UNIVERSITY OF PIRAEUS

DEPARTMENT OF BANKING & FINANCIAL MANAGEMENT

Setting Margins for freight futures: The IMAREX Case

Master Dissertation by:

Parianou Eirini MXAN/1129

Supervisor:

Skiadopoulos Georgios

Examination Committee:

Anthropelos Michael

Antzoulatos Aggelos

Skiadopoulos Georgios

PIRAEUS, FEBRUARY 2013

Setting margins for Freight Futures: The IMAREX case

The margin system is the clearing house's first line of defense against default risk. Various methods are used in setting margin requirements. One of the most valuable is Value at Risk which is the used tool in this dissertation and is applied in three FFA contracts from the Imarex Exchange. Backtesting is conducted through three statistical tests proposed by Christoffersen P.

Keywords: Imarex, Value at risk, margin, FFA, backtesting

Acknowledgements

I owe my gratitude to all these people who helped me accomplishing the demanding task of my dissertation.

My greatest gratitude is to my supervisor George Skiadopoulos for his guidance, practical comments and his always immediate answers to all my questions.

I would also like to thank all my teachers for educating me and providing me with all the supplies to fulfill my goals.

Last but not least, I would also like to thank George, my parents and all my friends for their love, patience and encouragement during this endeavor.

1. INTRODUCTION

The main characteristics of the shipping business are its highly volatile freight rates, seasonality, strong business cycles and capital intensiveness. 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 (Rasmussen and Tversland, 2007). Freight derivatives are financial instruments for trading in future levels of freight rates, for dry bulk carriers, tankers and containerships. These instruments are settled against various freight rate indices published by the Baltic Exchange (for Dry and most Wet contracts) & Platt's (Asian Wet contracts). Market participants that actively manage risk with such tools will then be less exposed to market volatility than they would otherwise be. More specifically, they can use them as a hedging tool, which means a tool with which they can minimize the portfolio price variance.

In order to face with uncertainties pertaining to the possibility of participants' failure to meet their obligations, derivative exchanges, along with clearing houses, use margins and therefore manage to decrease default risk. Clearing houses are financial institutions that provide netting and settlement for financial and commodities derivatives and security transactions. They do so by requiring margin deposits, by evaluating trades and collateral and by guarantying the cover of a probable loss using funds. Once a trade has been executed, it can be handed over to the clearing house, which stands between the two traders' clearing (or member) firms and assumes the legal counterparty risk for the trade.

Margins are significant tools in dealing with the default risk during each derivative transaction. In this point, exposition of the meanings of margin and exchanges' processes would be essential. When a participant enters a freight futures contract on an exchange, such as Imarex, the clearing house for that exchange, in our case Nos Clearing House, acts like counterparty to every contract. Imarex, and every futures exchange, acts as an intermediary and minimize the risk of default by either party. Parties can be ship-owners and operators, oil companies, trading companies, and grain houses using freight futures as tools for managing freight rate risk. The clearing house sets the initial amount, initial margin, and the exchange offers this initial amount of cash. At the end of every trading day, the exchange checks the freight futures price. If it has gone up, the exchange credits the counterparty's margin account, and if the price has gone down, the exchange debits the margin account. Every time the amount exceeds the initial margin level, the participant can withdraw the excessive amount. In the opposite, if the daily settlements reduce the margin below the maintenance level, the clearinghouse makes a margin call, asking for the deposit of additional funds to cover the margin. Maintenance margin is the fixed minimum amount that must be maintained in the margin account. The described process is called mark-to-market. On the delivery date, the amount exchanged is the spot value and the counterparty can decide if he closes the account and withdraws all the remaining funds or not.

Edwards (1983) and Bates and Craine (1998) present the pyramid structure of margin collection in which clearing house stands on top and collects and returns margin funds to clearing members, according to the daily mark-to market result. The clearing members commission merchants (FCM) collect margins from non FCM members, who execute trades though FCM members. At the base of the pyramid, all FCM members collect margins from their customers.

In the decision of margin requirement, the tradeoff between counterparty risk and limitation of trading volume is the most interesting and important factor. If margins are not high enough to cover a potential counterparty's payment failure, the exchange may have to deal with great losses. On the other side, too high margins can affect liquidity because the cost of funding will increase and a lot of participants will exit market.

The process of margin setting has occupied several papers, starting with Figlewski (1984). He assumes that margins should be set such that the probability of a loss large enough to drain margin before it is renewed, is less than an acceptable level.

Gay, Hunter and Kolb (1986), find that the margin level should be set such that the probability of a margin becoming inadequate to cover losses should be equal across time and contracts.

Fenn and Kupiec (1993) used two models based on the theory of efficient contracts (it proposes that in order to deal with increased price volatility, the number of settlements and margin levels should be increased). They blamed clearing houses for not changing the margins properly and they attributed the lack of frequently altering margins in the alternative methods of clearing houses to deal with risk: they strictly monitored their memberships, they asked for addition margin. The optimal level according to their research, assuming normality, was determined by factors like the conditional volatility of the contract, the settlement frequency and the ex-ante cost per unit deficit in the margin account.

Longin (1999) approaches margin setting through extreme value theory in order to find the margin level in a given probability value of margin violation. Comparing normal to extreme value distribution, he finds that in the observed period the results extracted from the second one, are more similar to the real margin levels and that normal distribution underestimates the optimal margin level.

