchased or sold, or if the hedge requires the futures contract to be closed out before maturity (Hull, 2003).

Cross hedges with considerable basis risk are usual in shipping. This is because the specific route and vessel a shipowner might want to hedge could differ significantly from the standard vessel and route(s) defined in the contract. Even if the routes were the same, each voyage would be unique in terms of cargo, loading and discharging ports, and loading date flexibility.

Finding the optimal futures contract to use as a hedging vehicle is thus a matter of finding the futures contract that gives the least basis risk when put in a portfolio together with the spot. When the best futures contract has been identified, the challenge is then to find the number of futures per unit of exposure in the spot market that minimises total return variance. This relationship is called the optimal hedge ratio. In the next chapter, we will present models to calculate the optimal hedge ratio.

4. Models for Calculating the Optimal Hedge Ratio

In traditional hedging it was assumed that the best hedge was achieved by taking a futures position equal in magnitude but of opposite sign to the position in the spot market. This is called a naive hedge or a 1:1 hedge. Spot risk is then entirely exchanged for basis risk. Later, it has been shown that a naive hedge most often does not give the greatest risk reduction. Researchers debated the issue extensively, and numerous hypotheses were put forth. When the modern portfolio theory was introduced by Markowitz in 1952, it was not long before the theory was applied to hedging as well. Several studies were made on the subject, but the breakthrough came when Louis H. Ederington in 1979 presented his framework for hedging. In the next chapters we will first present Ederington's model and hedging with a constant hedge ratio. We will then establish the need for time-varying hedge ratios and present models that deal with that specific issue. We will also present a measurement of the hedging effectiveness. When later using the different hedging models on actual data, we will compare them to see which one achieves the best hedging performance.

4.1 Ederington's Framework (1979)

In modern portfolio theory, the amount of the different assets in a portfolio is chosen so that the variance of the return is minimised. In our example with the shipowner who is long in freight and wants to use futures contracts to hedge his exposure, his return on the hedging portfolio between time t and time t – 1, $\Delta R_{H_t}(= R_{H_t} - R_{H_{t-1}})$, can be written as:[15]

$$\Delta R_{H_t} = (S_t - S_{t-1})X_S + (F_{t,T} - F_{t-1,T})X_F \tag{4.1}$$

where X_S and X_F represent the spot and futures market holdings. The spot market holdings, X_S, are viewed as fixed and the challenge is therefore to find the futures market holdings that minimise total variance as given by:

$$Var(\Delta R_H) = (X_S^2 \sigma_{\Delta S}^2) + (X_F^2 \sigma_{\Delta F}^2) + (2X_S X_F \sigma_{\Delta S, \Delta F}) \tag{4.2}$$

---

[15] This presentation follows the framework by Ederington (1979). As we will be studying the hedging effectiveness a posteriori, we have left expectations out of the formulas.

where $\sigma^2_{\Delta S}$, $\sigma^2_{\Delta F}$ and $\sigma_{\Delta S, \Delta F}$ represent the variances and the covariances of the price changes from time t = 0 to time t = 1.

The amount of futures market holdings divided by the amount of spot market holdings is called the hedge ratio. Algebraically, the hedge ratio is defined as h = X_F/X_S. The optimal hedge ratio, h*, is then the ratio that minimises the variance of the hedging portfolio.

Deriving the Optimal Hedge Ratio

To derive the optimal hedge ratio, h*, we first write Equation (4.1) in terms of price changes per unit of the underlying commodity:[16]

$$\frac{\Delta R_{H_t}}{X_S} = (S_t - S_{t-1}) + (F_{t,T} - F_{t-1,T})\frac{X_F}{X_S} \tag{4.3}$$

We define $(S_t - S_{t-1})$ as ΔS_t and $(F_{t,T} - F_{t-1,T})$ as ΔF_t and substitutes into (4.3):

$$\frac{\Delta R_{H_t}}{X_S} = \Delta S_t + h\Delta F_t \tag{4.4}$$

A hedger is concerned with minimising risk. We use Equation (4.4) to write the variance of the hedge portfolio return per unit of the commodity, $\sigma^2_H (= Var(\Delta R_{H_t} / X_S)$:

$$\sigma^2_H = \sigma^2_{\Delta S} + (h^2 \sigma^2_{\Delta F}) + (2h\sigma_{\Delta S, \Delta F}) \tag{4.5}$$

The value of h that minimises $\sigma_H{}^2$ is then found by taking the derivative of (4.5) with respect to h and setting it equal to zero:

$$\frac{d\sigma^2_H}{dh} = 2h\sigma^2_{\Delta F} + 2\sigma_{\Delta S, \Delta F} = 0$$

---

[16] Derivation of the optimal hedge ratio follows Stoll and Whaley (1993, p. 54), but with a slightly different notation and without considering the uncertainty of S_T and $F_{T,T}$ at time t < T.

When solving for h, we get the optimal hedge ratio, h*:

$$h^* = -\frac{\sigma_{\Delta S, \Delta F}}{\sigma_{\Delta F}^2}$$

(4.6)

We see from Equation (4.6) that the optimal hedge ratio depends on the covariance between the spot and futures price changes relative to the variance of the futures price change.

The optimal hedge ratio derived here considers a portfolio with only two assets: spot and futures. It is then implicitly assumed that the variance of this portfolio is independent of other risky assets, which the investor might have. In reality this is most often not the case. According to portfolio theory, one should then consider the investor's total portfolio[17] when minimising variance. When doing so, the optimal hedge ratio is likely to be different from what would be found using Equation (4.6) (Bond and Thompson, 1986).

Haigh and Holt (2002) study the hedging effectiveness of freight futures when hedging freight, exchange rates and commodity prices at the same time. They find that the contribution of the freight futures to risk reduction is reduced when used together with the other hedging instruments. This gives support to the portfolio theory. Ideally, one should then consider one's entire portfolio when calculating the optimal hedge ratio. As this would require an enormous number of calculations, we will in this thesis consider a portfolio consisting of only spot market holdings of an asset and futures contracts of the same asset.

4.1.1 Hedging Effectiveness

When employing the optimal hedge ratio in an actual hedge, it is interesting to see how effective the specific hedge is. We measure hedging effectiveness by how much of the variance that is eliminated by the hedge. The variance of the unhedged return per unit of the commodity is given by Var(U), whereas the variance of the risk minimising portfolio of spot and futures per unit of the commodity is given by σ_H^2. Hedging effectiveness, e, is then given by:

---

[17] According to theory, everyone should hold a very small piece of the market portfolio.

$$e = 1 - \frac{\sigma_H^2}{Var(U)} \tag{4.7}$$

$Var(U)$ is the same as Equation (4.5) with h = 0 $(Var(U) = \sigma_{\Delta S}^2)$. We substitute (4.5) into (4.7) and get:

$$e = 1 - \frac{\sigma_{\Delta S}^2 + h^2 \sigma_{\Delta F}^2 + 2h\sigma_{\Delta S,\Delta F}}{\sigma_{\Delta S}^2} = \frac{\sigma_{\Delta S}^2 - \sigma_{\Delta S}^2 - h^2 \sigma_{\Delta F}^2 - 2h\sigma_{\Delta S,\Delta F}}{\sigma_{\Delta S}^2} = \frac{-h^2 \sigma_{\Delta F}^2 - 2h\sigma_{\Delta S,\Delta F}}{\sigma_{\Delta S}^2}$$

We then substitute Equation (4.6) into the previous equation and get:

$$e = \frac{-(-\frac{\sigma_{\Delta S,\Delta F}}{\sigma_{\Delta F}^2})^2 \sigma_{\Delta F}^2 - 2(-\frac{\sigma_{\Delta S,\Delta F}}{\sigma_{\Delta F}^2})\sigma_{\Delta S,\Delta F}}{\sigma_{\Delta S}^2} = \frac{\frac{\sigma_{\Delta S,\Delta F}^2}{\sigma_{\Delta F}^2}}{\sigma_{\Delta S}^2}$$

The final expression for the hedging effectiveness is thus given by:

$$e = \frac{\sigma_{\Delta S,\Delta F}^2}{\sigma_{\Delta F}^2 \sigma_{\Delta S}^2} = \rho^2 \tag{4.8}$$

We see that the hedging effectiveness is given by the squared coefficient of correlation between the change in the spot price and the change in the future's price.

4.1.2 Finding the Optimal Hedge Ratio with a Regression Model

In practice, the optimal hedge ratio can also be found with a Classical Linear Regression Model.[18] Consider the following equation:

$$\Delta S_t = \alpha_0 + \alpha_1 \Delta F_t + u_t \; ; \; u_t \sim iid(0, \sigma^2) \tag{4.9}$$

---

[18] See Brooks (2002) for a presentation of regression theory.

This is a simple regression with an intercept term, α_0, a slope coefficient, α_1, and an error term, u_t. The errors are identically and independently distributed as specified. The regression thus assumes that changes in the spot price are proportional to changes in the futures price, but with a random element included.

To derive the optimal hedge ratio by using a regression model, we put Equation (4.9) into (4.4) as follows:

$$\frac{\Delta R_{H_t}}{X_S} = \Delta S_t + h\Delta F_t = \alpha_0 + \alpha_1 \Delta F_t + u_t + h\Delta F_t = \alpha_0 + (\alpha_1 + h)\Delta F_t + u_t \tag{4.10}$$

The variance of the risk minimising portfolio of spot and futures per unit of the commodity can be expressed as:

$$\sigma_H^2 = \alpha_1^2 \sigma_{\Delta F}^2 + h^2 \sigma_{\Delta F}^2 + 2h\alpha_1 \sigma_{\Delta F}^2 + \sigma_u^2 \tag{4.11}$$

The value of h that minimises σ_H^2 is then found by taking the derivative of Equation (4.11) with respect to h and setting it equal to zero:

$$\frac{d\sigma_H^2}{dh} = 2\alpha_1 \sigma_{\Delta F}^2 + 2h\sigma_{\Delta F}^2 = 0$$

We solve for h and get the optimal hedge ratio:

$$h* = -\frac{2\alpha_1 \sigma_{\Delta F}^2}{2\sigma_{\Delta F}^2} = -\alpha_1 \tag{4.12}$$

The optimal hedge ratio is thus the negative of the slope coefficient, α_1, from Equation (4.9). If, for example, $\alpha_1 = 1$, the optimal hedge ratio, h*, is - 1. In that case, we short one futures contract for every spot market holding.

4.1.2.1 The CLRM and Hedging Effectiveness

We find the hedging effectiveness in the CLRM approach as the coefficient of determination, R^2. R^2 is defined as the proportion of variability in a data set that is accounted for by the regression model. We derive this result as follows:

The hedging effectiveness, e, is still given by:

$$e = 1 - \frac{\sigma_H^2}{Var(U)}$$

Var(U) is $\sigma_{\Delta S}^2$ as before and σ_H^2 is given by Equation (4.11). The hedging effectiveness can then be written as:

$$e = 1 - \frac{\alpha_1^2 \sigma_{\Delta F}^2 + h^2 \sigma_{\Delta F}^2 + 2h\alpha_1 \sigma_{\Delta F}^2 + \sigma_u^2}{\sigma_{\Delta S}^2}$$

From Equation (4.12) we have that $h^* = -\alpha_1$. We insert into the previous equation and get:

$$e = 1 - \frac{\alpha_1^2 \sigma_{\Delta F}^2 + (-\alpha_1)^2 \sigma_{\Delta F}^2 + 2(-\alpha_1)\alpha_1 \sigma_{\Delta F}^2 + \sigma_u^2}{\sigma_{\Delta S}^2}$$

This gives us the following relation for the hedging effectiveness:

$$e = 1 - \frac{\sigma_u^2}{\sigma_{\Delta S}^2} = R^2 \tag{4.13}$$

When using a regression model to hedge, we thus find the optimal hedge ratio as the negative of the slope coefficient, α_1, and the hedging effectiveness as the coefficient of determination, R^2.

4.1.3 Assumptions Underlying the Regression Model

The Classical Linear Regression Model is based on five assumptions concerning the disturbance term, u_t, in Equation (4.9). If these five assumptions hold, the estimators α_0 and α_1 will have a number of desirable properties. First and foremost, the estimators will be the Best Linear Unbiased Estimators (BLUE) of their true values. An estimator is best if it has the minimum variance among the class of estimators. It is unbiased if the estimated coefficient values will be equal to their true values on average. In addition, if all the assumptions hold, it will be possible to make inferences about the actual values of the coefficients by using the estimated ones. We will now give a short presentation of the five assumptions.

Assumption 1: The Errors Have Zero Mean
This assumption can be written as:

$$\mathrm{E}(u_t) = 0 \tag{4.14}$$

If this assumption does not hold, the value of R^2, and thus our estimate of the hedging effectiveness, can be affected. In addition, it can lead to seriously biases in the slope coefficient estimates. Per definition, this assumption will never be violated if a constant term is included in the regression. In Equation (4.9), α_0 is the constant term, so this assumption is satisfied in our case.

Assumption 2: The Errors Have a Constant and Finite Variance Over All Values of x_t
This is known as the assumption of homoscedasticity. Technically, it is written as:

$$Var(u_t) = \sigma^2 < \infty \tag{4.15}$$

If the variance of the errors is not constant, they are so called heteroscedastic. There are two forms of heteroscedasticity. The general form of heteroscedasticity is when the variance of the error term depends on one or more of the explanatory variables. The other form is when the variance of the error term systematically varies over time. We then have so called ARCH-effects. If heteroscedasticity is present and not corrected for, the estimators will still be biased

and consistent[19], but they will not be BLUE. The estimators in the model will no longer have the minimum variance among all possible estimators. Thus, the standard errors of the coefficients are not correct and inferences can be misguiding.

White's test (White, 1980) can be used to test for general heteroscedasticity. The null hypothesis of the test assumes that the errors are independent of the explanatory variables, that they are homoscedastic and that the linear model is correctly specified. The test statistic is chi-square distributed with the number of degrees of freedom equal to the number of slope coefficients in the test regression.

A remedy by White (1980) can be used to correct for general heteroscedasticity in the regression: if the residuals of the estimated equation are positively related to the square of an explanatory variable, the standard errors of the slope coefficient will then be increased relatively to the ordinary regression standard errors. This makes inferences more conservative, and more evidence is needed to reject the null hypothesis of no explanatory power of the specific variable.

A more specific type of heteroscedasticity is the Autoregressive Conditional Heteroscedasticity (ARCH). If ARCH is present, the variance of the errors in a period will depend on the variance in prior periods. This does not in itself invalidate inferences about the coefficient estimates from the regression, but efficiency might be lost in the model. ARCH-effects are very common in financial time series. Engle's ARCH LM test (Engle, 1982) can be used to test for ARCH-effects.[20] The null hypothesis in the ARCH LM test is that there are no ARCH-effects up to the z'th lag. The test statistic is chi-square distributed with z degrees of freedom.

---

[19] When an estimator is consistent, it will converge to its true value as the sample size increases to infinity. See Brooks (2002) Chapter 3.6.1.

[20] LM is the Lagrange Multiplier. See Hamilton (1994) for more information.

Assumption 3: The Errors are Statistically Independent of One Another

This is the assumption of no autocorrelation and can be written as:

$$Cov(u_i, u_j) = 0 \qquad (4.16)$$

If a regression is still carried out when the assumption is violated, the coefficient estimates will still be unbiased, but they will not be BLUE. In addition, the coefficient of determination can be inflated if positive autocorrelation is present.

To test for autocorrelation, the Breusch-Godfrey test (Breusch, 1979 and Godfrey, 1978) can be employed. This is an LM test where the null hypothesis is that no autocorrelation is present up to a number of p lags. The LM test statistic is chi-square distributed with the number of degrees of freedom equal to p. The choice of the number of lags to test for could be based on specific assumptions regarding the nature of the autocorrelation. If, for example, one uses monthly data, twelve lags could be tested for. In practice, it is usually tested for several values of p. If no autocorrelation is present, the null hypothesis should not be rejected for any value of p.

Assumption 4: There is No Relationship between the Residuals and the Explanatory Variables

This assumption can be written as:

$$Cov(u_t, x_t) = 0 \qquad (4.17)$$

In the CLRM, it is assumed that the explanatory variables are non-stochastic. It is thus assumed that the regressors explain the regressand and not the other way around. If this assumption does not hold, the estimators will still be unbiased and consistent if the regressors are not correlated to the error terms, i.e. if Equation (4.17) holds.

Assumption 5: The Errors are Normally Distributed

This assumption can be written as:

$$u_t \sim \text{N}\left(0, \sigma^2\right)$$

The assumption of normality is necessary to perform joint and single hypothesis tests of the model parameters. For large sample sizes, deviation from normality has very little effect on inferences. The Central Limit Theorem[21] makes the test statistics asymptotically follow the appropriate distribution, even without normally distributed error terms. Jarque-Bera's test can be used to test for normality (see Chapter 5.1).

4.1.4 Correcting for Autocorrelation in a Regression

Herbst et al. (1989) argue that autocorrelation often is present when estimating the optimal hedge ratio with Ederington's Framework. If the model is estimated without correcting for autocorrelation, the standard errors of the coefficients in the estimation can be wrong. In addition, the coefficient of determination can be inflated and thus also the estimated hedging performance.