Except from the appropriate margin levels, a lot of research has been devoted to how margins affect and are affected by the prices and other financial trends, especially after the stock market crash of 1987[i]. Hardouvelis (1988) argued that history evidence supported the proposition that stock volatility could be controlled by the use of margin requirements. Unlike him, Kupiec and Sharpe (1991) in their equilibrium model, find that there is no unique theoretical relationship between margin requirements and price volatility. They come in agreement with Fishe, Goldberg, Gosnell and Sinha (1990) whose evidence does not support the hypothesis that margin changes can systematically affect price volatility. Garleanu and Pedersen (2011) argue that if speculators have no funding constraints, prices cannot be affected by any margin change. In agreement with Acharya, Lochstoer and Ramadorai (2011), they note that changes in margin requirements affect asset prices and especially they do so because of speculators, who have capital-constraints and are more vulnerable to margin increases.

Hedeegard (2011) in his analysis in margin setting finds that exchanges response to volatility changes by changing margin levels. According to his paper, liquidity decreases as margins increase and open interest is harmed as holding a position becomes too expensive. As far as prices are concerned, he observes that prices are abnormally high right after the margin increase, while this does not happen after an announcement of margin decrease. The price impact is even bigger in long-term contracts which are hold by speculators with long positions. Hedeegard in his research computes the percentage maintenance margin requirements as 2.5 times the daily standard deviation of returns in his sample of 16 future contracts. The aforementioned measure is used in this dissertation as an alternative method of setting margins in the three future contracts we investigate.

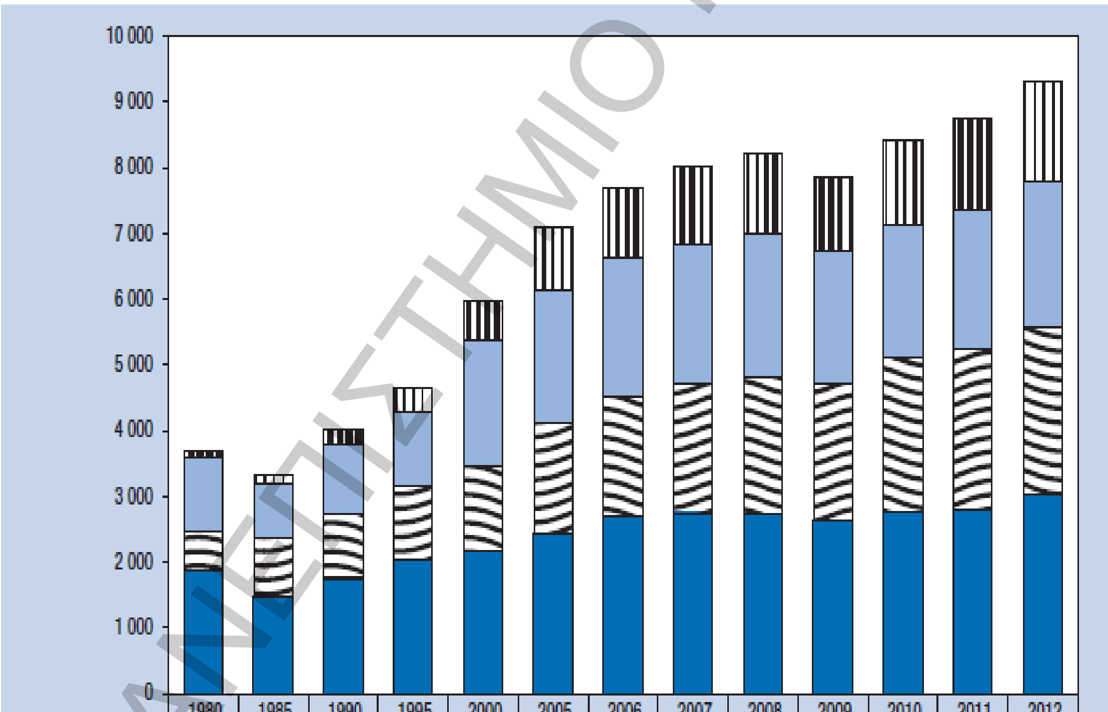
Daskalaki and Skiadopoulos (2012) in their commodity futures analysis find a positive relationship between changes in margins and prices (negative if prices are presented as returns). Positive relation is being shown between margins and volatility as well. Moreover, they also refer to speculators' behavior. When margins increase, they decrease their open positions, harming the trading volume and liquidity of the market but despite of that, margining can prevent excessive speculation. Lastly, the influence of margins is sensed in positive and large margins unlike small and negative ones.

The purpose of this dissertation is a distribution of margins setting in the freight derivatives market and especially IMAREX exchange through implementing various methods of Value at Risk model. The reader enters the world of shipping by reading chapter 2 in which the freight market is presented, along with its vital component, the freight rate, the emerging freight derivatives market and the significant "players" Imarex-Nos-Balting exchanges. In chapter 3 we explain the VaR models and Backtesting procedure, the most preferable of which we implement in chapter 4 and 5 (Dataset and Methodology), in order to decide and conclude in chapters 6 and 7 of the most appropriate in freight margins setting.

2. DESCRIPTION OF FREIGHT MARKET

2.1.General Presentation

Maritime transport is the backbone of international trade and a key engine driving globalization as described by United Nations Conference on Trade and Development in 2012 Report of Maritime Transport. During the last century, the shipping industry has experienced enormous development: sizes of ships 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 tremendously. Around 80% of global trade by volume and over 70% by value is carried by sea and is handled by ports worldwide.