The problem of autocorrelation can be remedied, however. Autocorrelation in the residuals of an estimated model implies that a richer structure is present in the dependent variable and that more information is available in the sample about that particular structure, than has already been captured by the model. If autocorrelation is present when estimating optimal hedge ratios, the richer structure might be captured by including lags of the error term in the regression equation. The model is then written as follows:

$$\Delta S_t = \alpha_0 + \alpha_1 \Delta F_t + u_t$$

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \ldots + \phi_p u_{t-p} + \varepsilon_t; \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \tag{4.18}$$

---

[21] The Central Limit Theorem states that a sum of *n* independent and identically distributed stochastic variables will be approximately normally distributed, independently of the distribution of the population, as long as *n* is sufficiently large and the sum of the variables has a finite variance.

where ϕ_i (i = 1,2,…,p) are parameters and ε_t is an error term. The remaining notation is as explained for Equation (4.9). This model is then a linear regression model with AR(p) errors, where AR means Auto Regressive. We will call this model LRM AR(p), for short.

The optimal number of lags to include in the model depends on the form of the autocorrelation. In practice, a so called information criterion is often used. Schwarz Information Criterion is for example defined as:

$$SIC = -2\frac{L}{T} + k\frac{\ln(T)}{T} \qquad (4.19)$$

where L is the value of the log-likelihood function with k parameters estimated using T observations. We see that the criterion penalises an increase in the number of parameters (lags). The model with the number of lags that minimises the value of SIC is then selected.

Correcting for autocorrelation in this way will most likely improve the hedging effectiveness in out-of-sample hedging. This is because the regression model with AR(p) errors is better able to capture the information present in the data. In-sample, however, the standard CLRM hedge ratio will always give the best hedging performance among all constant hedge ratio models.

4.2 Hedging with a VAR model and a VECM

Hedging with Ederington's Framework, as presented in Chapter 4.1, has been greatly debated in the literature. In addition to problems with autocorrelation, the possible presence of cointegration of spot and futures prices might lead to further misspecifications of the model. Gosh (1993a and 1993b) find empirically that the estimate of the optimal hedge ratio is biased downwards if spot and futures prices are cointegrated, but a model without an error correction term is still specified. This result is also supported by Brenner and Kroner (1995). We will now present models that try to remedy the mentioned problems.

4.2.1 A VAR Model for Finding the Optimal Hedge Ratio

If autocorrelation is present when estimating optimal hedge ratios, the richer structure might be captured by the following bivariate Vector Autoregressive (VAR) model:

$$\Delta S_t = c_s + \sum_{i=1}^{k} \beta_{si} \Delta S_{t-i} + \sum_{i=1}^{k} \theta_{si} \Delta F_{t-i,T} + u_{st} ; \qquad u_{st} \sim iidN(0, \sigma_u^2)$$

$$\Delta F_{t,T} = c_f + \sum_{i=1}^{k} \beta_{fi} \Delta S_{t-i} + \sum_{i=1}^{k} \theta_{fi} \Delta F_{t-i,T} + u_{ft} ; \qquad u_{ft} \sim iidN(0, \sigma_u^2) \qquad (4.20)$$

where c is the intercept term and β_s, β_f, θ_s and θ_f are parameters. u_{st} and u_{ft} are the error terms.

In the VAR model all of the variables are endogenous. This is necessary because the two equations in the model are estimated simultaneously. This fact makes the VAR model less restrictive than the regression models presented earlier.

The VAR model seeks to remedy the problem of autocorrelation by allowing the value of a variable to depend on its own lags and the lags of the other variables. This implies that the model might be able to capture more features of the data and thus offers a richer structure.

When using a VAR model, the optimal hedge ratio can be derived as:

$$h* = -\frac{Cov(u_{st}, u_{ft})}{Var(u_{ft})} \qquad (4.21)$$

After having found the optimal hedge ratio, the hedging effectiveness is calculated by using Equation (4.7).

4.2.2 The Vector Error Correction Model

4.2.2.1 The Concept of Stationarity

Financial time series very often show properties of nonstationarity. To explain this term, we start the other way around and define stationarity. A data series is said to be weakly stationary

if its mean, variance and autocovariances are constant for each given lag. If they are not, the series is nonstationary. When a system is nonstationary, shocks to the system will not die away over time, but will persist and never disappear.

A regression on nonstationary time series can lead to spurious results. The regression can then have significant coefficient estimates and a high coefficient of determination (R^2), even when the two variables in the regression are unrelated. In addition, it can be proved that the standard assumptions for asymptotic analysis will not be valid. A standard remedy for the presence of stationarity is to difference the time series. If a series is nonstationary, but becomes stationary after differencing once, it is said to have *one unit root*. The time series is then I(1). In general, if a time series has to be differenced *n* times to become stationary, the time series has *n* unit roots and is I(n).

Most financial time series have one unit root and are stationary after being differenced once (Brooks, 2002). As we see from the Equations (4.9), (4.18) and (4.20), both the spot and the futures prices are differenced one time. Therefore, spurious results from a regression will probably not be a problem, but this has to be tested. The test for stationarity will be explained briefly in Chapter 5.1.

When specifying a model in first differences, however, any long-run properties of the time series will be lost (Engle and Granger, 1987). The model will have no long-run solution and the series could wander apart without bound.[22] The VAR model specified in Equation (4.20) is in first differences and it therefore only captures the short-term properties of the spot and futures prices. To account for a possible long-run solution, we have to check for cointegration of the variables and include a correction term in the model.

4.2.2.2 Cointegration and the Error Correction Term

The concept of cointegration can be explained as follows. If two or more series are themselves nonstationary, I(1), but a linear combination of them is stationary, I(0), then the series are said to be cointegrated. In such cases, there will be a long-run equilibrium of the variables and they would have a constant mean that would be returned to frequently. This

---

[22] A long-run solution to a model is one where $Y^* = Y_t = Y_{t-1} = ... = Y_{t-i}$.

50

long-run property is an effect that we would like to include in our model and that would make the model more flexible and efficient.

Spot and futures prices of storable commodities are expected to be cointegrated because they are prices for the same asset at different points in time. Given the release of new information, the spot and futures prices will be affected in very similar ways. Their long-run relationship will be given by the cost-of-carry (Brooks, 2002).

Brenner and Kroner (1995) show that spot and futures prices of investment assets will be cointegrated if the interest differential is stationary. For storable commodities, they show that spot and futures prices are cointegrated only if the interest rate itself is stationary. As interest rates most often are found to be nonstationary, but stationary after differencing once, empirical studies find evidence of cointegration for investment assets, but most often not for storable commodities.[23]

For freight, which is a non-storable commodity, the spot and futures prices of freight are not linked by the cost-of-carry relationship, but by the unbiasedness hypothesis. Given that this hypothesis holds, we will also find evidence of cointegration in this case. Kavussanos and Nomikos (2000b) find cointegration between the spot assessments and futures prices for most of the routes underlying the BIFFEX, but not for all.

Cointegration can be tested by using Johansen's test (1988). This test will be explained in Chapter 4.2.2.4

To account for cointegration and include the long-run properties of the time series, an Error Correction Term (ECT) can be added to the VAR model. We define the ECT as:

$$Z_{t-1} = S_{t-1} + \beta_1 + \beta_2 F_{t-1,T} \tag{4.22}$$

where Z_{t-1} is a linear combination of the two series, S_{t-1} and $F_{t-1,T}$. β_1 is a constant and β_2 is the cointegrating vector. S_{t-1} and $F_{t-1,T}$ are I(1) and Z_{t-1} is I(0).

---

[23] See for example Baillie and Myers (1991), who do not find cointegration in spot and futures prices for commodities. Kroner and Sultan (1993), however, find cointegration in currency markets, as would be expected.

4.2.2.3 Specifying the VECM

We then specify the new model, a bivariate Vector Error Correction Model (VECM), as follows:

$$\Delta S_t = c_s + \sum_{i=1}^{k} \beta_{si} \Delta S_{t-i} + \sum_{i=1}^{k} \theta_{si} \Delta F_{t-i,T} + \gamma_s Z_{t-1} + u_{st}$$

$$\Delta F_{t,T} = c_f + \sum_{i=1}^{k} \beta_{fi} \Delta S_{t-i} + \sum_{i=1}^{k} \theta_{fi} \Delta F_{t-i,T} + \gamma_f Z_{t-1} + u_{ft} \tag{4.23}$$

where u_{st} and u_{ft} are white-noise disturbance terms, β_s, β_f, θ_s and θ_f are parameters and Z_{t-1} is the ECT, which measures how the dependent variable adjusts to the previous period's deviation from long-run equilibrium. γ is then the coefficient that measures the speed of adjustment to long-run equilibrium. A greater value of γ implies a greater response to the deviation from long-run equilibrium.

The variance minimising hedge ratio can be calculated similarly to Equation (4.21) by using the residuals obtained when estimating Equation (4.23). The hedging effectiveness is calculated by Equation (4.7) as before.

4.2.2.4 Johansen's Test for Cointegration

The question is then how to test for cointegration and choose between using a VAR model or a VECM. One alternative is to use the Engle and Granger procedure (Engle and Granger, 1987). Another alternative is the Johansen test (Johansen, 1988), which has been shown by Gonzalo (1989) to provide superior inference. To test for cointegration, Johansen proposed to specify the following VECM:

$$\Delta X_t = \alpha_1 + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \Pi X_{t-1} + u_t \quad ; \qquad u_t \sim iid(0, \Sigma) \tag{4.24}$$

where α_1 is a constant, $X_t = (S_t \ F_{t,T})'$ is the vector of spot and futures prices and Γ_i and Π are 2 x 2 coefficient matrices measuring the short- and long-run adjustment of the system to changes in X_t. u_t is a 2 x 1 vector of white noise residuals and Σ is a 2 x 2 variance/covariance matrix.

In Johansen's test, five different assumptions can be made regarding the deterministic trend in the time series. The test gives different results depending on which one of the five assumptions that are made. We refer the reader to Asteriou (2006) for an explanation of the five assumptions. As suggested by Johansen (1992), which of the five assumptions or models to choose can be decided by using the Pantula principle.[24] The number of lags to include in the test can be based on Information Criteria, for example SIC.

Having chosen the appropriate model for the test, the Johansen procedure then tests for cointegration by examining the rank of Π. For this, the λ_{max} and λ_{trace} statistics are estimated. The max test, with the corresponding λ_{max} statistic, is a test for rank(Π) = r against the null hypothesis that rank(Π) = r + 1. The trace test then examines the null hypothesis that the number of cointegrating vectors is less than or equal to r, with the alternative hypothesis that the number of cointegrating vectors is greater than r.

Johansen's test subsequently use the information of rank(Π) to conclude on the question of cointegration. If the rank of Π is 0, there are no cointegrating relationships. We will then specify a VAR model in first differences. If the rank of Π is 1, a single cointegrating relationship is present. ΠX_{t-1} is now the error correction term and Π can be factorised into two separate 2 x 1 matrices, Φ and β. We then have that $\Pi = \Phi\beta'$, where Φ is the vector of error correction coefficients measuring the speed of convergence to the long-run steady state and β' represents the vector of cointegrating parameters. For rank(Π) = 1, we will specify a VECM to calculate the hedge ratio. If the rank of Π is 2, all the variables in X_{t-1} are I(0), and we will specify a VAR model in levels.

4.3 Time-Varying Hedge Ratios and ARCH Models

When autocorrelation and cointegration of spot and futures prices are present, the specified VECM model will capture the structure of the time series better than the CLRM first presented. The VECM still only calculates a constant hedge ratio. This fact has been criticised by several researchers. Kroner and Sultan (1993) argue that asset prices are characterised by time-varying distributions and that the optimal hedge ratio should also be time-varying. The riskiness of each of these assets changes as new information is received by the market. This

---

[24] For a simple explanation of the Pantula principle, see Asteriou (2006).

view is also supported by Bollerslev (1990). In Figure 4.1, we see that the variance of the price changes for the P1A route seems to be greater in some periods than in others.[25] This indicates that we should specify a model that captures this information.

Figure 4.1: Weekly Price Changes for the P1A Route

Source: The Baltic Exchange
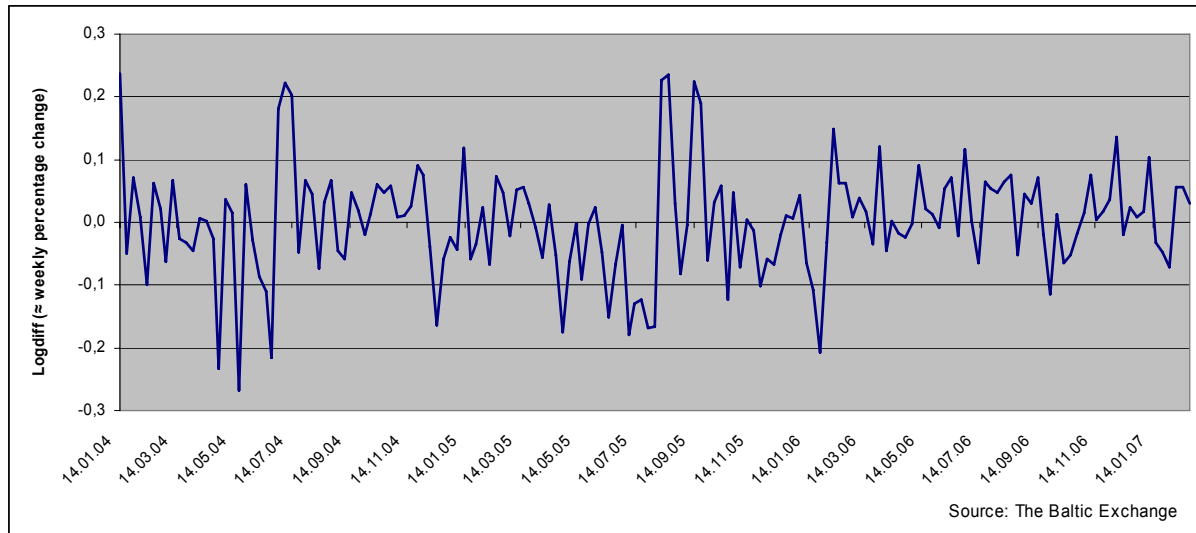
4.3.1 The ARCH-model

To capture the effects of time-varying distribution of the errors in a model, Robert Engle developed the ARCH[26] class of models in 1982. In the simplest case, modelling the second moments of a univariate model, we assume that the conditional variance at time t depends on the squared errors from the preceding *p* periods. The error at time t depends on the information given in the market in the previous period. This can be written as:

$$u_t | \Omega_{t-1} \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + ... + \alpha_p u_{t-p}^2 \qquad (4.25)$$

where α_0 is a constant, u_{t-1}^2 is the squared error from period t-1, Ω_{t-1} is the information in the market at time t-1, σ_t^2 is the conditional variance at time t and α_i (i = 1, 2, 3, …, p) are coefficients.

---

[25] The great volatility of the freight rates were shown and explained extensively in Chapter 2.1.4.
[26] ARCH = autoregressive conditional heteroscedasticity.

The model in Equation (4.25) is called ARCH of order p or ARCH(p). An ARCH(p) representation of the second moments is not very much used in practice. This is mainly due to the large number of parameters that have to be estimated when many lags are included in the variance equation.

When estimating ARCH-models in practice, the residuals from the estimated mean equation are used as input. The mean equation can for example be the regression from Equation (4.9).

4.3.2 The GARCH-model

In 1986, Tim Bollerslev generalized the ARCH(p) model by including q number of past conditional variances in the equation. This model is called GARCH. The variance equation from a univariate GARCH(p,q) model is written as:

$$u_t \big| \Omega_{t-1} \sim N(0, \sigma_t^2)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + ... + \alpha_p u_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + ... + \beta_p \sigma_{t-q}^2 \qquad (4.26)$$

where β_i (i = 1, 2, 3, …, q) are coefficients and the other notation is as explained for Equation (4.25). The most commonly used version of the GARCH model is the GARCH(1,1). The GARCH(1,1) is better than an ARCH model as it is more parsimonious and avoids overfitting. In fact, it can be shown that a GARCH(1,1) model, where only three parameters have to be estimated, allows an infinite number of past squared errors to influence the present conditional variance (Brooks, 2002).

4.3.3 BEKK

As both the VAR model and the VECM presented earlier are bivariate, we need to specify a bivariate GARCH model to estimate the conditional moments. In line with previous studies of hedging in the shipping business by Kavussanos and Nomikos (2000a, 2000b and 2000c), we specify the conditional second moments of spot and futures returns as a GARCH(1,1) model with a BEKK (Engle and Kroner, 1995) representation. For both the VAR model from Equation (4.20) and the VECM model from Equation (4.23) the BEKK specification can be written as

$$u_t = \begin{pmatrix} u_{st} \\ u_{ft} \end{pmatrix} \bigg| \Omega_{t-1} \sim distr(0, H_t)$$

$$H_t = C'C + A'u_{t-1}u'_{t-1}A + B'H_{t-1}B \tag{4.27}$$

where C is a 2 x 2 upper triangular matrix, and A and B are 2 x 2 diagonal coefficient matrices. C', A', B' and u_{t-1}' are the transposed of the matrices. In the BEKK representation, the diagonal conditional variances are a function of their own lagged values and their own lagged error terms, while the conditional covariance is a function of lagged covariances and lagged cross-products of the u_{t-1}'. In addition, the formulation guarantees that H_t is positive and allows the conditional covariance to change sign over time.