	1980	1985	1990	1995	2000	2005	2006	2007	2008	2009	2010	2011	2012
Container	102	152	234	371	598	969	1 076	1 193	1 249	1 127	1 275	1 385	1 498
Other dry cargo	1 123	819	1 031	1 125	1 928	2 009	2 112	2 141	2 173	2 004	2 027	2 090	2 219
Five major bulks	608	900	988	1 105	1 295	1 709	1 814	1 953	2 065	2 085	2 335	2 477	2 547
Oil and gas	1 871	1 459	1 755	2 050	2 163	2 422	2 698	2 747	2 742	2 642	2 772	2 796	3 033

Figure 1-International seaborne trade per main cargo. (Millions of tons loaded)
The figure shows the trade per main cargo. From grain to crude oil, iron ore to chemicals, seaborne trade amounted to more than 8.7 billion tonnes in 2011 from 4 billion in 1990 and 7.1 billion in 2005.

In order to better understand the shipping market, the distribution of different cargo-ships is important. The cargoes are divided in five major categories:

1. General Cargo Vessels
2. Tankers
3. Dry-bulk Carriers
4. Multipurpose Vessels (different types of cargo like liquid and general)
5. Reefer Ships(carrying temperature controlled goods)

General Cargo Vessels carry packaged items like chemicals, foods, furniture, machinery, motor vehicles, footwear, garments, etc in general, nonspecialized stowage areas or standard shipping containers; e.g., boxes, barrels, bales, crates, packages, bundles, and pallets.

Liquid Cargo Carriers, or tankers, are specifically designed to transport liquid cargoes in bulk. Tankers can range in size of capacity from several hundred tons, which includes vessels for servicing small harbors and coastal settlements, to several hundred thousand tons, for long-range haulage. The tanker market, which encompasses the transportation of crude oil and petroleum products, represents approximately one third of the world seaborne trade volume. In general, smaller tankers carry "clean" cargoes (refined products, such as gasoline, diesel fuel, or jet fuel). Large tankers generally carry "dirty" (black oil or crude oil) cargoes. Tankers of less than 100,000 dwt are referred to as either "clean" or "dirty". Clean tankers carry refined petroleum products such as gasoline, kerosene or jet fuels, or chemicals. The so-called dirty vessels transport products such as heavy fuel oils or crude oil. Larger tankers usually only carry crude oil. Major types of tanker-ships include the oil tanker, the chemical tanker, and Gas carrier.

Dry bulk trades comprise iron ore, coal, grain, timber, steel and other similar cargoes which are shipped in bulk as opposed to carried in containers or other unit loads. Three are the main categories of dry bulk vessels: Handies (10 - 49,999 dwt), Panamax (50 - 79,999 dwt), Post-Panamax (80,000 – 109,999 dwt), Capesize (110,000-199,000 dwt) and Very Large Ore Carriers (VLOC) which carry over 200,000 dwt. The dry bulk shipping market is by far the largest sector of the world’s shipping market in terms of cargo volume and weight.

Specialized types of cargo vessels include container ships and bulk carriers (technically tankers of all sizes are cargo ships, although they are routinely thought of as a separate category).

2.2.The freight Rate

The most important subject in maritime transport is the **freight rate**, which is the price at which a certain cargo is delivered from one point to another. The activity, in which freight rate is finalized, is called chartering. A charterer, who is maybe a cargo owner as well, employs an intermediary, a shipbroker to find a ship and transfer the cargo from one place to another at a fixed freight. A charterer may also be a party without a cargo who takes a vessel on charter for a specified period from the owner and then trades the ship to carry cargoes at a profit above the hire rate. There are four main types of chartering: the voyage chartering, where the vessel is being hired for a voyage, time chartering, hiring the vessel for a specific period of time, trip time chartering, for a route only, and bare boat chartering for which no crew or provisions are part of the agreement.

The freight rates combine a very unpredictable area and vary overtime due to a number of factors. The intended destination is an important factor when it comes to calculating ocean freight rates. In simple terms, the longer the journey is, the exorbitant the ocean shipping rates are and vice-versa. Any extra charge levied by port authorities like the security service charges also tends to affect the ocean freight rate. For certain goods, the season becomes a very important factor. Grains and fruits transported during a particular freight season will have higher cargo rates. Ocean freight rate depends on the fluctuating rate of exchanges as well, and therefore is likely to be levied on the latest prevailing exchange rate. Fines in delays and terminal fees while embarking the journey from a port and after reaching the intended destination, and changes in fuel prices affect ocean freight rate as well. Moreover, if the shipper does not have enough goods to fill the containers to their optimum capacity, it will affect the freight charges by way of the shipper having to pay more in spite of lesser quantity.

From a microeconomic prospect, the freight rate is determined by supply and demand forces for freight services. Supply in shipping is the available capacity for carrying cargoes and demand derives from the need for goods to be transported. Because of the fact that the cost of sea transport is low as a proportion of the total cost of the final good and the sea transport cannot be easily substituted, demand is considered to be inelastic. During recessions, when many vessels are laid up, the supply is elastic. But when the market is strong and all ships are in service, the supply becomes inelastic. So a j-supply-curve is shaped. Referring to the supply, a distinction between short and long run must be made. In the short run, the fleet cannot expand, because it takes some years for ships to be built. So, in the short run, the total supply is almost fixed. But in the long run, where new ships can enter the market, the supply curve can shift to the right which will force freight rates to be lessened (if demand remains unchanged). The five most important factors generally affecting demand in the shipping market are:

1. The global economy: generates the bulk of demand, either through imports of raw materials or trade in finished products. Political factors are also very important and cover the strategies adopted by a government (Branch 1998).