4.3.4 Lee's GARCH-X model

For a specification with a VECM as the mean equation, Lee (1994) shows that the BEKK model can be augmented to include the squared lagged error correction term (ECT), Z^2_{t-1}, in the conditional variance equation. The idea is that the magnitude of the deviations from long-run equilibrium might affect the conditional variance and that inclusion of the ECT therefore will improve the efficiency of the model. This augmentation of the BEKK model is often called GARCH-X. The variance equation for this model is written as:

$$H_t = C'C + A'u_{t-1}u'_{t-1}A + B'H_{t-1}B + D'Z^2_{t-1}D \tag{4.28}$$

where D is a 1 x 2 vector of coefficients, Z^2_{t-1} is the squared ECT and the other notation is as explained for Equation (4.27).

In the study by Kavussanos and Nomikos (2000b) mentioned earlier, they find that a VECM-GARCH-X model outperforms the regression, VAR and Vector Error Correction models most of the time (see Table 3.2).

4.3.5 Estimation of ARCH/GARCH Models

The ARCH class of models are estimated by using a maximum likelihood technique. Simply put, it involves finding the most likely values of the parameters given the actual data. The

maximum likelihood technique works by first forming a log-likelihood function (LLF). The LLF is then minimised to find the estimate of the parameter values. Minimisation of the LLF can be done manually through derivation, but in practice, econometrics packages like EViews use numerical procedures to minimise the LLF.[27]

4.3.6 Hedge Ratio and Hedging Performance

When using GARCH models to specify the conditional second moments of the time series, the optimal hedge ratio will consequently be time-varying and given by the following formula:

$$
h_t^*\big|\Omega_{t-1} = -\frac{Cov(\Delta S_t, \Delta F_{t,T}\big|\Omega_{t-1})}{Var(\Delta F_{t,T}\big|\Omega_{t-1})}
\tag{4.29}
$$

The return of the time-varying hedge portfolio will then be given as:

$$
r_{H_t} = \Delta S_t + \left(h_t^*\big|\Omega_{t-1}\right)\Delta F_{t,T}
\tag{4.30}
$$

and the hedging performance is calculated as:

$$
e = 1 - \frac{Var(r_{H_t})}{\sigma_{\Delta S}^2}
\tag{4.31}
$$

As reported earlier, Kavussanos and Nomikos (2000a and 2000b) compared the hedging effectiveness of constant hedge ratio models to that of time-varying hedge ratio models. They found the time-varying hedge ratio models to be superior in most cases. Time-varying models are, as previously argued, better able to capture information present in the data than the constant models. We would therefore expect to find that time-varying models are best in our study as well.

---

[27] For more information on the maximum likelihood technique, see Brooks (2002) or Hamilton (1994).

5. Testing the Hedging Performance of the PM4TC

In this section we will perform our empirical analyses on the hedging performance of the IMAREX PM4TC freight futures contract.[28] First, we will describe the properties of the data series of spot and futures prices. Then, we will study the hedging performance by using the models with constant hedge ratio presented in Chapter 4.1 and 4.2. Finally, we will employ the more advanced models presented in Chapter 4.3. To choose the optimal model to specify, we will perform tests for ARCH effects, cointegration and for the optimal number of lags to include in a VAR/VECM model.

5.1 Properties of the Data Series

The IMAREX PM4TC freight futures contract has The Baltic Exchange price assessments of the four Panamax routes P1A, P2A, P3A and P4 as the underlying.[29] We have been supplied with spot price data from The Baltic Exchange for the four routes. IMAREX has supplied us with freight futures prices for the PM4TC contract. The composition of the PM4TC Freight Futures contract is shown in Table 5.1.

Table 5.1: Composition of the IMAREX PM4TC Freight Futures Contract

Route	Vessel size	Route description	Weight
P1A	74.000 MT	Skaw-Gibraltar to Skaw-Gibraltar, Transatlantic round voyage (50-60 days)	25 %
P2A	74.000 MT	Skaw-Gibraltar to Taiwan-Japan, (60-65 days)	25 %
P3A	74.000 MT	Transpacific round voyage, (35-50 days)	25 %
P4	74.000 MT	Delivery Japan-Korea, redelivery Skaw-Gibraltar (50-60 days)	25 %

Source: IMAREX

In our analyses we will use a data sample from Wednesday 7 January 2004 to Wednesday 21 February 2007. Data is available for a longer period of time, but as we can see from Figure 3.2, it was only from the beginning of 2004 that the contracts were getting liquid. Liquidity is a desirable property as it facilitates a more efficient, correct and frequent pricing of the freight

---

[28] See Appendix 2 for a description of the PM4TC contract details.
[29] The standard ship The Baltic Exchange assesses the Panamax routes on is defined as follows: 74,000 mt dwt, "built in first class competitive yard" 89,000 cbm grain, max. loa 225 m, draft 13.95 m, 14.0 knots on 32/28 fuel oil laden/ballast and no diesel at sea. Non coated Not ice classed. 5 years old. Special survey passed. Delivery prompt (2/3 months), charter free. 2% total commission (The Baltic Exchange).

futures. We have therefore chosen to analyse the period from January 2004 and onwards. We will perform in-sample analyses only.

When assessing the hedging effectiveness of the PM4TC with both constant and time-varying hedge ratios, we will assume that a shipowner operates constantly in the spot market on one of the four routes. We then assume that he uses the PM4TC contract to hedge his exposure. In the hedge, he weekly rolls over to the futures contract with the shortest delivery period and that, at the same time, is closest to expiration.

Price changes are the main focus when analysing hedging and finding optimal hedge ratios. We therefore have to decide on the length of the interval to measure the price changes over. Daily, weekly and bi-weekly price differences are most often used. To remove bid/ask problems and day-of-the-week effects, we will use weekly price differences in our analyses.[30] We calculate the price differences from Wednesday to Wednesday to get rid of possible weekend effects. If a specific Wednesday is a holiday, the first observation prior in time will be used in its place. Log-differences are used to find the price changes from week to week.

A problem with the PM4TC contracts is how to splice the data into one continuous time series. As mentioned before, we have used the contracts with the shortest delivery period and that, at the same time, is closest to maturity to calculate the price differences from one week to the next.[31] On Wednesday before a contract's last trading day, we roll the hedge over to the following contract.[32] To remedy the problem of having observations of two different futures prices each time the hedge is rolled over, the futures price series is converted into an index by setting the value of the first observation to 100 and adjusting it by the percentage change of the futures price each week.

In Table 5.2, we show summary statistics for the time series under study. We see that the mean of the price changes is close to zero, but that the standard deviations are very large. For the route P4 we observe a maximum price change in one week of 45,7%. The same route has

---

[30] See Stoll and Whaley (1993) pp. 57-60 for a more detailed discussion on what time interval to use for price differences.
[31] Quarterly contracts are used in the time period before April 2005. IMAREX then started to offer monthly contracts as well. The monthly contracts are therefore used from that point and forward.
[32] From February 2006, the 20th of each month is the last trading day for the monthly contracts. Prior in time, the 15th of each month was the last trading day.

also experienced a 35,9% decline in prices over a week. We see similar tendencies for the other routes, something which supports our presentation of the highly volatile shipping business.

Table 5.2: Summary Statistics

	P1A	P2A	P3A	P4	AVG4TC	IFF PM4TC	Critical values 5 %	1 %
N	163	163	163	163	163	163		
Mean	0,00044	0,000558	-0,001227	-0,001314	-0,000372	0,002092		
Maximum	0,237484	0,268826	0,419213	0,457495	0,307584	0,371064		
Minimum	-0,26803	-0,308963	-0,345201	-0,358861	-0,287798	-0,228259		
Standard Deviation	0,087519	0,085032	0,107436	0,102571	0,088935	0,083958		
Skewness	0,021017	-0,164239	-0,074772	0,305067	0,005733	0,38795		
Kurtosis	4,138268	4,72638	4,839805	5,858843	4,559133	5,251066		
Jarque-Bera	8,811656	20,97461	23,14087	58,03646	16,51072	38,50412	5,99	9,21
ADF in levels	-2,057203 (1)	-2,035091 (1)	-1,919743 (0)	-2,141508 (1)	-2,099818 (1)	-1,252537 (0)	-2,88	-3,47
ADF in 1st differences	-10,47391 (0)	-11,19867 (0)	-9,981985 (1)	-10,53794 (0)	-11,03066 (0)	-13.07265 (0)	-2,88	-3,47

Data are weekly, log differenced prices. (Except for the test statistics for "ADF in levels", where log prices before differencing are used)

AVG4TC is a time series consisting of the average of P1A, P2A, P3A and P4.

IFF is short for IMAREX freight futures.

ADF is the augmented Dickey-Fuller test. We have included an intercept term in the test. The null hypothesis is that a unit root is present. The lag length (in parenthesis) is automatic and based on Schwarz Information Criterion.

Numbers written in red indicate rejection of the null hypothesis.

Skewness is the third moment of the data and measures the asymmetry of the distribution of the series around its mean. Kurtosis is the fourth moment of the data and measures the peakedness or flatness of the distribution of the series. In a symmetric normal distribution, skewness should be zero and kurtosis should be three. The Jarque-Bera (Bera and Jarque, 1980) test for normality measures the difference of the skewness and kurtosis of the series with those from the normal distribution. The null hypothesis of the test is that the data is symmetrically and normally distributed, and the Jarque-Bera test statistic is chi-square distributed with two degrees of freedom. We reject the null hypothesis if the value of the test statistic is above the critical value for the test.[33] We see from Table 5.2 that none of the time series are normally distributed. All of the routes have a kurtosis greater than three. This means that the data is more peaked (leptokurtic) than the normal distribution.

To test for stationarity of the time series, we use the Augmented Dickey Fuller Test (Dickey and Fuller, 1979 and 1981). The null hypothesis in the ADF test is that the data series has a unit root.[34] In Table 5.2, we have first shown the ADF test statistic for the data in levels (on

---

[33] See for example Brooks (2002) Chapter 4.9.1 for more information on the Jarque-Bera test statistic.

[34] A series has a unit root if shocks to the system persist and never die away. In the regression $y(t) = \alpha_1 y_{t-1} + ut$, we have a unit root if $\alpha_1 = 1$. The system is then nonstationary. See Brooks (2002) Chapter 7.1 for more information on stationarity and unit roots.

log form, but before differencing). As we can see, the data series are all nonstationary. A standard remedy for this is to difference the time series once and test again. The test statistic for this is also shown in Table 5.2, and we see that the data has become stationary after differencing once. The data is then said to be I(1). Considering the problem of stationarity, inferences will be valid when using weekly log differences in a regression.

5.1.1 Expectations of the Hedging Effectiveness

As discussed earlier, it is the correlation between the spot price and the hedging instrument that is important. To examine the potential for hedging with the IMAREX PM4TC freight futures contract, we have calculated the coefficient of correlation between the weekly spot price changes and the weekly futures price changes over the sample. The results are shown in Table 5.3.

Table 5.3: Correlation between the T/C Routes and the IMAREX PM4TC Contract

	P1A	P2A	P3A	P4	AVG4TC
IFF PM4TC	0,5625	0,5548	0,5402	0,5261	0,5783

We see that the correlation is just above 0,5 for all routes and for the average of the routes. The correlation is highest for route P1A and we therefore expect the best hedging performance for this route. The correlation is even greater for the average of the routes (AVG4TC), but not by much. We had initially expected that this correlation would be much higher, as the IFF PM4TC contract is settled on the average of the four Panamax time charter routes. When we only get a correlation of 0,5783, it might be because the IFF PM4TC is settled on the average of the daily AVG4TC assessments over the settlement month. Weekly futures price changes should thus be smoother than weekly spot price changes.

In Figure 5.1, we have shown the movements of the P1A price assessment and the IFF PM4TC price. To make them directly comparable, we have transformed both time series into indices by setting the values of the first observations to 100 and then adjusting by the percentage changes of the prices each week.

Figure 5.1: Movements of the P1A Price and the IMAREX PM4TC Freight Futures
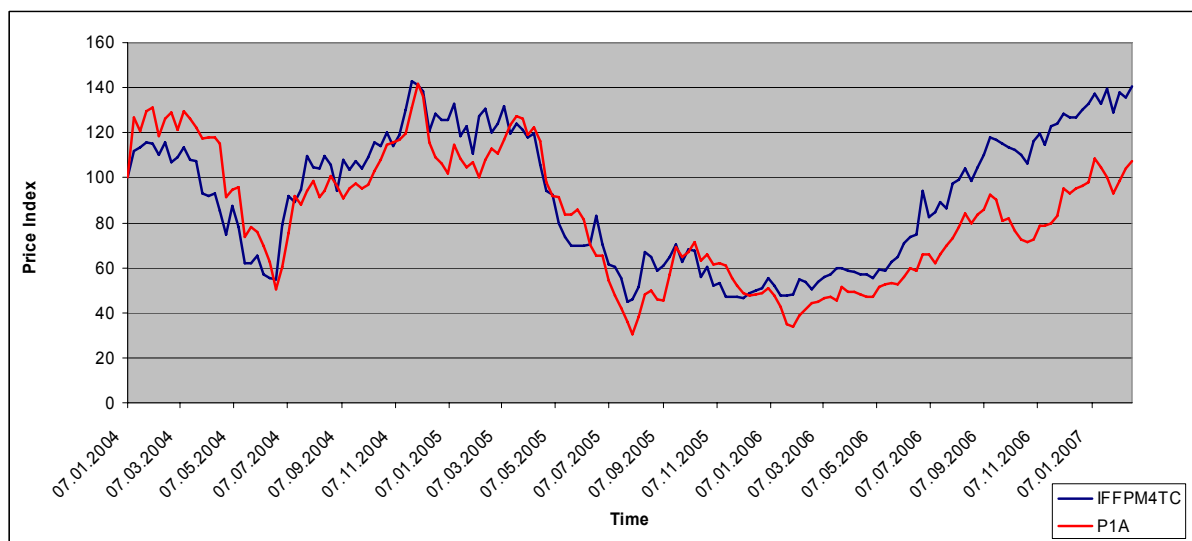
We see that the two prices move together, but that movements of the P1A price are far from perfectly mirrored by the futures price. This is both due to the fact that the IFF PM4TC is an average of the four Panamax T/C routes and because of the averaging over the settlement month. Hedging one of the four routes with the futures contract will thus be a cross hedge that we do not expect a very high performance from. As discussed in Chapter 3.5.2, we would expect at least equally good and perhaps better hedging performance than shown for the BIFFEX contract in Table 3.3.

5.2 Hedging with Constant Hedge Ratio

In this chapter we use Ederington's Framework[35] to calculate the hedging effectiveness of the PM4TC contract. We study how much a shipowner can reduce his variance by hedging with the PM4TC when initially operating spot on each of the four underlying routes (P1A, P2A, P3A and P4). We also consider the hedging performance for a shipowner who might have an equal position in each of the four routes (AVG4TC).

When calculating the hedging effectiveness, we use the regression Equation (4.9). We then assume that a shipowner operates one of the four Panamax routes constantly in the time

---

[35] See Chapter 4.1.

period under study. To hedge his exposure in spot, he uses the PM4TC contract. From Equation (4.9), we see that he must hedge with α_1 PM4TC contracts for each position in spot.

We will now test our estimated models for the assumptions underlying a regression. Then we will present the hedging effectiveness of the PM4TC contract.

5.2.1 Residual Diagnostics on the Estimated Models

In Table 5.4, we have shown the residual diagnostics performed on the regression models. One regression model is specified for each of the four Panamax routes. In addition, we have specified a model on the average of the spot assessments of the four routes (AVG4TC). The tests performed are those presented in Chapter 4.1.3. For our analyses of the hedging effectiveness, it is important that the hedge ratio (i.e. the negative of the slope coefficient in the regression) is unbiased and consistent and that the coefficient of determination is correct. We will now describe the results of the diagnostics tests.