2. Raw materials available: some raw materials, especially agricultural products, are subject to seasonal fluctuations. In addition, raw materials' trading is affected by overall demand and availability.

3. Average haul: expresses the distance travelled by a ship before it reaches its final destination. For example, periodic closures of the Suez Canal have increased the average sea travel distance between the Arabian Gulf and Europe from 6,000 miles to 11,000 miles.

4. Random shocks: shocks that affect the stability of the economic system such as natural disasters, wars, economic crises and the like.

5. Transport costs: raw materials will only be transported from remote places of origin if transport costs have been reduced to an acceptable level.

The five principal factors on the supply side are:

1. The world fleet: The size and composition of the world fleet clearly reflects the current supply of vessels.

2. Fleet productivity: even though the size of the fleet is fixed, ship productivity provides an element of flexibility.

3. Shipyard production: shipbuilding plays an active role in the adjustment of the fleet. However, it usually takes between one and three years from the ship is ordered until delivery takes place.

4. Scrapping: scrapping also plays a key role in fleet growth.

5. Freight earnings: Freight rates motivate shipowners to adjust their capacity in the short term and to identify savings and improvement in the longer term.

Cargo owners, charterers, shipbrokers, ship owners all transact in the market and their requirements try to balance supply and demand which in the end combined can explain the equilibrium freight level at which the demanded quantity equals the supplied quantity. A lot of studies describe this interaction such as this of Hawdon (1978), Strandenes (1984), Beenstock and Vergottis (1989), Kavussanos and Alizadeh (2002) and Stopford (2009)

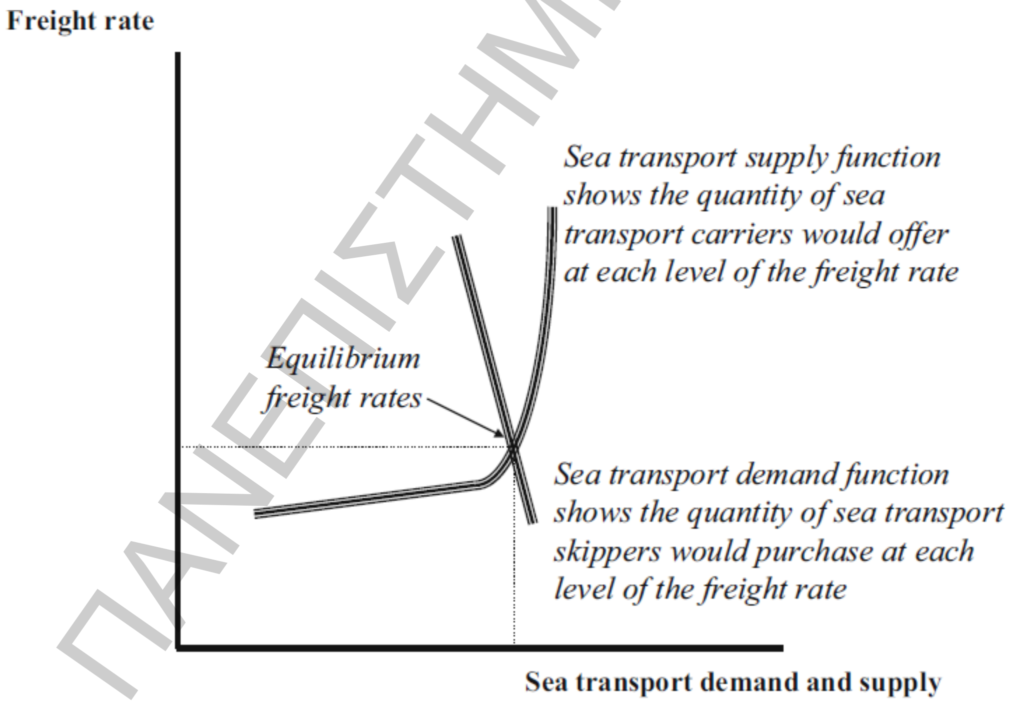


Figure 2- Schematic presentations of equilibrium of freight.
Source: Shipping and Transport Logistics (Lun, Y.H. Venus, Cheng, T.C. Edwin (Eds.)

2.3. The freight derivatives market

The 2008 collapse of the shipping market affected all market participants in terms of cash flows, valuation of collaterals, credit ratings and credit profiles, but most of all it affected their entrepreneurial spirit and confidence in a market once promising windfall earnings. When a physical market goes through such a deep recession, it automatically affects negatively all related traded derivative markets which can post exaggerated losses due to the use of leverage. In the recent case however of the Shipping Market Collapse and despite the fact that currently volumes are very slow (especially in the OTC options market) we will argue that the crisis actually paved the way for deeper market penetration of freight derivatives (expected at the next booming cycle) as they offer a number of key advantages which were overlooked so far by ship-owners but are currently considered a prerequisite as business plans are being revised and positions restructured. Below, we present the price graphs of both the dry bulk market main indices (Capesize and Panamax) and the tanker market (Clean and Dirty Tanker) (Figure 3).