Table 5.4: Residual Diagnostics on the Estimated Models

	P1A	P2A	P3A	P4	AVG4TC	Critical values 5 %	1 %
White's test	0,88741	1,31772	5,08146	6,54070	3,34817	5,99	9,21
Breusch-Godfrey (1)	0,12721	0,00352	1,39421	0,05949	0,99323	3,84	6,64
Breusch-Godfrey (2)	2,26947	0,57183	9,82774	8,49946	7,03668	5,99	9,21
Breusch-Godfrey (4)	4,55759	5,67572	10,16091	9,33616	8,40378	9,49	13,28
Breusch-Godfrey (12)	21,81290	20,71095	17,93795	15,57729	17,25685	21,03	26,22
Breusch-Godfrey (26)	36,03618	35,00944	31,52281	24,65843	30,10397	38,89	45,64
Breusch-Godfrey (52)	53,56495	55,60313	53,54890	48,87705	52,85532	69,83	78,62
Skewness	-0,10993	-0,18287	-0,19662	0,10065	-0,09385		
Kurtosis	3,79594	4,04315	3,59923	4,49341	3,60773		
Jarque-Bera	4,63095	8,29896	3,48894	15,42238	2,74771	5,99	9,21
ARCH (1)	5,38165	0,89024	7,28240	15,79443	11,52708	3,84	6,64
ARCH (2)	6,08642	2,74611	10,48338	16,11186	11,34663	5,99	9,21
ARCH (5)	11,17998	15,96962	12,10436	17,31714	13,00158	11,07	15,09

For White's test we have shown the chi-square test statistics. Two degrees of freedom for all routes.
Numbers in parenthesis indicate the number of lags in the specific test.
Numbers written in red are above the critical value. The null hypothesis is then rejected.

Tests for Heteroscedasticity

White's test for general heteroscedasticity shows that the null hypothesis of homoscedasticity is rejected only for P4. As discussed, the coefficient estimates will still be unbiased and consistent, so this does not pose a problem.

To check for autoregressive conditional heteroscedasticity, we employ Engle's ARCH LM test. We test for $z = 1$, 2 and 5 and see that the null hypothesis of no ARCH-effects is rejected for all routes. The coefficient estimates are still unbiased and consistent, but the estimated regression model might not be efficient. An ARCH representation of the second moments might incorporate these effects and result in a more efficient model.

The Breusch-Godfrey Test for Autocorrelation

We use the Breusch-Godfrey test to examine if autocorrelation is present in the residuals. As we use weekly data, we find it appropriate to test for 1, 2, 4, 12, 26 and 52 lags. We see that the null hypothesis of no autocorrelation is rejected for all routes but the P2A. The coefficient estimates will still be unbiased and consistent, but the coefficient of determination might be inflated. Our results of the hedging performance might therefore look better than they in fact are.

Test for Normality

To test for normality, we use the Jarque-Bera test. We see that the null hypothesis of normality is rejected for the routes P2A and P4. All of the routes have a greater kurtosis than three. The distributions are thus more peaked than the normal distribution. None of the routes looks to be particularly skewed. Because of the Central Limit Theorem, we can still make valid inferences, even though the residuals are not normally distributed.

To conclude, we can say that the coefficient estimates in our models are unbiased and consistent, but that autocorrelation in route P1A, P3A, P4 and AVG4TC might inflate the coefficients of determination. Autocorrelation can be dealt with by either including lagged errors in the regression model or by estimating a VAR model. In Chapter 5.2.2, we estimate both CLRMs and regression models with AR(p) errors.[36] We will estimate VAR models in Chapter 5.3.

The residual diagnostics also showed presence of ARCH-effects. This implies that incorporating a GARCH error structure might make the models more efficient. In addition,

---

[36] Residual diagnostics and lag length selection for the LRM AR(p) models are shown in Appendix 3. The chosen models do not show signs of autocorrelation, but there are still ARCH effects present. Normality is rejected for P1A, P2A and P4.

spot and futures prices could be cointegrated, justifying the use of a VECM-GARCH-X model. We will test for the most appropriate specification in Chapter 5.3.

5.2.2 Hedging Performance of the PM4TC

In Table 5.5, we show the results of our hedging analyses with constant hedge ratios. First, we show variance comparisons between the unhedged portfolios and each of the three estimated constant-hedge strategies.[37] We then show the estimated values of the hedge ratios. Finally, we present the calculated hedging performance of the IMAREX PM4TC for each of the four routes and for the average of them.

Table 5.5: Hedging Effectiveness

Variance Comparisons					
	P1A	P2A	P3A	P4	AVG4TC
Unhedged	0,007612653	0,007186156	0,011471654	0,010456200	0,007860894
Naive Hedge	0,006402283	0,006318622	0,008792544	0,008455680	0,006282778
CLRM Hedge	0,005203767	0,004974109	0,008124541	0,007561702	0,005231541
LRM AR(p) Hedge	0,005204968	0,004978078	0,008125506	0,007565809	0,005231541

Hedge Ratio					
	P1A	P2A	P3A	P4	AVG4TC
Naive Hedge	1,0	1,0	1,0	1,0	1,0
CLRM Hedge	0,5864	0,5619	0,6912	0,6428	0,6126
LRM AR(p) Hedge	0,5995	0,5381	0,7030	0,6670	0,6124
Std.Err. CLRM Hedge	0,0679	0,0664	0,0849	0,0819	0,0681
t-stat. CLRM Hedge	8,6330	8,4616	8,1442	7,8504	8,9954
Std.Err. LRM AR(p) Hedge	0,0637	0,0643	0,0835	0,0804	0,0667
t-stat. LRM AR(p) Hedge	9,4130	8,3666	8,4220	8,2940	9,1757

Variance Reduction					
	P1A	P2A	P3A	P4	AVG4TC
Naive Hedge	15,90 %	12,10 %	23,30 %	19,10 %	20,10 %
CLRM Hedge	31,64 %	30,78 %	29,18 %	27,68 %	33,45 %
LRM AR(p) Hedge	31,63 %	30,73 %	29,17 %	27,64 %	33,45 %

The percentages have been rounded to two decimal places.

CLRM is short for Classical Linear Regression Model.

LRM AP(p) is short for Linear Regression Model with AR(p) lags.

Red colouring denotes the route with the lowest hedge effectiveness.

Green colouring denotes the route with the highest hedge effectiveness.

---

[37] The minimum variance hedge ratio (CLRM Hedge) is calculated from Equation (4.9). The naive hedge equals Equation (4.9) with $\alpha_0 = 0$ and $\alpha_1 = 1$. The hedge ratio for the LRM AR(p) model equals α_1 in Equation (4.18).

We see from Table 5.5 that the naive hedges, the CLRM hedges and the LRM AR(p) hedges all reduce the variance of the portfolios compared to the unhedged cases. The hedging performances of the naive hedges are, however, much poorer than for both of the regression hedges.

The estimated hedge ratios are all statistically significant. The CLRM hedge performs best for route P1A, where the variance is reduced by 31,64%. Hedging performance is even better for the average of the four routes (AVG4TC), where the variance reduction is 33,45%. The worst hedge is achieved for route P4 with 27,68% variance reduction. For the LRM AR(p) hedges, we see that the hedging performances is equal or lower compared to the CLRM hedges. This is as it always will be in in-sample studies.

If we compare to the results from the study of the BIFFEX by Kavussanos and Nomikos (2000b), we see that the IMAREX PM4TC performs much better than the BIFFEX did when The Baltic Freight Index was the underlying (Table 3.2). After the BFI was exchanged for the BPI, the hedging performance of the BIFFEX improved considerably (see Table 3.3). For some routes, it was in fact better than what we have calculated for the IMAREX PM4TC contract.

5.3 Hedging with Time-Varying Hedge Ratios

In this chapter we will employ time series models that try to remedy the problems of heteroscedasticity and autocorrelation that we experienced with the regression model. We will use a VAR model to better capture the information in the data and correct for autocorrelation. If cointegration is present, we will extend the VAR model to a VECM. Heteroscedasticity will be dealt with by employing a GARCH specification of the second moments.

5.3.1 Choosing the Most Appropriate Model

We will use Johansen's test (Johansen, 1988) for cointegration to choose between specifying a VAR model or a VECM. To perform the Johansen test in EViews, we have to find the optimal number of lags to include. In order to do this, we estimate VAR models for the different routes with lag length from 0 to 12. The VAR model with the lowest value of the

Schwarz Information Criterion indicates the appropriate number of lags to include. The results are shown in Table A4.1. We see that two lags are appropriate for all routes.

The Pantula principle is then used to choose the appropriate assumptions for the Johansen test. The results from the selection procedure are shown in Table A4.2. We see that model 2 is chosen in all cases. Model 2 has an intercept, but no trend in the cointegrating equation and no intercept or trend in the VAR.

We see that there is no evidence of cointegration for any of the routes (Table A4.2). This implies that a VAR model is more appropriate than a VECM. The VAR model will be specified with two lags for all routes, as suggested by the values of the SIC in Table A4.1. As we earlier found presence of autoregressive conditional heteroscedasticity, we will specify the second moments with a GARCH model. Because no cointegration was found, Lee's GARCH-X model is not appropriate; we will thus use the BEKK model presented in Chapter 4.3.3.

To find the optimal time-varying hedge ratios, we will therefore employ a VAR-GARCH model. The VAR model and its notations are given by Equation (4.20). The BEKK model and its notations are given in Equation (4.27). The following specification is then the final model we will use in our analysis:

$$\Delta S_t = c_s + \sum_{i=1}^{k} \beta_{si} \Delta S_{t-i} + \sum_{i=1}^{k} \theta_{si} \Delta F_{t-i,T} + u_{st} \qquad \Delta F_{t,T} = c_f + \sum_{i=1}^{k} \beta_{fi} \Delta S_{t-i} + \sum_{i=1}^{k} \theta_{fi} \Delta F_{t-i,T} + u_{ft}$$

$$u_t = \begin{pmatrix} u_{st} \\ u_{ft} \end{pmatrix} \bigg| \Omega_{t-1} \sim distr(0, H_t)$$

$$H_t = \begin{pmatrix} h_{SS,t} & h_{SF,t} \\ h_{SF,t} & h_{FF,t} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{pmatrix}' \begin{pmatrix} c_{11} & c_{12} \\ 0 & c_{22} \end{pmatrix} + \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}' u_{t-1} u'_{t-1} \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix}' H_{t-1} \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix}$$

We estimate the model with a maximum likelihood procedure. The estimated coefficients and their standard deviations are shown in Appendix 4 in Table A4.3.

5.3.2 Diagnostics and Comments to the Estimated Models

In this chapter we will perform residual diagnostics on the estimated models. We also study the time-varying hedge ratios. This includes testing the hedge ratio series for stationarity and comparing the series to the constant hedge ratios.

Residual diagnostics tests are made on the standardised ARCH residuals, $u_t / \sqrt{h_t}$, and the standardized squared ARCH residuals, $(u_t / \sqrt{h_t})^2$. We have tested for normality, autocorrelation and ARCH effects. The results are shown in Table 5.6. We see that most of the residuals are not normally distributed. As the sample size is reasonably large, this might not be a major problem. Further, the tests show no evidence of autocorrelation or ARCH effects. This indicates that the VAR-GARCH models are well specified and that we have remedied the problems with the regression models first estimated.

Table 5.6: Residual Diagnostics

	P1A		P2A		P3A		P4		AVG4TC		Critical Value	
	Spot	Futures	Spot	Futures	Spot	Futures	Spot	Futures	Spot	Futures	5 %	1 %
Jarque-Bera	2,97646	19,87372	9,72177	17,24131	3,89578	29,09550	3,52186	36,87716	8,74580	25,47555	5,99	9,21
Q(12)	9,84529	6,79128	11,97337	6,93861	15,48354	6,37302	14,42235	5,70116	9,92666	6,15537	36,42	42,98
Q^2(12)	8,58819	9,33364	8,88421	9,61875	11,36925	6,53204	13,69788	7,20001	8,44447	7,36270	36,42	42,98
ARCH(1)	0,05880	0,00846	0,00941	0,01257	0,03305	0,33538	0,07799	0,31241	0,09203	0,07416	3,84	6,64
ARCH(2)	0,15478	1,04440	0,38969	1,06630	0,10993	1,69234	0,10257	1,34436	0,24609	1,33515	5,99	9,21
ARCH(5)	3,62896	4,09506	4,49161	3,83974	1,02931	3,05005	0,63696	2,87902	2,95712	3,32604	11,07	15,09

Jarque-Bera is a test for normality. The test statistic is chi-square distributed with two degrees of freedom. Under the null hypothesis, the residuals are normally distributed.

Q and Q^2 are Ljung-Box test statistics. The Ljung-Box test statistic is asymptotically chi-square distributed with degrees of freedom equal to the number of autocorrelations. The null hypothesis is that no autocorrelation is present up to order k (where k in this case is 12).

ARCH(p) is Engle's LM test for conditional heteroscedastisity. The null hypothesis is that there is no ARCH up to order p in the residuals. The LM test statistic is asymptotically chi-square distributed with degrees of freedom equal to the number of p.

Red colouring denotes rejection of the null hypothesis.

After having estimated the models, we calculate the time-varying hedge ratios with Equation (4.29). In Table A4.4, we have shown the constant hedge ratios calculated from the CLRM, LRM AR(p) and VAR models. In addition, the averages of the time-varying hedge ratios are shown along with their standard deviations and ADF tests for stationarity in the series on level. We see that the averages of the time-varying hedge ratios are close to the values of the constant hedge ratios. The ADF tests show that the series of hedge ratios are stationary for all routes. This implies that the series of hedge ratios are mean-reverting and thus that the impact of shocks to the series gets smaller and smaller and eventually dies away.

In Figure 5.2, we illustrate by plotting both the time-varying hedge ratio and the constant CLRM hedge ratio for route P1A. We see clearly that the VAR-GARCH hedge ratio varies as new information arrives in the market. This should make the hedging performance of the model better, and we therefore expect a higher hedging effectiveness compared to hedging with a constant hedge ratio.

Figure 5.2: Hedge Ratios for Route P1A (Time-Varying and Constant)
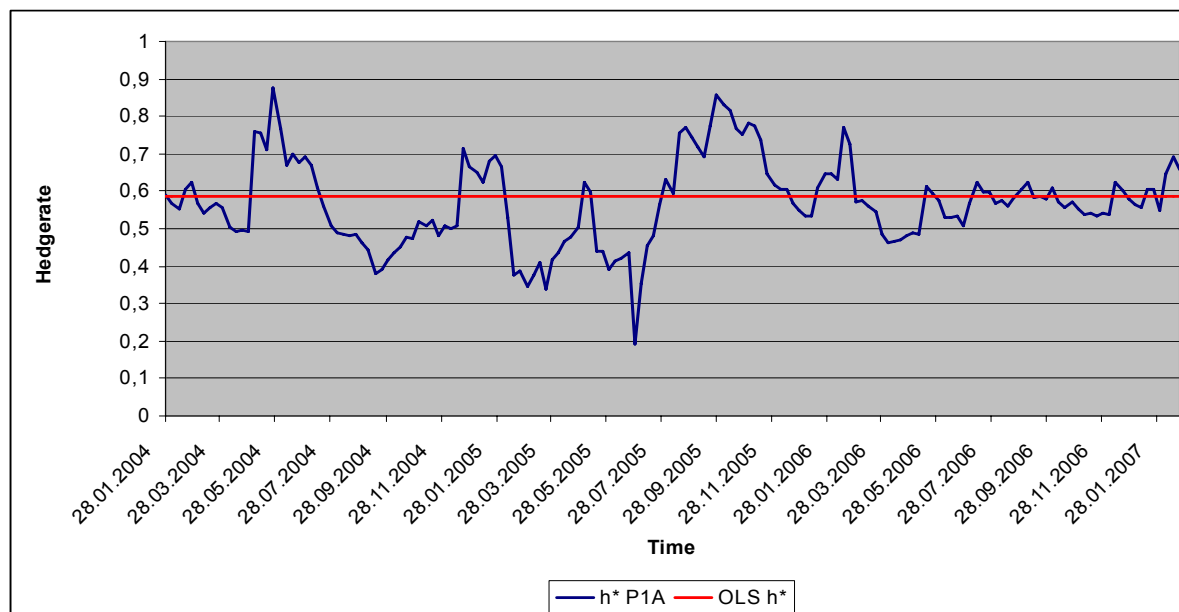
It is, however, by no means certain that a VAR-GARCH model will outperform a model with a constant hedge ratio. The added complexity in estimating the GARCH model might reduce the hedging effectiveness. Several studies have been made on the subject in other markets. In the wheat futures market, for example, Myers (1991) finds that constant hedge ratios are superior to time-varying hedge ratios. Garcia et al. (1995) come to the same conclusion for the soybean futures market. Studies that show best performance for time-varying hedge ratios are for example Kroner and Sultan (1993), who study foreign currency futures, and Gagnon and Lypny (1995), who study the hedging of short-term interest risk.

5.3.3 Hedging Performance of the PM4TC

The hedging performances of all the estimated models are shown in Table 5.7. Hedging performances for the VAR-GARCH models have been calculated by using Equations (4.30)

and (4.31).[38] For comparison purposes, we have also shown the hedging performances estimated by using a VAR model for each of the four routes. Residual diagnostics are not shown for the VAR models.

Table 5.7: Hedging Effectiveness

Variance Comparisons					
	P1A	**P2A**	**P3A**	**P4**	**AVG4TC**
Unhedged	0,007612653	0,007186156	0,011471654	0,010456200	0,007860894
Naive Hedge	0,006402283	0,006318622	0,008792544	0,008455680	0,006282778
CLRM	0,005203767	0,004974109	0,008124541	0,007561702	0,005231541
LRM AR(p)	0,005204968	0,004978078	0,008125506	0,007565809	0,005231541
VAR(2)	0,005203990	0,004974319	0,008131571	0,007580428	0,005232330
VAR(2)-GARCH	0,005008602	0,004754647	0,008023494	0,007326590	0,005049526

Variance Reduction					
	P1A	**P2A**	**P3A**	**P4**	**AVG4TC**
Naive Hedge	15,90 %	12,10 %	23,30 %	19,10 %	20,10 %
CLRM	31,64 %	30,78 %	29,18 %	27,68 %	33,45 %
LRM AR(p)	31,63 %	30,73 %	29,17 %	27,64 %	33,45 %
VAR(2)	31,64 %	30,78 %	29,12 %	27,50 %	33,44 %
VAR(2)-GARCH *	31,78 %	30,16 %	29,50 %	29,60 %	34,26 %

The percentages have been rounded to two decimal places.

Green colouring denotes the route with the highest hedge effectiveness.

Red colouring denotes the route with the lowest hedge effectiveness.

Yellow colouring denotes the model with the best hedging performance.

In an in-sample study of hedging performance, the LRM AR(p) hedges and the VAR hedges can never be better than the CLRM hedges. This is because the CLRM estimates the variance minimising hedge ratio for that particular sample. No other constant hedge ratio can be more effective. In Table 5.7, we see that the hedging performances of the LRM AR(p) and the VAR models are equal or slightly below the performances of the CLRMs. In out-of-sample studies and when hedging in practice, however, LRM AR(p) or VAR models might be better. This is due to the fact that the models are richer and that they might better capture structures in the data, like for example autocorrelation.

We see from Table 5.7 that, except for the naive hedges, the hedging performances are approximately 30% for all routes. This is about the same as what was achieved for the BIFFEX in the period when the BPI was underlying the contract (see Table 3.3). For the BIFFEX, however, the spread of the hedging performance for the different routes were larger. For example, the performance for Route 1A was only 24,55%, whereas the performance for

---

[38] Due to uncertainties regarding the best way to calculate the hedging performance of the VAR-GARCH model, Table A4.5 with a different calculation method is shown in Appendix 4.

Route 3A was 39,95% (Kavussanos and Nomikos, 2000b). This is about 10% better performance than what is achieved by hedging route P3A with IFF PM4TC.

We see from Table 5.7 that the VAR-GARCH model is better than the other models for AVG4TC and for all routes but for the P3A, where the CLRM performs slightly better. The difference in hedging performance between the VAR-GARCH model and the CLRM ranges from -0,62% for route P2A to +1,92% for route P4. Compared to the total hedging effectiveness of around 30%, an increase in performance of below 2% is not very much.

The VAR-GARCH model achieves the best hedging effectiveness for route P1A with 31,78% variance reduction. The poorest performance is achieved for route P3A with 29,50%. For the average of the four routes, AVG4TC, the hedging performance is 34,26%. This is better performance than for any of the individual routes.

According to the analyses in this chapter, we conclude that time-varying hedge ratios outperform constant hedge ratios in the freight futures market. The estimated hedging performance of the IFF PM4TC contract is about 30% for all routes. This is very low compared to other futures markets, like for example futures contracts on exchange rates. There are, however, no better alternatives for hedging in the shipping market. As discussed in Chapter 3.4, there are only minor differences between the IFF PM4TC contract and a comparable FFA contract. We therefore assume that the FFA contracts do not offer better performance. A theoretical alternative is of course to tailor-make an FFA contract on a particular route, but it would probably be difficult to find a counterpart and the price would be high. The alternative is therefore not a practical one. Choosing between hedging with freight futures from IMAREX or OTC FFA contracts is then mostly a choice of removing counterparty risk or not.

The analyses we performed were in-sample. To conclude which of the models are best in practice, out-of-sample tests should be performed. We briefly show how to perform out-of-sample tests in Appendix 5. Based on our results, however, we will recommend a hedger to use a VAR-GARCH model to calculate hedge ratios.

6. The Unbiasedness Hypothesis

In this chapter we will test if the unbiasedness hypothesis holds for the IMAREX PM4TC freight futures contract. First, we will present the conditions necessary for the unbiasedness hypothesis to hold. Then, we will describe the properties of the data series of spot and futures prices. Finally, we will employ cointegration techniques to test if the conditions are fulfilled for the PM4TC contract.

6.1 Conditions Necessary for the Unbiasedness Hypothesis to Hold

In Chapter 3.4.1 and 3.4.2 we presented the unbiasedness hypothesis. Chapter 3.5.1 reviewed prior studies of the hypothesis. In Equation (3.4), the pricing function of futures prices were given as:

$$F_{t,T} = E_t(S_T)$$

which states that the futures price is an unbiased predictor of the future spot price. Rational expectations and no risk premium are assumed. The hypothesis implies that the futures price at time t for delivery at time t + 1 differs from the spot price realised at time t + 1 only by a random error, u_{t+1}. This can be written as:

$$S_{t+1} = F_{t,t+1} + u_{t+1}; \qquad u_{t+1} \sim iid(0, \sigma^2) \tag{6.1}$$

To test the hypothesis, one traditionally studied the following relationship:

$$S_{t+1} = \beta_1 + \beta_2 F_{t,t+1} + u_{t+1}; \qquad u_{t+1} \sim iid(0, \sigma^2) \tag{6.2}$$

where β_1 and β_2 are parameters. It was then tested to see if $\beta_1 = 0$ and $\beta_2 = 1$ simultaneously. If they were, the hypothesis of unbiasedness was not rejected.

If spot and futures prices are nonstationary, as they most often are, the calculated standard errors of the parameters in Equation (6.2) are not valid. Inferences about β_1 and β_2 can therefore not be performed. In such cases, cointegration techniques might be employed. This of course requires that the two series are cointegrated.

Generally, there is no long-run relationship between two nonstationary variables. There is then no constant mean that will be returned to frequently, and the series are free to wander apart without bound. Consequently, the unbiasedness hypothesis will not hold. If spot and futures prices are cointegrated, however, there will be a long-run relationship between them. The hypothesis will then hold under certain conditions, which we will explain in a moment.

The long-run relationship between spot and futures prices is given by the Error Correction Term. The ECT was defined in Equation (4.22) as:

$$Z_{t-1} = S_{t-1} + \beta_1 + \beta_2 F_{t-1,T}$$

If the relationship between spot and futures prices is specified as a VECM, like in Equation (4.23), the ECT will correct for deviations from the long-run relationship. The unbiasedness hypothesis will hold if the parameters in the ECT are as follows, $\beta_1 = 0$ and $\beta_2 = -1$.

We have thus established two conditions that are necessary for the unbiasedness hypothesis to hold. First, spot and futures prices have to be cointegrated. Second, the cointegrating equation has to be on the following form:

$$Z_{t-1} = S_{t-1} - F_{t-1,T} \tag{6.3}$$

To test the hypothesis, we will first test spot and futures prices for stationarity. If they are nonstationary of the same order, we will use Johansen's test for cointegration. Tests on the parameters will be performed if cointegration is found.[39]

---

[39] How to perform inferences on the parameters is explained in Johansen and Juselius (1990).

If the unbiasedness hypothesis is found not to hold, this is either because the market participants are not behaving rationally, because there is a risk premium in the futures price or because of a combination of both. A risk premium might arise if there are more long hedgers than short hedgers, or if there are more short hedgers than long hedgers. Speculators will then have to close the gap and will demand a risk premium to do so. Lack of rationality, on the other hand, might for example stem from the fact that the settlement price for the PM4TC contracts are calculated as the average of the spot price assessments over the settlement month, and that the market participants do not fully take this averaging into consideration when trading futures contracts.

Any test of the unbiasedness hypothesis is a joint test for risk premium and lack of rationality. If the unbiasedness hypothesis is rejected, we will therefore not be able to conclude which of the two assumptions that are violated, or if they both are violated.

Another point that is interesting to notice is that the futures prices – even when they are biased estimators of futures spot prices – might still predict future spot prices better than forecasts from time series models with present and past spot prices. Such time series models might for example be ARIMA, exponential smoothing or error correction models. We will not elaborate further on this topic.

6.2 Properties of the Data Series

In this analysis we are going to test if one, two and three months to maturity futures contracts are unbiased predictors of the spot price at maturity. The time periods for each of the three data sets are shown in Table 6.1. As IMAREX first started trading monthly contracts in the beginning of April 2005, our first observations are for the futures prices 29 April 2005. Because we have monthly observations, the number of observations is just above 20 for each of the series. This is much fewer observations than we had when analysing the hedging effectiveness, and this might possibly result in small-sample problems when performing the analyses.

Table 6.1: Presenting the Data Series

Series	N	First spot observation	Last spot observation	First futures observation	Last futures observation
Series 1	23	AVG4TC May 2005	AVG4TC March 2007	29 April 2005	28 February 2007
Series 2	22	AVG4TC June 2005	AVG4TC March 2007	29 April 2005	31 January 2007
Series 3	21	AVG4TC July 2005	AVG4TC March 2007	29 April 2005	29 December 2006

Series 1: Spot at maturity and futures prices one month before maturity.
Series 2: Spot at maturity and futures prices two months before maturity.
Series 3: Spot at maturity and futures prices three months before maturity.

N is the number of observations.

AVG4TC May 2005 is the average of The Baltic Exchange spot price assesments over the number of trading days in May etc.

Futures prices have been supplied by IMAREX, spot prices by The Baltic Exchange.

In Chapter 5, the spot price was given as The Baltic Exchange price assessment on a particular day. In this chapter, however, the spot price is given as the average of the spot price assessments over the settlement month of the corresponding futures contract. This is because the spot price average over the settlement month is the actual settlement price of the futures contracts.[40] When testing the unbiasedness hypothesis, this is the correct spot price to use in the analysis. Because of this, we have higher expectations of finding cointegration between spot and futures prices in this analysis, than we had when analysing hedging effectiveness (Chapter 5).

In Table 6.2, we present summary statistics on the logarithmic first differences of the data series. We see that the mean is around zero, but that the standard deviations are all very large. This fact is also seen in the large maximum and minimum values. From the Jarque-Bera test, we see that the null hypothesis of normality is rejected only for the spot price in the 2-Month Price Series. The large differences in the Jarque-Bera test statistic for the three spot price series are probably due to the low number of observations in the data samples.

The Augmented Dickey Fuller tests show that spot and futures price series are all nonstationary on level, but becomes stationary after differencing once. Nonstationarity of the same order is a prerequisite for cointegration.

---

[40] See Appendix 2.

Table 6.2: Summary Statistics

	N	Mean	STD	Max	Min	Skewness	Kurtosis	Jarque-Bera	ADF lvl	ADF 1diff
1-Month Price Series										
Spot	22	0,0202	0,1521	0,2652	-0,4200	-1,0329	4,3633	5,6157	-0,3736	-3,4136
Futures	22	0,0175	0,1900	0,3483	-0,4193	-0,1368	2,6727	0,1668	-0,8813	-4,2009
2-Month Price Series										
Spot	21	0,0304	0,1479	0,2652	-0,4200	-1,2369	5,2213	9,6717	-0,3736	-3,4136
Futures	21	0,0088	0,1802	0,3382	-0,3015	0,3091	2,0711	1,0895	-1,1343	-3,6960
3-Month Price Series										
Spot	20	0,0530	0,1087	0,2652	-0,1559	-0,0903	2,3615	0,3669	-0,2175	-3,4136
Futures	20	0,0121	0,1813	0,3413	-0,3234	0,1986	2,1991	0,6659	-1,2332	-3,5848

N is the number of observations in the sample. Max and Min are the maximum and minimum values in the sample. Mean, STD, Skewness and Kurtosis are respectively the first, second, third and fourth moments of the data series.

The Jarque-Bera test for normality is chi-square distributed with two degrees of freedom. The 5% critical value is 5,99. Red colouring denotes rejection of the null hypothesis of normality.

ADF lvl is the augmented Dickey-Fuller test on level. ADF 1diff is the same test on first differenced data. We have included an intercept term in the test. The null hypothesis is that a unit root is present. The lag length is automatic and based on SIC.

Numbers written in red indicate rejection of the null hypothesis.

All series are measured in logarithmic first differences.

We then test for cointegration by using Johansen's procedure. The selection of lag length, selection of the proper assumptions to include in the test and the Johansen test itself are shown and commented in Appendix 6. As we see from Table A6.2, spot and futures prices are cointegrated in all cases. One lag and Model 2 are chosen for all three series.

The necessary condition of cointegration is thus fulfilled. To test for the second condition, that the cointegration equation is on the form given in Equation (6.3), Vector Error Correction Models have to be estimated and inferences made about the parameters. First then can we conclude whether the futures prices are biased or unbiased estimators of future spot prices.

6.3 Testing the Unbiasedness Hypothesis

Having identified that the spot and futures price series are I(1) and that they are cointegrated, we specify the following VECMs:

VECM for Spot and 1-Month Futures Prices

$$\begin{pmatrix} \Delta S_t \\ \Delta F_{t-1,t} \end{pmatrix} = \begin{pmatrix} \Gamma_{11} \Gamma_{12} \\ \Gamma_{21} \Gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta S_{t-1} \\ \Delta F_{t-2,t-1} \end{pmatrix} + \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} \left(1 \, \beta_1 \, \beta_2\right) \begin{pmatrix} S_{t-1} \\ 1 \\ F_{t-2,t-1} \end{pmatrix} + \begin{pmatrix} u_{st} \\ u_{ft} \end{pmatrix}; \begin{pmatrix} u_{st} \\ u_{ft} \end{pmatrix} \sim iid(0, \Sigma) \qquad (6.4)$$

VECM for Spot and 2-Month Futures Prices

$$\begin{pmatrix} \Delta S_t \\ \Delta F_{t-2,t} \end{pmatrix} = \begin{pmatrix} \Gamma_{11} \Gamma_{12} \\ \Gamma_{21} \Gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta S_{t-1} \\ \Delta F_{t-3,t-1} \end{pmatrix} + \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} \left(1 \, \beta_1 \, \beta_2\right) \begin{pmatrix} S_{t-1} \\ 1 \\ F_{t-3,t-1} \end{pmatrix} + \begin{pmatrix} u_{st} \\ u_{ft} \end{pmatrix}; \begin{pmatrix} u_{st} \\ u_{ft} \end{pmatrix} \sim iid(0, \Sigma) \qquad (6.5)$$

VECM for Spot and 3-Month Futures Prices

$$\begin{pmatrix} \Delta S_t \\ \Delta F_{t-3,t} \end{pmatrix} = \begin{pmatrix} \Gamma_{11} \Gamma_{12} \\ \Gamma_{21} \Gamma_{22} \end{pmatrix} \begin{pmatrix} \Delta S_{t-1} \\ \Delta F_{t-4,t-1} \end{pmatrix} + \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} \left(1 \, \beta_1 \, \beta_2\right) \begin{pmatrix} S_{t-1} \\ 1 \\ F_{t-4,t-1} \end{pmatrix} + \begin{pmatrix} u_{st} \\ u_{ft} \end{pmatrix}; \begin{pmatrix} u_{st} \\ u_{ft} \end{pmatrix} \sim iid(0, \Sigma) \qquad (6.6)$$

where ΔS_t is the logarithmic spot price difference, and $\Delta F_{t-1,t}$, $\Delta F_{t-2,t}$ and $\Delta F_{t-3,t}$ are the logarithmic futures price differences at time t-1, t-2 and t-3, respectively. Γ_{ir} (i = 1, 2 and r = 1, 2) are coefficients measuring the short-run adjustment of the system to changes in the vector of spot and futures prices. Φ_1 and Φ_2 are the error correction coefficients measuring the speed of convergence to the long-run steady state. The matrix $\left(1 \, \beta_1 \, \beta_2\right)$ contains the cointegrating parameters, where the coefficient of S_{t-1} is normalised to be unity. u_{st} and u_{ft} are white noise residuals, and Σ is a $2 X 2$ variance/covariance matrix.

6.3.1 Coefficient Estimates and Residual Diagnostics on the VECMs

The three VECMs are then estimated. The coefficient estimates are shown in Table 6.3. We see from the table that β_1 is around 0 for all three series. According to the t-statistic, the hypothesis of $\beta_1 = 0$ is not rejected for any of them. In addition, we see that β_2 is around -1 for all three series. The hypothesis of $\beta_2 = 0$ is rejected in all cases. A test of $H_0 : \beta_2 = -1$ is not performed here. These results indicate that the unbiasedness hypothesis might hold. To be sure, however, we have to perform a test that $\beta_1 = 0$ and $\beta_2 = -1$ simultaneously. This is done in Chapter 6.3.2.