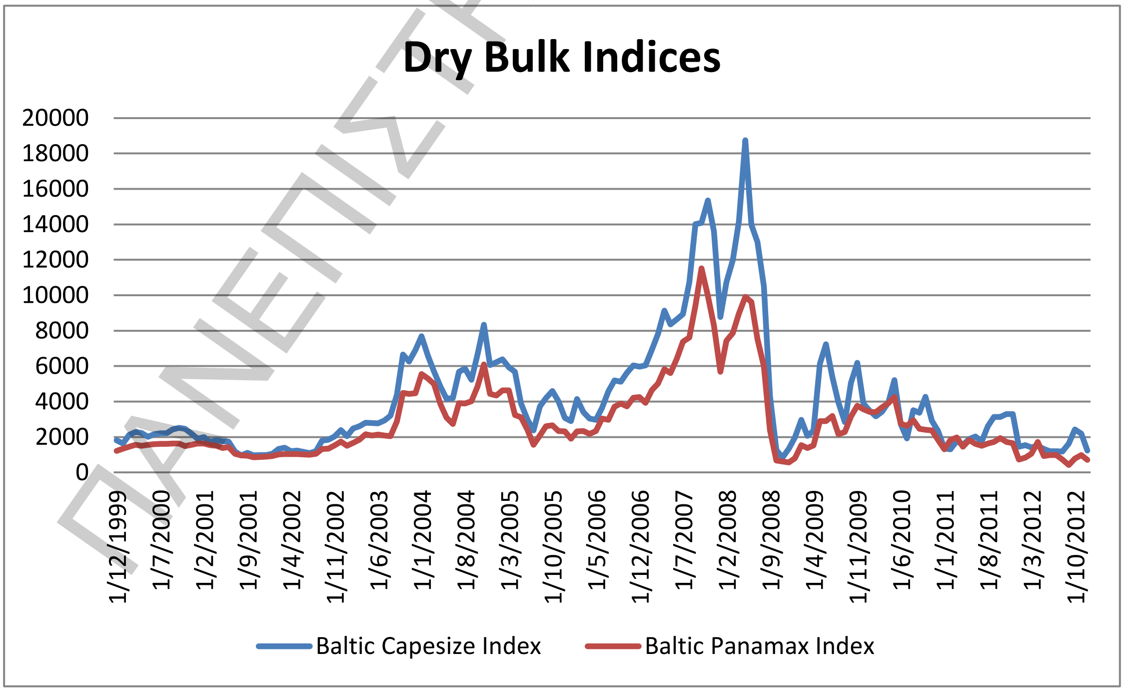


Figure 3.a – Baltic Capesize and Baltic Panamax Indices

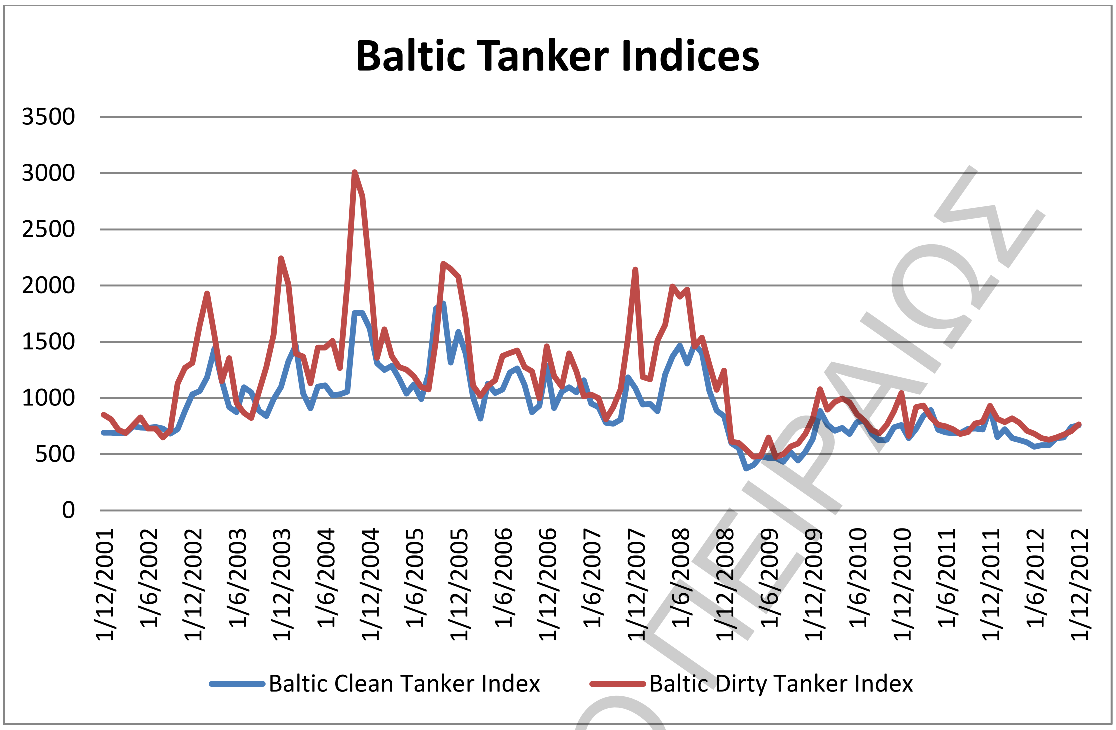


Figure 3.b – Baltic Clean and Dirty Tanker Indices. As we can see, the collapse in the freight market started to appear in the beginning of 2008 and deteriorate during mid-2008, where it collapsed below 1.000 points first time in October 2008. On the other hand, the tanker market, which peaked during 2004 due to the high global economic activity, followed the collapse one month later, at November 2008.