Table 6.3: Coefficient Estimates

1-Month Price Series Coefficient Estimates							
						$\beta' = (1\ \beta_1\ \beta_2)$	
	Γ_{11}	Γ_{12}	Γ_{21}	Γ_{22}	Φ_1	Φ_2	β_1	β_2
Coeff	1,1380	-0,5806	0,1667	0,0300	-0,8668	1,5503	-0,2799	-0,9728
STD	0,9085	0,4745	0,4075	0,2129	1,1151	0,5002	0,1405	0,0141
t-stat	(1,2526)	(-1,2234)	(0,409)	(0,1408)	(-0,7773)	(3,0995)	(-1,9920)	(-69,0943)

2-Month Price Series Coefficient Estimates							
						$\beta' = (1\ \beta_1\ \beta_2)$	
	Γ_{11}	Γ_{12}	Γ_{21}	Γ_{22}	Φ_1	Φ_2	β_1	β_2
Coeff	0,3109	-0,0691	-0,6569	0,1698	-0,1519	1,1122	0,1692	-1,0177
STD	0,3769	0,1556	0,2438	0,1007	0,2674	0,1730	0,5875	0,0590
t-stat	(0,8249)	(-0,4440)	(-2,6943)	(1,6866)	(-0,5682)	(6,4301)	(0,2879)	(-17,2534)

3-Month Price Series Coefficient Estimates							
						$\beta' = (1\ \beta_1\ \beta_2)$	
	Γ_{11}	Γ_{12}	Γ_{21}	Γ_{22}	Φ_1	Φ_2	β_1	β_2
Coeff	0,3636	-0,0191	-0,8768	0,1144	-0,0520	0,6838	-0,0277	-0,9989
STD	0,2967	0,1594	0,2483	0,1334	0,1357	0,1136	1,3306	0,1339
t-stat	(1,2255)	(-0,1199)	(-3,5317)	(0,8579)	(-0,3830)	(6,0212)	(-0,0208)	(-7,4593)

t statistic for the null hypothesis coeff = 0 is is parenthesis.
Red colouring denotes rejection of the null hypothesis.

The coefficient Φ_2 is positive and statistically significant for all series, whereas Φ_1 is insignificant. This is in accordance with convergence to long-run equilibrium. For example, if $S_{t-1} > F_{t-2,t-1}$ then the price of the futures contract will increase in the next period to approach long-run equilibrium. The spot price in the next period will not be significantly affected by

the ECT. Thus, only the futures price is responding to deviations from long-run equilibrium. These findings also indicate unbiasedness in the futures prices.

The residual diagnostics on the estimated models are shown in Table 6.4. We see that the Q-tests show that no autocorrelation is present up to order 5 for any of the three series. The bivariate LM-test for autocorrelation of order one, however, show evidence of autocorrelation for the two-month price series. The low values of the Q^2-statistics suggest that conditional homoscedasticity cannot be rejected.

Table 6.4: Residual Diagnostics

Residual Diagnostics - 1-Month Price Series						
	Q(5)	Q^2(5)	Normality	White's test*	Normality*	LM(1)*
u_{st}	1,4248	1,2372	4,9939	27,9109	9,2000	6,7070
u_{ft}	2,3044	5,6248	4,2061			
5% c.v.	11,07	11,07	5,99	40,11	9,49	9,49

Residual Diagnostics - 2-Month Price Series						
	Q(5)	Q^2(5)	Normality	White's test*	Normality*	LM(1)*
u_{st}	4,6337	4,6537	24,2166	35,2392	47,3210	13,7606
u_{ft}	8,1213	8,1213	23,1044			
5% c.v.	11,07	11,07	5,99	40,11	9,49	9,49

Residual Diagnostics - 3-Month Price Series						
	Q(5)	Q^2(5)	Normality	White's test*	Normality*	LM(1)*
u_{st}	5,6929	5,6929	18,7693	33,2774	43,7551	4,8002
u_{ft}	2,0181	2,0181	24,9858			
5% c.v.	11,07	11,07	5,99	40,11	9,49	9,49

Q and Q2 are Ljung-Box test statistics. The Ljung-Box test statistic is asymptotically chi-square distributed with degrees of freedom equal to the number of autocorrelations. The null hypothesis is that no autocorrelation is present up to order k (where k in this case is 5).

Normality is the Doornik and Hansen (1994) test for normality, distributed as chi-square(2). Normality* is the bivariate test, distributed as chi-square(4).

LM is the Breusch-Godfrey test for autocorrelation of order 1.

Red colouring denotes rejection of the null hypothesis.

When autocorrelation is present in the residuals of the estimated models, the estimated standard errors of the coefficients will be incorrect. Inferences about the coefficients will thus be invalid.

When testing for normality, we find normality for the 1-Month Price Series, but not for the other two. The results are the same for the univariate and bivariate test for normality. Non-normality might be a problem when making inferences about coefficients. This is because the calculated test statistics might not be asymptotically distributed as assumed by the different tests, and inferences will then be invalid.

White's test show no signs of heteroscedasticity for any of the three series.
Based on the tests, inferences about the coefficients from Equation (6.4) should be valid, but inferences about the coefficients from Equation (6.5) and (6.6) might not be. The main problem is probably the small data sample, which can only be fully remedied by waiting a few years and performing the tests with more observations in the data sample. Despite of these problems, we will perform hypothesis tests on all three series. We will, however, be very cautious about making conclusions concerning the two- and three-month series.

6.3.2 Results from the Tests of the Unbiasedness Hypothesis

In this chapter we perform hypothesis tests to find out if one, two and three months to maturity futures contracts are unbiased predictors of the spot price at maturity. In able to do this, we have to perform a joint test on the null hypothesis that $\beta_1 = 0$ and $\beta_2 = -1$ in the cointegrating vector. We will also test the null hypothesises that $\beta_1 = 0$ and $\beta_2 = -1$ separately.

To perform the tests, we estimate the VECMs with and without restrictions on the cointegrating vector. The maximized value of the (Gaussian) log-likelihood function is noted and then used as input in the test statistic. The LR test statistic is then given as $-2(L_r - L_u)$, where L_r and L_u are the maximized values of the log-likelihood function of the restricted and unrestricted regressions, respectively. The test statistic is asymptotically chi-square distributed with degrees of freedom equal to the number of restrictions placed on the cointegrating vector.

In Table 6.5, the results from the hypothesis tests are shown. We see that the hypothesis $\beta_1 = 0$ is not rejected for any of the three series at the 5% significance level. For the one-month price series, however, the null hypothesis is rejected at the 10% significance level.

The hypothesis $\beta_2 = -1$ is not rejected for any of the three series at the 5% significance level. Also here, the null hypothesis is rejected for the one-month price series at the 10% significance level.

Moving on to the null hypothesis that $\beta_1 = 0$ and $\beta_2 = -1$ simultaneously, we see that the hypothesis is not rejected for any of the three series at the 5% significance level. This was expected as none of the separate tests were rejected. For the one-month price series, the null hypothesis is yet again rejected at the 10% significance level. For the two- and three-month price series, the unbiasedness hypothesis is far from being rejected.

Table 6.5: Testing the Unbiasedness Hypothesis

	1-Month Price Series Hypothesis Tests on β´		
	H_0: $\beta_1 = 0$	H_0: $\beta_2 = -1$	H_0: $\beta_1 = 0$ and $\beta_2 = -1$
	3,4782	3,3256	5,6118
c.v. 5%	[3,84]	[3,84]	[5,99]
c.v. 10%	[2,71]	[2,71]	[4,61]

	2-Month Price Series Hypothesis Tests on β´		
	H_0: $\beta_1 = 0$	H_0: $\beta_2 = -1$	H_0: $\beta_1 = 0$ and $\beta_2 = -1$
	0,0916	0,1000	0,2796
c.v. 5%	[3,84]	[3,84]	[5,99]
c.v. 10%	[2,71]	[2,71]	[4,61]

	3-Month Price Series Hypothesis Tests on β´		
	H_0: $\beta_1 = 0$	H_0: $\beta_2 = -1$	H_0: $\beta_1 = 0$ and $\beta_2 = -1$
	0,0004	0,0001	0,1682
c.v. 5%	[3,84]	[3,84]	[5,99]
c.v. 10%	[2,71]	[2,71]	[4,61]

Red colouring denotes rejection of the null hypothesis.

The results in Table 6.5 support the hypothesis that one, two and three months to maturity futures contracts are unbiased predictors of the spot price at maturity. The hypothesis of unbiasedness is least significant for one month to maturity futures contracts and most significant for three months to maturity contracts.

We remember from Table 6.4 that the model for the one-month price series was found to be well specified. Inferences about the model parameters should therefore be valid. A correction to the test-statistics for small-sample problems could perhaps improve the properties of the tests, but we have not performed such a correction in this thesis.

The models for the two- and three-month prices series were not found to be well specified, however. The residual diagnostics in Table 6.4, showed that the residuals from the two models were not normally distributed. In addition, we found presence of autocorrelation for the two-month price series model. These circumstances might invalidate inferences about the parameters in the models, and tests of the unbiasedness hypothesis for the two- and three-month price series might thus be invalid.

We therefore conclude that one month to maturity futures contracts appear to be unbiased predictors of the spot price at maturity. Hedgers can then use the market efficiently without paying any risk premium. We are sceptical to the results from the tests on two and three months to maturity contracts and choose therefore not to draw any conclusions.

7. Summary and Conclusions

In this thesis we have tested the hedging effectiveness and the unbiasedness hypothesis of the IMAREX PM4TC freight futures contract. We used both constant and time-varying models to calculate the hedge ratios and test the hedging performance. The time-varying hedge ratios calculated from a VAR-GARCH model were shown to be superior in four out of five cases. The constant hedge ratios and most of the time-varying hedge ratios were found to be below one, implying that a naive hedge is never recommendable.

For the four routes included in the PM4TC contract, the hedging performance were found to range from 29,50% for route P3A to 31,78% for route P1A. For the average of the four routes, the hedging effectiveness was found to be 34,26%. When using the PM4TC freight futures contract in a continuously rolling hedge, this is therefore the greatest variance reduction that can be expected. In practice, however, a shipowner will often use the PM4TC contract to hedge a different route than the four T/C routes. His basis risk will then be even larger and the hedging effectiveness is thus expected to be lower.

The calculated hedging performance is much lower than what is found in other futures markets. However, we are not surprised by the results, as other studies have found the hedging performance of freight futures to be in the same range.[41] In the shipping market there are, however, no better alternatives for hedging; FFA contracts will only offer comparable variance reduction.

The liquidity of the futures contracts by themselves is not much better than for the BIFFEX. However, we argued that the FFA contracts and freight futures are so similar in specifications that it really is the total liquidity in the market that is of essence regarding efficient pricing. In addition, IMAREX has had a steadily growing customer base and trading volumes. We therefore believe that IMAREX will not be brought down by low liquidity.

The unbiasedness hypothesis states that the futures prices should give an unbiased prediction of future spot prices. The tests in this study found that one month to maturity futures contracts are unbiased predictors of the spot price at maturity. The same result was indicated by the

---

[41] See for example Kavussanos and Nomikos (2000b).

tests of two and three months to maturity contracts, but due to a small data sample and residual diagnostics problems, we have not drawn any conclusions. The tests should therefore be performed again in a few years, when more observations are available.

These findings imply that a hedger can trade in one month to maturity futures contracts without paying a risk premium. A rolling hedge can as a result be executed efficiently. When deciding how often to roll the hedge over, there will be a trade-off between paying more transactions costs and not being optimally hedged at all times. Further studies could be performed to identify how often a hedge should be rolled over to maximise the utility in the trade-off situation.

Another implication of the findings is that all market participants can use the futures prices as a forecast of future spot prices, even if they do not actually trade in futures themselves. This is the price discovery role of the futures contracts. Physical market decisions can thus be guided by examining the futures prices. A suggestion to further studies is to test if the futures prices are able to forecast future spot prices better than forecasts from time series models based on present and past spot prices.

Readings

Abeele, Philippe Van Den (2005): *FFA – The Principles of Forward Freight Agreements*. BIMCO Centenary. May 2005

Alizadeh, Amir H. and Nikos K. Nomikos (2002): *Risk Management in the Shipping Industry: Theory and Practice*. (Chapter 31 in Grammenos, Costas Th. (ed.): The Handbook of Maritime Economics and Business. Lloyds of London Press)

Asteriou, Dimitrios (2006): *Applied Econometrics: A Moders Approach Using EViews and Microfit*. Palgrave Macmillan, New York

Baillie, R. T. and R. J. Myers (1991): *Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge*. (In: Journal of Applied Econometrics. Vol. 6, pp.109-124)

Baltic Exchange, The <http://www.balticexchange.com/ > 12.02.2007

Batchelor, Roy, Amir H. Alizadeh and Ilias D. Visvikis (2007): *Forecasting Spot and Forward Prices in the International Freight Market*. (In: International Journal of Forecasting. Vol. 23, No.1, pp. 101-114)

Benstock, M, and Vergottis, A. (1989): *An Econometric Model of the World Market for Dry Cargo Freight and Shipping*. (In: Applied Economics. Vol. 21, pp. 339-356)

Bera, Ankil K. and Carlos M. Jarque (1980): *Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals*. (In: Economics Letters. Vol. 6, pp. 255-259)

Bollerslev, Tim (1986): *Generalized Autoregressive Conditional Heteroskedasticity*. (In: Journal of Econometrics. Vol. 31, pp. 307-327)

Bollerslev, Tim (1990): *Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach*. (In: Review of Economics and Statistics. Vol. 72, pp. 498-505)

Bond, Gary E. and Stanley R. Thompson (1986): *Optimal Commodity Hedging within the Capital Asset Pricing Model*. (In: Journal of Futures Markets. 1986, Vol. 6, No. 3, pp. 421-432)

Brenner, Robin and Kenneth F. Kroner (1995): *Arbitrage, Cointegration and Testing the Unbiasedness Hypothesis in Financial Markets*. (In: Journal of Financial & Quantitative Analysis. Vol. 30, No. 1, pp. 23-42)

Breusch, T. S. (1979): *Testing for Autocorrelation in Dynamic Linear Models*. (In: Australian Economic Papers. Vol. 17, pp. 334-355)

Brooks, Chris (2002): *Introductory Econometrics for Finance*. Cambridge University Press. Cambridge

Bøe, Mikal (2005): *Dry Bulk Shipping Derivatives – A Screen Market in the Making*. (In: Commodities Now. December 2005, pp. 89-91)

Clarkson (1991): *The Clarkson VLCC Quality Survey*. Clarkson Research Studies Ltd., London

Clarksons Securities Limited, <http://www.clarksonsecurities.com> 12.02.2007

Dickey, David A., and Fuller, Wayne A. (1979): *Distribution of Estimators for Time Series Regressions with a Unit Root*. (In: Journal of the American Statistical Association. Vol. 74, pp. 427-31)

Dickey, David A. and Wayne A. Fuller (1981): *Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root*. (In: Econometrica. Vol. 49, No. 4, (July, 1981), pp. 1057-1072)

Doornik, J. A. and H. Hansen (1994): *A Practical Test for Univariate and Multivariate Normality*. Working Paper, Nuffield, Oxford

Ederington, Louis H. (1979): *The Hedging Performance of the New Futures Markets*. (In: The Journal of Finance. Vol. 34, No. 1, (March, 1979), pp. 157-170)

Engle, Robert F. (1982): *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*. (In: Econometrica. Vol. 50, No. 4. (July, 1982), pp. 987-1008)

Engle, Robert F. and C. W. J. Granger (1987): *Co-Integration and Error Correction: Representation, Estimation, and Testing*. (In: Econometrica. Vol. 55, No. 2, (March, 1987), pp. 251-276)

Engle, Robert F. and Kenneth F. Kroner (1995): *Multivariate Simultaneous Generalized ARCH*. (In: Economic Theory. Vol. 11, pp. 122-150)

EViews (2005): *EViews 5.1 Help File*

Fama, Eugene F. (1991): *Efficient Capital Markets: II*. (In: The Journal of Finance. Vol. 46, No. 5, (December, 1991), pp. 1575-1617)

French, Kenneth R. (1983): *A comparison of Futures and Forward Prices. (In: Journal of Financial Economics*. Vol. 12, No. 3, (November, 1983), pp. 311-342)

Gagnon, L. and G. Lypny (1995): *Hedging Short-Term Interest Risk under Time-Varying Distributions*. (In: Journal of Futures Markets. Vol. 15, pp. 767-783)

Garcia, P., J. Roh and M. Leuthold (1995): *Simultaneously Determined, Time-Varying Hedge Ratios in the Soybean Complex*. (In: Applied Economics. Vol. 27, pp. 1127-1134)

Glen, David and Brendan Martin (2002): *The Tanker Market: Current Structure and Economic Analysis*. (Chapter 12 in Grammenos, Costas Th. (ed.): The Handbook of Maritime Economics and Business. Lloyds of London Press)

Godfrey, L.G. (1978): *Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables*. (In: Econometrica. Vol. 46, pp. 1293-1302)

Gonzalo, J. (1989): *Comparison of Five Alternative Methods of Estimating Long Run Equilibrium Relationships*. Working Paper. University of California at San Diego

Gosh, A. (1993a): *Hedging With Stock Index Futures: Estimation and Forecasting with Error Correction Model*. (In: Journal of Futures Markets. Vol. 13, pp. 743-752)

Gosh, A. (1993b): *Cointegration and Error Correction Models: Intertemporal Causality between Index and Futures Prices*. (In: Journal of Futures Markets. Vol. 13, pp. 193-198)

Gretton, Jim (2002): *Portfolio Management of Tanker Freight Risk*. Intertanko's Rotterdam Tanker Event. April 2002

Haigh, Michael S. (2000): *Cointegration, Unbiased Expectations and Forecasting in the BIFFEX Freight Futures Market*. (In: The Journal of Futures Markets. Vol. 20, No. 6, pp. 545-571)

Haigh, Michael S. and Matthew T. Holt (2002): *Hedging Foreign Currency, Freight and Commodity Futures Portfolios: A Note*. The University of Maryland (Department of Agricultural and Resource Economics), College Park

Hamilton, James D. (1994): *Time Series Analysis*. Princeton University Press. Princeton, New Jersey

Herbst, A., D. Kare and J. Marshall (1989): *A Time-varying, Convergence Adjusted, Minimum Risk Futures Hedge Ratio*. (In: Advances in Futures and Options Research. Vol. 6, pp. 137-155)

Hull, John C. (2003): *Options, Futures and Other Derivatives*. 5th International edition. Pearson Education.