The collapse of the shipping market brought into the surface the need for proper and sufficient risk management from the shipping companies, which can do so by entering into forward contracts on the freight of specific routes, the so called Forward Freight Agreements or FFAs. The main uses of FFAs, apart from risk management and hedging, are also speculation and arbitrage. Their main advantage to ship-owners and charterers is the management of freight risk. The choice of a ship-owner to use freight derivatives through a financial institution in order to fix revenues - rather than use a physical charterer - automatically diversifies overall market and counterparty risk, as spot physical revenues are pegged to the shipping cycle from the vessel operation, while future expected revenues (through freight derivatives) are pegged to the banking cycle. Especially the recent increased use of clearing-houses for freight derivatives transactions, not only diversifies but actually minimizes counterparty risk, as ship-owners peg their potential hedging revenues against a pool of market counterparties rather than a single corporate entity. The 2008 94.4% drop of the BDI and BDTI resulted so far to a number of defaults in the physical shipping market (Armada Pte Singapore, Atlas Shipping, Industrial Carriers, etc), one default in the Banking Industry (Lehman Brothers) and no defaults of Clearing Houses which were stress-tested from the fierce freight indices drop and the Lehman default.

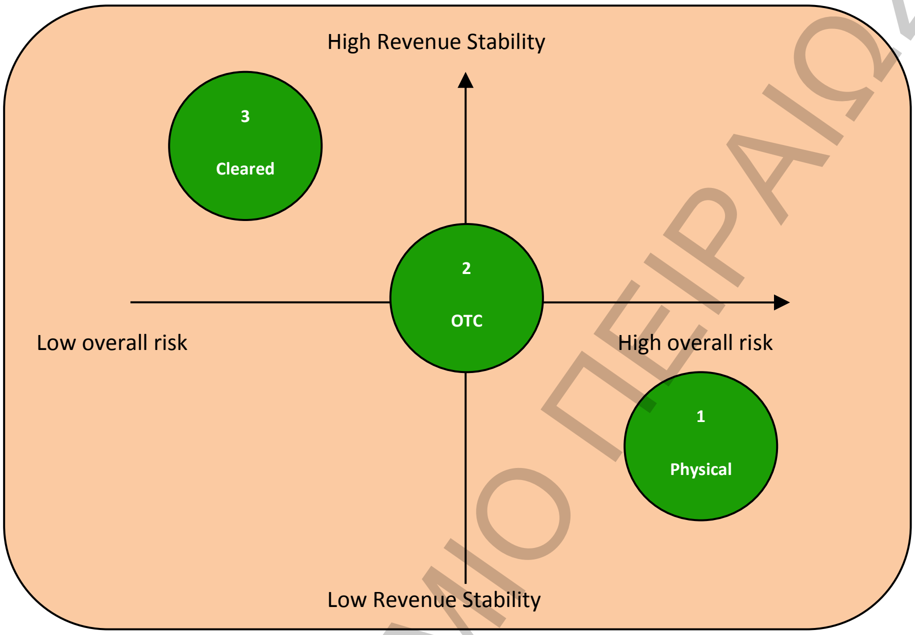


Figure 4 – Risk/Reward Matrix. Securing future revenues through Freight Derivatives (preferably cleared), improves the credit profile of the company as it is practically repositioned to a safer part of the Risk/Return Matrix

FFAs also contribute both to transparency and increased bargaining power of the market having maintained the anonymity of the players and providing easy access to counterparties. Having the ability to negotiate rates both in the spot (physical) market and in the freight derivatives (paper) market, ship-owners enjoy the benefit to: (a) cross-check freight rates from different market makers and acquire bargaining power against the physical charterers (b) trade in the freight derivatives market under conditions of anonymity and therefore do not reveal sensitive strategic positioning to competition. (c) execute derivative transactions immediately and therefore capture arbitrage opportunities which was not possible in the physical markets due to more bureaucratic procedures (d) take early redemption of derivative positions if desired (i.e.:. in the case of loan restructuring or profit taking)

Most notably, FFAs can secure budgeting by fixing costs and lock out future revenue stability. Expected revenues to maturity of the financial agreement can be secured, in contrast to the physical agreements where "*re-negotiation*" of rates at each time of crisis is a common market practice (charterers can exercise pressure from their dominant position in the physical market to reach a more favorable settlement). In contrast to the physical market, freight financial agreements can be considered "carved in stone" as agreed rates are not negotiable after inception. The use of freight derivatives can provide powerful tools in the hands of sophisticated ship-owners to optimize their risk/return ratio at any market conditions since specific instrument characteristics when implemented in the right strategy can improve ship-owners risk-profile and revenues as we will attempt to illustrate in our examples below using real market data.

The fierce collapse of the BDI and BDTI indices definitely left a deep mark in the shipping industry, however moving away from the "eye of the storm" we can identify the market need for a new paradigm where transactions are more flexible and transparent and where ship-owners can have more diverse income streams and better grip of market dynamics. An expansion of mark-to-market clearing in types of derivatives like FFA would enforce this dynamic.

Against traditional practices through purely physical transactions where the Charterers were ruling the game, freight derivatives offer the benefit of the choice to the Owners to take advantage of market conditions themselves and improve considerably their risk/return ratio while at the same time be more prepared to risk manage their positions against a future economic slowdown. The extraordinary potential of freight derivatives was revealed amid this recent collapse of the Shipping Market and this is why despite current drop of volumes we expect FFAs to prove the real winners in the medium term. Exchanges, clearinghouses like NOS, and margining systems will stand by them in the difficult recovery period.

2.4. Imarex, Nos, and the Baltic Exchange

As explained above, no freight future hedging tool can be adopted without the presence of an exchange and a clearing house. The two respective institutions involved in this dissertation are International Maritime Exchange, Imarex, and Nos Clearing House. History of Nos begins in 1987 as a central counterpart clearing house, a CCP, for the Norwegian equity derivatives market. Norwegian listed derivatives are traded at the Oslo Stock Exchange, and NOS was the clearing house for this market since its origin in 1990. NOS cleared both listed and OTC equity derivatives, bond futures, and stock borrowing and lending contracts. NOS sold the financial derivatives clearing business to VPS (the Norwegian CSD) in September 2006.

In November 2001, NOS meets Imarex, as it started offering clearing services to the international maritime freight derivatives market in cooperation with the International Maritime Exchange. This initiative has developed into the current global market for dry and tanker FFA trading. NOS has also been offering clearing services to the international bunker fuel oil derivatives market since December 2005. In September 2006 NOS and International Maritime Exchange merged ownership, creating the Imarex Group. In 2007, Imarex Group changed name to Imarex ASA. In 2010, Imarex ASA and its maritime freight derivatives restructured and integrated with Spectron. In 2011, Imarex sold its subsidiary Spectron Group to futures broker Marex. The end of clearing activities took place in 2012. IMAREX ASA announced that the agreement between the Company as seller and NASDAQ OMX Stockholm AB as buyer regarding NOS Clearing ASA has been closed and become effective on July 2, 2012

The Baltic Exchange is an independent exchange that provides trading, settlement and informative services to maritime finance participants. It has a long history originated back in the 18^{th} century which was first has been established under the name of "Virginia and Baltic Coffee House". It is the main calculation agent of the freight market indices which are broadly used in the shipping market. Namely:

- Baltic Dry Index (BDI)
- Baltic Panamax Index (BPI)

- Baltic Capesize Index (BCI)
- Baltic Supramax Index (BSI)
- Baltic Handysize Index (BHSI)
- Baltic Dirty Tanker Index (BDTI)
- Baltic Clean Tanker Index (BCTI)

Especially for the last two tanker indices, which are the primary focus of this paper, these are formulated combining the average freight of the following routes:

Dirty Tanker Routes	Clean Tanker Routes
1) TD1 Arabian Gulf to USGC (VLCC)	1) TC1 ME Gulf to Japan (75,000 mt)
2) TD2 Arabian Gulf to Singapore (VLCC)	2) TC2 Continent to USAC (37,000 mt)
3) TD3 Arabian Gulf to Japan (VLCC)	3) TC3 Caribbean to USAC (38,000 mt)
4) TD4 West Africa to USGC (VLCC)	4) TC4 Singapore to Japan (30,000 mt)
5) TD5 West Africa to USAC (Suezmax)	5) TC5 Middle East/Japan (55,000 mt)
6) TD6 Black Sea/Mediterranean (Suezmax)	6) TC6 Cross Mediterranean (30,000 mt)
7) TD7 North Sea to Continent (Aframax)	7) TC7 Singapore to EC Australia (30,000 mt)
8) TD8 Kuwait to Singapore (Aframax)	8) TC8 Arabian Gulf to UK/Continent (65,000 mt)
9) TD9 Caribbean to USGC (Aframax)	9) TC9 Baltic to UK/Continent (22,000 mt)
10) TD10D Caribbean to USAC (50,000 mt)	10) TC10 South Korea to NOPAC WC (40,000mt)
11) TD12 ARA to USGC (Panamax)	11) TC11 South Korea to Singapore (40,000 mt)
12) TD14 SE Asia to EC Australia (80,000 mt)	12) TC12 India/Far East-Japan WC (35,000mt)
13) TD15 West Africa (VLCC)	13) TC14 USGC to UK/Continent (38,000mt)
14) TD16 Black Sea to Mediterranean (30,000 mt)	
15) TD17 Baltic to UK/Continent (100,000 mt)	
16) TD18 Baltic to UK/Continent (30,000 mt)	
17) TD19 Cross Mediterranean	

Table 1- Dirty and Clean Tanker Routes composing BCTI and BDTI indices.

A general discussion of the freight market reveals the need of trying to understand and measure its high risks in order to find ways of protections like optimal margining systems. VaR models are precious tools in this try.

3. THE VAR MODEL

3.1. General Presentation

Value at Risk (VaR onwards) is a statistical method that provides the analyst with a metric of the maximum loss that can occur in a financial asset, or portfolio of assets, at any given time. It was first introduced amid the US stock market crash of 1987, as a risk management tool which could provide assistance in portfolio management. By this time, the quantitative analyst society were trying to explain the occurrence of very rare events in the financial market, which seemed to appear more often than they "should" given the standard volatility estimation assumptions. The normal distribution assumption of stock returns seemed more and more unsuitable to explain extreme events[ii], as the persistence of fat tails pinpointed the difference. VaR was developed as a systematic tool to deal with such extreme events, in the trading desks of Banker's Trust in the late 1980's and further developed by J.P. Morgan. In 1994, J.P Morgan gave public access to its so called Risk Metrics Group and that was the first time were VaR left the trading desks and entered the broad society. In 1997, the Securities and Exchanges Commission (SEC onwards), which is the regulator of financial institutions in the US, enforced a rule of mandatory disclosure of the VaR exposure of the financial institutions derivatives portfolios. From 1999, where the Basel Commission, the European regulator of financial institutions, gave further impetus to the use of VaR, making it now the most common risk metric in the financial world. In mathematical terms, VaR is the maximum loss that an asset or portfolio of assets can incur in a given time period, with X% confidence accuracy, as the following equation shows:

$$\Pr ob\left[\Delta P\left(\Delta t,\Delta X\right) > -VaR\right] = 1-a \quad (3.1)$$

where, ΔP(Δt, ΔX) is the change in the value of the portfolio in question, expressed as a function of the time interval, Δt and the change of the random state variables, ΔX and 1-α is the confidence level. Confidence interval defines the percentage of the time that an investor should not lose more than the VaR amount. The most widely used assumption of the VaR model is that of normality. VaR assumes that the distribution of asset returns follows a normal distribution. As a result, the part of the tail that corresponds to 5% of the observations is -1.65 standard deviations from the mean. This level is used in most of the VaR models and is the main criticized point of the VaR, as in reality the distribution of asset returns has fat tails. "Fat tails" refers to the fact that large market moves occur more frequently than what would occur if market returns were normally distributed