IMAREX, <www.imarex.com> 12.02.2007

Johansen, Søren (1988): *Statistical analysis of cointegration vectors*. (In: Journal of Economic Dynamics and Control. Vol. 12, No. 2-3, (June-September 1988), pp. 231-254)

Johansen, Søren and Katarina Juselius (1990): *Maximum Likelihood Estimation and Inference on Cointegration – With Applications to the Demand for Money. (In: Oxford Bulletin of Economics and Statistics*. Vol. 52, No. 2, pp. 169-210)

Johansen, Søren (1992): *Determination of Cointegration Rank in the Presence of a Linear Trend*. (In: Oxford Bulletin of Economics and statistics. Vol. 54, pp. 383-397)

Kavussanos, Manolis G. and Nikos K. Nomikos (1999): *The Forward Pricing Function of the Shipping Freight Futures Market*. (In: The Journal of Futures Markets. Vol. 19, No. 3, (May, 1999), pp. 353-376)

Kavussanos, Manolis G. and Nikos K. Nomikos (2000a): *Constant vs. Time-varying Hedge Ratios and Hedging Efficiency in the BIFFEX Market*. (In: Transportation Research Part E 36, pp. 229-248)

Kavussanos, Manolis G. and Nikos K. Nomikos (2000b): *Hedging in the Freight Futures Market*. (In: Journal of Derivatives. Vol.8, No. 1 (Fall, 2000), pp. 41-58)

Kavussanos, Manolis G. and Nikos K. Nomikos (2000c): *Futures Hedging Effectiveness when the Composition of the Underlying Asset Changes; the Case of BIFFEX Contract*. (In: The Journal of Futures Markets. Vol. 20, No. 8, pp. 775-801)

Kavussanos, Manolis G. and Nikos K. Nomikos (2001): *Price Discovery, Causality and Forecasting in the Freight Futures Market*. EFMA 2001 Lugano Meetings, Available at SSRN: <http://ssrn.com/abstract=271072> (23.01.07)

Kavussanos, Manolis G. (2002): *Business Risk Measurement in the Cargo Carrying Sector of the Shipping Industry*. (Chapter 30 in Grammenos, Costas Th. (ed.): The Handbook of Maritime Economics and Business. Lloyds of London Press)

Kavussanos, Manolis G. and Ilias D. Visvikis (2004): *Market Interactions in Returns and Volatilities between Spot and Forward Shipping Freight Markets*. (In: Journal of Banking & Finance. Vol. 28, pp. 2015–2049)

Kavussanos, Manolis G., David Menachof and Ilias D. Visvikis (2004): *The Unbiasedness Hypothesis in the Freight Forward Market: Evidence from Cointegration Tests*. (In: Review of Derivatives Research. Vol. 7, No. 3, pp. 241–266)

Kavussanos, Manolis G. and Ilias D. Visvikis (2006a): *Shipping Freight Derivatives - A Survey of Recent Evidence*. (In: Maritime Policy & Management. Vol. 33, No. 3, (July 2006), pp. 233-255)

Kavussanos, Manolis G. and Ilias D. Visvikis (2006b): *Derivatives and Risk Management in Shipping*. Witherbys Publishing, London

Kroner, K and J. Sultan (1993): *Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures*. (In: Journal of Financial and Quantitative Analysis. Vol. 28, pp. 535-551)

Lee, Tae-Hwy (1994): *Spread and Volatility in Spot and Forward Exchange Rates*. (In: Journal of International Money and Finance. Vol. 13, No. 3, pp. 375-383)

LIFFE, <www.euronext.com> 26.04.2007

Ljung, M. and G. Box (1978): *On a Measure of Lack of Fit in Time Series Models*. (In: Biometrika. Vol. 65, pp. 297-303)

Margrabe, W. (1976): *A Theory of Forward and Futures Prices*. Working-paper. The Wharton School, University of Pennsylvania, Philadelphia

Markowitz, Harry M. (1952): *Portfolio Selection* (In: Journal of Finance. Vol. 7, No. 1, pp. 77-91)

McDonald, Robert L. (2003): *Derivatives Markets*. International edition. Pearson Education

Myers, R. (1991): *Estimating Time-Varying Optimal Hedge Ratios on Futures Markets*. (In: Journal of Futures Markets. Vol. 11, pp. 39-53)

NOS, <www.nos.no> 12.02.2007

Schwarz, Gideon (1978): *Estimating the Dimension of a Model*. (In: Annals of Statistics. Vol. 6, No. 2, pp. 461-464)

Skjetne, Runar A. (2005): *Applied Financial Risk Management for the Shipping Industry Using IMAREX Derivatives*. Master thesis. Norwegian School of Economics and Business Administration (NHH), Bergen

Sonin, Adam (2005): *Managing Risk with Forward Freight Agreements*. (In: Commodities Now (June, 2005), pp. 1-2)

Stoll, H.R. and Robert E. Whaley (1993): *Futures and Options: Theory and Applications*. South-Western Publishing Co, Cincinnati, Ohio

Stopford, Martin (1997): *Maritime Economics*. Routledge

Strandenes, Siri P. (2002): *Economics of the Markets for Ships*. (Chapter 9 in Grammenos, Costas Th. (ed.): The Handbook of Maritime Economics and Business. Lloyds of London Press)

Sutcliffe, Charles and Nikolaos Zacharatos (2002): *Is the Forward Rate for the Greek Drachma Unbiased? A VECM Analysis with both Overlapping and Non-Overlapping Data*. (In: Journal of Financial Management and Analysis. Vol. 15, No. 1, pp. 27-37)

Thuong, L. T. and Visscher, S. L. (1990): *The Hedging Effectiveness of Dry-Bulk Freight Rate Futures.* (In: Transportation Journal. Vol. 29, pp. 58-65)

White, H. (1980): *A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity*. (In: Econometria. Vol. 48, pp. 817-838)

Wikipedia, <www.wikipedia.com> 24.04.2007

Working, Holbrook (1953): *Futures Trading and Hedging*. (In: The American Economic Review, Vol. 43, No. 3. (June, 1953), pp. 314-343)

Appendix 1: IMAREX Freight Futures Liquidity

Table A1.1: IMAREX Trade Volume
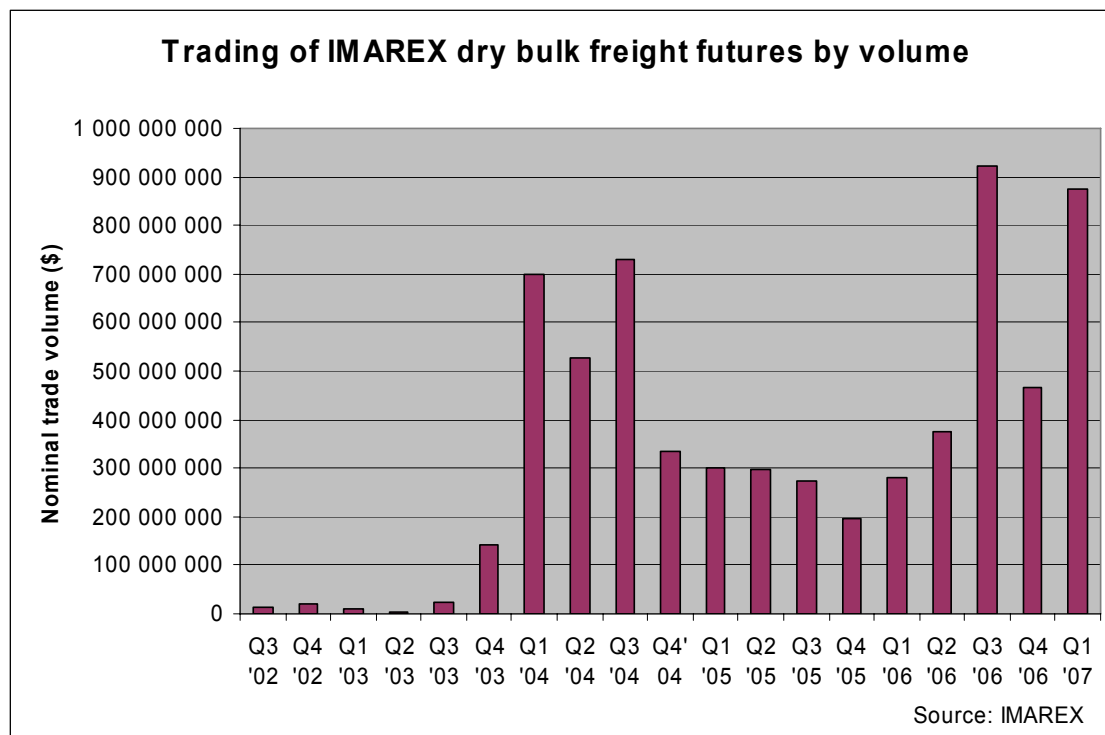
Table A1.2: IMAREX Number of Trades of Dry Bulk Freight Futures
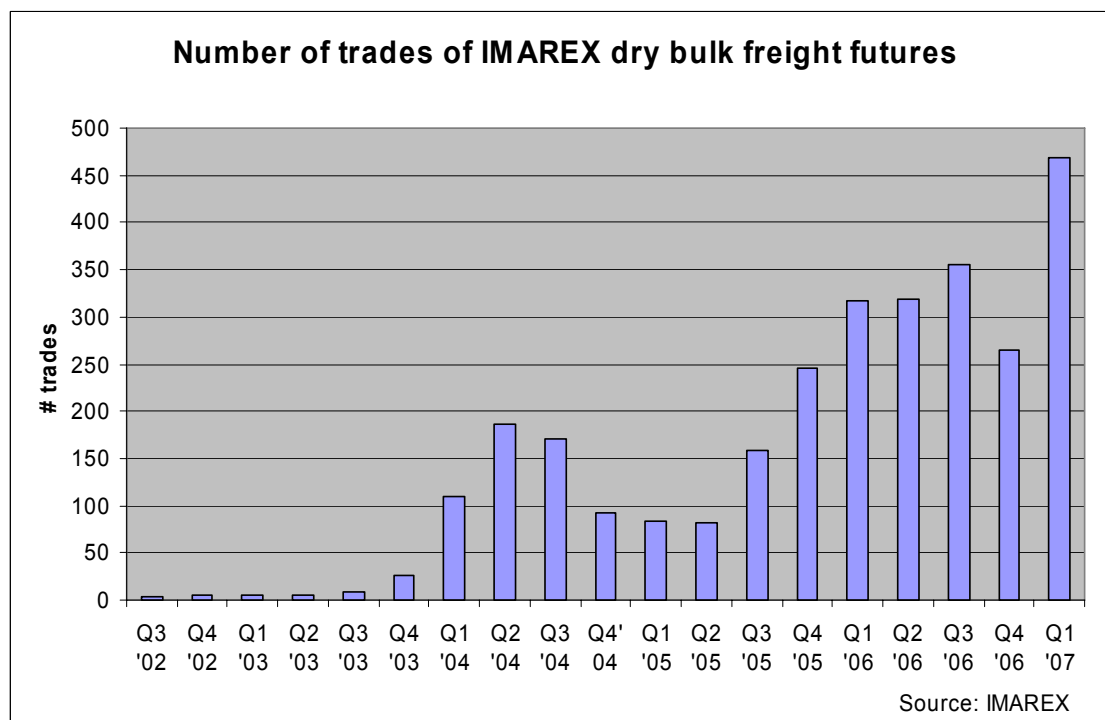
Appendix 2: IMAREX's Freight Futures Descriptions

The following information is found in Appendix 5 to the IMAREX Rulebook as per 12 March 2007.

A.2.1 How Closing Prices are Set
Closing Price is set to:

The best bid, if last price < best bid

The best offer, if last price > best offer

or else use Last Price.

If no Last Price exists, the previous Closing Price is used in the formulae above.

A.2.2 PM4TC Product Specification

Lot size
1 lot = 1 day

Price quotation
USD/day

Minimum price fluctuation
USD 25.00

Contract value
Lots x Lot size x Price

Delivery period
Month: First Index Day of month to last Index Day of month

Quarter: First Index Day of quarter to last Index Day of quarter

Half Year: First Index Day of half year to last Index Day of half year

Year: First Index Day of year to last Index Day of year

There is no cascading of Products.

Product structure
Month: 4 consecutive Month Contracts. Last trading day is the 20th of the month in question. If this date is a non-trading day, the Last trading day is defined as the nearest trading day prior to the 20th.

Quarter: 4 consecutive quarterly Contracts. When 1/3 of the quarter Contract is settled/delivered (after 1 month), the Contract will be taken off screen and a new quarter Contract will be introduced.

Half Year: 2 consecutive half-yearly Contracts. When 1/3 of the half year Contract is settled/delivered (after 2 months), the half year Contract will be taken off screen and a new half year Contract will be
introduced.

Year: 3 yearly Contracts commencing January each year and 1 yearly Contract commencing July each year. When 1/3 of the yearly Contract is settled/delivered (after 4 months), the yearly Contract will be taken off screen and a new yearly Contract will be introduced.

Last trading day on screen
Month: Last trading day is the 20th of the month in question. If this date is a non-trading day, the last trading day is defined as the nearest trading day prior to the 20th.

Quarter: When 1/3 of the Contact is delivered. Last trading day in the first month of the Contract.

Half Year: When 1/3 of the Contract is delivered. Last trading day in the second month of the Contract.

Year: When 1/3 of the Contract is delivered. Last trading day in the forth month of the Contract.

Last day for reporting for clearing
Last day of the Delivery period for the relevant Product

Final settlement day
Last settlement day in the Delivery Period.

Settlement price
The average of Spot Prices for the relevant Underlying Product in the Delivery Period monthly weighted over the number of calendar days in each month.

Minimum lots per contract
Trayport:
1 lot in all Products

Clearing:
0.01 lot in all Products

Appendix 3: Lag Length Selection and Diagnostics of the LRM AR(p)

To choose which lags to include in the model, we start by estimating the following regression model for p = 1, 2, 3, …, 12:

$$\Delta S_t = \alpha_0 + \alpha_1 \Delta F_t + u_t$$

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + ... + \phi_p u_{t-p} + \varepsilon_t; \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \tag{A3.1}$$

where $\phi_1, \phi_2, ..., \phi_p$ are parameters and ε_t is an error term. The remaining notation is as explained for Equation (4.9).

Then we exclude the least significant AR terms one by one until the AR terms are all significant at the 5% significance level. This process is called "data mining" and might invalidate inferences about the parameters. As we are not going to make inferences, but just want to remove the autocorrelation, "data mining" is not a big problem. Residual diagnostics of the chosen models are shown in Table A.3.1.

Table A3.1: Residual Diagnostics on the Estimated Models

	P1A	P2A	P3A	P4	AVG4TC	Critical values 5 %	1 %
The AR terms in the model	2, 10, 12	10	2, 9	2	2		
SIC for the estimated model	-2,34633	-2,44224	-1,90395	-2,00627	-2,37251		
White's test	0,38025	2,14144	2,92491	5,18781	2,21925	5,99	9,21
Breusch-Godfrey (1)	0,25664	0,01404	1,22463	0,02362	0,52828	3,84	6,64
Breusch-Godfrey (2)	1,04768	1,91367	1,31400	0,02426	1,69349	5,99	9,21
Breusch-Godfrey (4)	1,24835	5,42176	2,64115	1,27317	3,18994	9,49	13,28
Breusch-Godfrey (12)	8,45544	15,78157	5,54277	8,27279	12,65300	21,03	26,22
Breusch-Godfrey (26)	20,67231	32,37250	17,94368	14,85885	21,35236	38,89	45,64
Breusch-Godfrey (52)	36,21730	44,59626	40,58233	39,52820	43,32549	69,83	78,62
Skewness	-0,40948	-0,36572	-0,23000	-0,02376	-0,26977		
Kurtosis	3,61008	3,89455	3,23639	4,33835	3,70828		
Jarque-Bera	6,56158	8,51203	1,71626	12,03094	5,31810	5,99	9,21
ARCH (1)	4,30822	0,75668	7,26407	12,29674	5,97077	3,84	6,64
ARCH (2)	5,22006	2,90435	7,13932	12,36057	6,11562	5,99	9,21
ARCH (5)	17,70956	19,16299	7,54703	12,66307	8,87262	11,07	15,09

For White's test we have shown the X^2 test values. Two degrees of freedom for all routes.

Numbers in parenthesis indicate the number of lags in the specific test.

Numbers written in red are above the critical value. The null hypothesis is then rejected.

As we see from the table, we do not find evidence of autocorrelation. However, we still have ARCH-effects. Modelling the second moments with an ARCH model is therefore necessary.

Appendix 4: Statistics for the VAR-GARCH Model

Table A4.1: Schwarz Information Criterion for Choosing the Optimal Lag Length

Lag length	P1A	P2A	P3A	P4	AVG4TC
0	0,15921	0,41501	0,11109	0,20060	-0,16925
1	-4,53103	-4,56640	-3,99943	-4,07038	-4,48380
2	-4,73888	-4,73254	-4,17147	-4,30149	-4,71305
3	-4,63599	-4,62688	-4,11126	-4,25499	-4,64076
4	-4,52680	-4,52208	-3,99653	-4,13113	-4,52983
5	-4,40315	-4,41866	-3,88587	-4,01775	-4,41600
6	-4,31390	-4,33255	-3,76726	-3,90204	-4,30623
7	-4,22969	-4,25685	-3,65969	-3,80284	-4,20553
8	-4,10272	-4,14803	-3,53938	-3,68481	-4,08980
9	-3,98315	-4,03191	-3,46803	-3,59151	-4,00356
10	-3,92508	-3,98362	-3,45443	-3,56331	-3,99110
11	-3,80584	-3,86354	-3,32584	-3,43495	-3,87012
12	-3,66682	-3,73136	-3,19625	-3,32063	-3,74111

Green colouring shows the minimum value of the Schwarz criterion
and thus the chosen number of lags.

Table A4.2: Results from Johansen's Test for Cointegration

P1A (trace)					P1A (max)			
No CE	Model 2	Model 3	Model 4		No CE	Model 2	Model 3	Model 4
0	15,1509	14,9291	22,1996		0	12,1327	12,1224	18,6065
c.v.	20,2618	15,4947	25,8721		c.v.	15,8921	14,2646	19,3870
1	3,0183	2,8068	3,5931		1	3,0183	2,8068	3,5931
c.v.	9,1645	3,8415	12,5180		c.v.	9,1645	3,8415	12,5180

P2A (trace)					P2A (max)			
No CE	Model 2	Model 3	Model 4		No CE	Model 2	Model 3	Model 4
0	17,7171	17,5195	26,1158		0	13,5042	13,5040	21,4830
c.v.	20,2618	15,4947	25,8721		c.v.	15,8921	14,2646	19,3870
1	4,2130	4,0156	4,6328		1	4,2130	4,0156	4,6328
c.v.	9,1645	3,8415	12,5180		c.v.	9,1645	3,8415	12,5180

P3A (trace)					P3A (max)			
No CE	Model 2	Model 3	Model 4		No CE	Model 2	Model 3	Model 4
0	9,1440	8,7737	12,3827		0	7,3684	7,0803	9,0279
c.v.	20,2618	15,4947	25,8721		c.v.	15,8921	14,2646	19,3870
1	1,7756	1,6934	3,3548		1	1,7756	1,6934	3,3548
c.v.	9,1645	3,8415	12,5180		c.v.	9,1645	3,8415	12,5180

P4A (trace)					P4A (max)			
No CE	Model 2	Model 3	Model 4		No CE	Model 2	Model 3	Model 4
0	10,7403	10,3334	12,8137		0	7,9857	7,6147	8,7014
c.v.	20,2618	15,4947	25,8721		c.v.	15,8921	14,2646	19,3870
1	2,7546	2,7187	4,1123		1	2,7546	2,7187	4,1123
c.v.	9,1645	3,8415	12,5180		c.v.	9,1645	3,8415	12,5180

AVG4TC (trace)					AVG4TC (max)			
No CE	Model 2	Model 3	Model 4		No CE	Model 2	Model 3	Model 4
0	9,6144	9,1920	13,2916		0	7,9051	7,8170	10,3237
c.v.	20,2618	15,4947	25,8721		c.v.	15,8921	14,2646	19,3870
1	1,7093	1,3750	2,9678		1	1,7093	1,3750	2,9678
c.v.	9,1645	3,8415	12,5180		c.v.	9,1645	3,8415	12,5180

2 lags are included in the test.

Selection of the appropriate model is done by applying the Pantula Principle.
The model selected is coloured green.

Red colour denotes rejection of the null hypothesis.

C.V. is short for critical value (at the 5% significance level).

Table A4.3: Maximum Likelihood Estimates of VAR-GARCH Models

	P1A	P2A	P3A	P4	AVG4TC
			Conditional Mean Equation		
c_s	-0,0014 (0,0061)	-0,0014 (0,0060)	-0,0023 (0,0077)	-0,0021 (0,0073)	-0,0021 (0,0062)
$\beta_{s,1}$	-0,0525 (0,1031)	-0,0812 (0,1004)	-0,0978 (0,0976)	-0,0184 (0,0979)	-0,1537 (0,1043)
$\beta_{s,2}$	-0,0023 (0,0851)	-0,0481 (0,0857)	-0,0844 (0,0853)	-0,0668 (0,0831)	-0,0452 (0,0859)
$\theta_{s,1}$	0,4946 (0,0951)	0,4531 (0,0912)	0,5801 (0,1140)	0,5548 (0,1078)	0,5689 (0,0975)
$\theta_{s,2}$	0,0318 (0,0997)	-0,0060 (0,0955)	-0,0531 (0,1226)	-0,0281 (0,1154)	0,0453 (0,1058)
c_f	0,0010 (0,0066)	0,0010 (0,0066)	0,0015 (0,0066)	0,0017 (0,0066)	0,0012 (0,0066)
$\beta_{f,1}$	-0,1790 (0,1119)	-0,1912 (0,1106)	-0,0744 (0,0841)	-0,0623 (0,0885)	-0,1467 (0,1116)
$\beta_{s,2}$	0,0792 (0,0923)	0,0738 (0,0944)	0,1344 (0,0735)	0,1868 (0,0752)	0,1348 (0,0919)
$\theta_{f,1}$	0,0824 (0,1032)	0,0838 (0,1005)	0,0352 (0,0982)	0,0238 (0,0975)	0,0732 (0,1044)
$\theta_{f,2}$	0,0047 (0,1082)	0,0036 (0,1052)	-0,0868 (0,1056)	-0,1189 (0,1043)	-0,0446 (0,1132)
			Conditional Variance Equation		
c_{11}	0,0173 (0,0204)	0,0305 (0,0079)	0,0758 (0,0110)	0,0716 (0,0075)	0,0321 (0,0092)
c_{12}	0,0364 (0,0333)	0,0179 (0,0059)	0,0302 (0,0076)	0,0344 (0,0085)	0,0195 (0,0057)
c_{22}	0,0001 (7,9048)	0,0206 (0,0106)	0,0000 (8,9498)	0,0001 (4,0060)	0,0212 (0,0077)
a_{11}	0,1442 (0,0677)	0,4212 (0,1104)	0,4411 (0,0926)	0,5302 (0,0878)	0,3530 (0,0909)
a_{22}	0,4215 (0,1112)	0,2814 (0,0882)	0,3773 (0,0995)	0,3583 (0,0926)	0,3280 (0,0785)
b_{11}	0,9543 (0,0886)	0,8190 (0,0882)	0,4610 (0,1952)	0,3466 (0,1900)	0,8416 (0,0801)
b_{22}	0,7646 (0,1146)	0,8994 (0,0684)	0,8600 (0,0570)	0,8401 (0,0637)	0,8799 (0,0598)
LL	438,8018	405,991	368,2778	382,6468	410,6859
SIC	-5,262984	-4,852849	-4,381434	-4,561046	-4,911535

Estimation period is 04/02/2004 to 27/02/2007 (N=164).

LL is the maximum value of the log-likelihood function.

Numbers in parenthesis are the standard deviations of the parameter estimates.

Table A4.4: Summary Statistics on the Estimated Hedge Ratios

Route	Model	Mean	STD	ADF(lags)
P1A	CLRM	0,5864	-	
	LRM AR(p)	0,5995	-	
	VAR	0,5920	-	
	VAR-GARCH	0,5693	0,1135	-3,4540 (0)
P2A	CLRM	0,5619	-	
	LRM AR(p)	0,5381	-	
	VAR	0,5564	-	
	VAR-GARCH	0,5471	0,1340	-4,0541 (0)
P3A	CLRM	0,6912	-	
	LRM AR(p)	0,7030	-	
	VAR	0,7229	-	
	VAR-GARCH	0,7481	0,2020	-5,1829 (0)
P4	CLRM	0,6428	-	
	LRM AR(p)	0,6670	-	
	VAR	0,6945	-	
	VAR-GARCH	0,6919	0,1879	-7,1082 (0)
AVG4TC	CLRM	0,6126	-	
	LRM AR(p)	0,6124	-	
	VAR	0,6232	-	
	VAR-GARCH	0,6174	0,1242	-3,9907 (0)

Mean and STD are the mean and standard deviation of the series.

ADF is the augmented Dickey-Fuller test on the level of the series.
An intercept term is included in the ADF regressions. The ADF lag
length is determined by minimising SIC.

Red colouring siginfies non-stationarity.

Table A4.5: Hedging Effectiveness (Alternative Hedge Ratio Calculation)

	Variance				
	P1A	P2A	P3A	P4	AVG4TC
VAR(2)-GARCH	0,005008602	0,004754647	0,008023494	0,00732659	0,005049526

	Variance Reduction				
	P1A	P2A	P3A	P4	AVG4TC
VAR(2)-GARCH	34,21 %	33,84 %	30,06 %	29,93 %	35,76 %

The percentages have been rounded to two decimal places.

Green colouring denotes the route with the highest hedge effectiveness.

Red colouring denotes the route with the lowest hedge effectiveness.

Comments to Table A4.5:

Due to the inclusion of two lags in the VAR-GARCH model, we are unable to calculate a
hedge ratio for the first three weeks of the data sample. This is two weeks more than for the
CLRM. When calculating the hedging effectiveness with Equation (4.31), the question is then

how many observations to include in the variance calculation of the unhedged portfolio. In Table 5.7 we have left out the first two observations of the unhedged portfolio when calculating the hedging effectiveness of the VAR-GARCH model (not when calculating the effectiveness of the other models). There are then 161 observations of both the unhedged and the hedged portfolio in the calculation. In Table A4.5, however, we have included all 163 observations of the unhedged portfolio when calculating the variance. Conversely, all 161 observations of the hedged portfolio are included. Normally, the calculated hedging effectiveness should not differ much from Table 5.7 to Table A4.5, but we see here that it does. This is because the first observation of the price differences is among the most extreme observations in the entire sample.

The variance of the unhedged spot portfolio for route P1A when 161 observations are included is only 96,4% of the variance when all 163 observations are included in the variance calculation. Including all 163 observations in the unhedged portfolio variance and only 161 observations in the VAR-GARCH hedged portfolio will therefore inflate the actual performance of the VAR-GARCH model. The performance shown in Table A4.5 will consequently be overstated.

Leaving out two observations is normally not advisable either. As both the constant hedge ratio and the time-varying hedge ratio is estimated on the same data sample, all observations should intuitively be included when calculating the unhedged portfolio variance as well. Neither of the calculations will therefore give the "correct" answer.

The main objective of the analysis is to find the model that achieves the best hedging performance. It is therefore important that the "true" conclusion of the test is not altered by the explained problems.

In our opinion, the hedging effectiveness will clearly be overstated in Table A4.5. In Table 5.7 we can not say for certain, but the effectiveness is probably around the correct value. We therefore choose to report Table 5.7 in the main text of the thesis.

Appendix 5: How to Perform Out-of-Sample Tests

The tests of the hedging effectiveness in this thesis are in-sample analyses. The conclusion as to whether constant or time-varying hedge ratios are better than the other can not with absolute certainty be used on hedging in practice. In-sample analyses will give a very good indication, but out-of-sample analyses should be performed to support the conclusion.

To perform an out-of-sample analysis, some part of the data sample should be withheld, for example one third of the observations. The remaining data are then used to estimate the conditional models. The one-step-ahead forecasts of the variance and covariance for the VAR-GARCH model from Chapter 5.3.1 are then calculated as follows:

$$E(h_{SF,t+1}|\Omega_t) = c_{11}c_{12} + a_{11}a_{22}u_{st}u_{ft} + b_{11}b_{22}h_{SF,t} \tag{A5.1}$$

$$E(h_{FF,t+1}|\Omega_t) = c_{12}^2 + c_{22}^2 + a_{22}^2 u_{ft}^2 + b_{22}^2 h_{FF,t} \tag{A5.2}$$

The optimal hedge ratio at time t+1 is then given as:

$$E(h_{t+1}^*|\Omega_t) = \frac{E(h_{SF,t+1}|\Omega_t)}{E(h_{FF,t+1}|\Omega_t)} \tag{A5.3}$$

The procedure is repeated the following week with the new observation in the data set. This process is continued for all out-of-sample observations. The return of the time-varying out-of-sample hedge portfolio is then calculated, and, finally, its variance is compared to the variance of the unhedged portfolio to find the hedging performance. This hedging performance is then compared to an out-of-sample performance for the constant hedge ratio models.

When hedging in practice, the same procedure as described here is used to find the optimal hedge ratio at each given time.

Appendix 6: Testing the Unbiasedness Hypothesis

Table A6.1: Schwarz Information Criterion for Choosing the Optimal Lag Length

Lag length	1-month	2-month	3-month
0	-1,377898	-0,022644	0,458047
1	-4,518415	-3,095536	-2,143075
2	-4,548495	-2,914495	-2,684785
3	-4,283253	-3,188721	-2,440487

Green colouring shows the minimum value of the SIC.

Comments to Table A6.1:

We see that on the basis of the Schwarz Criterion, 2, 3 and 2 lags should be used for the 1, 2 and 3 month series respectively. When specifying a VECM, an observation is lost for each lag that is included. Due to the very low number of observations, we want to loose as few as possible. In Table A6.1, we see that the value of the SIC is about the same for 1, 2 and 3 lags for all series. We therefore choose to include 1 lag in our tests.

Inclusion of too few lags might result in autocorrelation in the residuals, which again might invalidate inferences about the model parameters. Inferences might not be very good anyway, due to the small data sample, so this will be a trade-off between autocorrelation and even fewer observations.

Table A6.2: Results from Johansen's Test for Cointegration

1-month (trace)				1-month (max)			
No CE	Model 2	Model 3	Model 4	No CE	Model 2	Model 3	Model 4
0	29,0783	28,0419	48,2784	0	27,7357	27,6745	40,7987
c.v.	20,2618	15,4947	25,8721	c.v.	15,8921	14,2646	19,3870
1	1,3426	0,3674	7,4798	1	1,3426	0,3674	7,4798
c.v.	9,1645	3,8415	12,5180	c.v.	9,1645	3,8415	12,5180

2-month (trace)				2-month (max)			
No CE	Model 2	Model 3	Model 4	No CE	Model 2	Model 3	Model 4
0	30,7873	25,8565	47,5755	0	25,8886	25,7782	45,4348
c.v.	20,2618	15,4947	25,8721	c.v.	15,8921	14,2646	19,3870
1	4,8987	0,0784	2,1408	1	4,8987	0,0784	2,1408
c.v.	9,1645	3,8415	12,5180	c.v.	9,1645	3,8415	12,5180

3-month (trace)				3-month (max)			
No CE	Model 2	Model 3	Model 4	No CE	Model 2	Model 3	Model 4
0	25,8191	22,4012	36,1369	0	22,6603	22,3853	33,2771
c.v.	20,2618	15,4947	25,8721	c.v.	15,8921	14,2646	19,3870
1	3,1588	0,0159	2,8598	1	3,1588	0,0159	2,8598
c.v.	9,1645	3,8415	12,5180	c.v.	9,1645	3,8415	12,5180

1 lag is included in the test.

Selection of the appropriate model is done by applying the Pantula Principle.
The model selected is coloured green.

Red colour denotes rejection of the null hypothesis.

C.V. is short for critical value (at the 5% significance level).

Comments to Table A6.2:

In Table A6.2, the trace and max statistics are shown for the three series. The Pantula principle is used to choose the correct assumptions for the test. We see that model 2 is chosen in all cases. Model 2 has an intercept, but no trend in the cointegrating equation and no intercept or trend in the VAR. Cointegration is found for all three series. In small samples, the Johansen test will show evidence of cointegration too often. As the test statistics in Table A6.2 are all much higher than the critical value, this is probably not a problem.

In Table A6.2, one lag is included in the tests. We also tried to include the number of lags as implied by the SICs from Table A6.1 in the tests. Then, cointegration was not found for series 1 and the evidence was inconclusive for series 2. As economic theory hypothesises that futures and spot prices should be cointegrated, we attribute the finding of no cointegration in this case to the low number of observations in the data sample. Cointegration was found for series 3, but there were more signs of autocorrelation than when only 1 lag was included. The only good way to perform the tests without these problems is to wait a few years and test again with more observations included.

Thomas Samdal Rasmussen
Eivind Tversland

NHH

2007