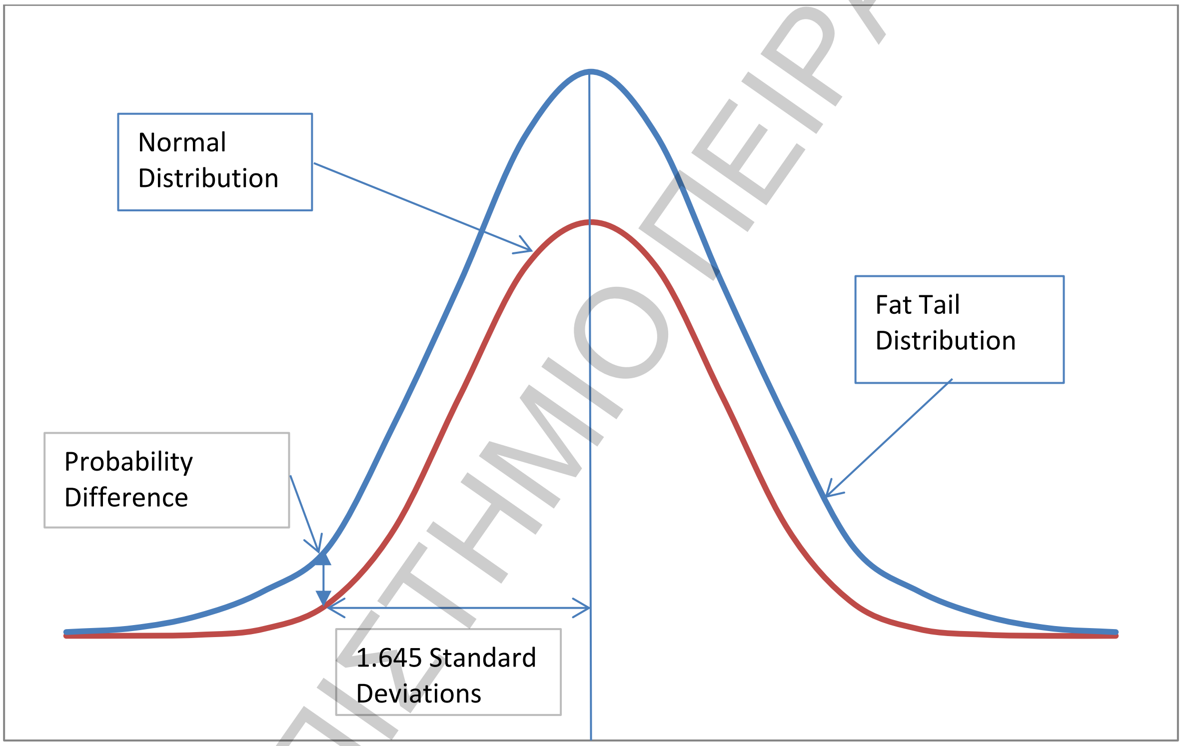


Figure 5-Fat tailed vs. normal distribution. We present the normal distribution (red curve) in comparison to a fat tailed distribution (blue curve). It can been seen that the calculation of VaR under the assumption of normality, yields different results in comparison to the fat tailed distribution creating inference problems i.e. the 1.65 standard deviations correspond to more than 5% of the observations under the fat tailed distribution

It can be measured using either the variance – covariance method, the historical simulation method or the Monte Carlo approach and will be discussed in the following section.

3.2. Variance – Covariance method

Variance – and the closely standard deviation- is a measure of how spread out a distribution is. They can be considered as measures of volatility. The variance is computed as the average squared deviation of each number from its mean:

$$\sigma^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2 \tag{3.2}$$

Under the Var-Covar method, which is the most straightforward approach, one needs to identify just three data series; the weight of each asset in the portfolio, the standard deviation of the asset returns and the correlation between the assets within the portfolio.

$$VaR = \Phi^{-1}\ a\ \cdot \sqrt{\sum_{i=1}^{N} v_i^2 \cdot \sigma_i^2 + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} V_i \cdot V_j \cdot \sigma_i \cdot \sigma_j \cdot \rho_{i,j}} \tag{3.3}$$

where, $\Phi^{-1}(a)$ is the inverse function of the cumulative distribution function of standard normal distribution (for example $\Phi^{-1}(0.95) = 1.65$) , V_i is the value of the asset position i, σ_i is the volatility of the returns of the asset position i, V_j is the value of the asset position j, σ_j is the volatility of the returns of the asset position j and $\rho_{i,j}$ is the correlation of returns of assets i and j.

The first term of the above equation multiplies the squared asset position value with the volatility of its returns, calculating the P&L of the position assuming zero correlation. The second term of equation (3.3) adjusts for the diversification effect with the inclusion of the correlation coefficient. If the correlation coefficient is positive, then the total VaR increases, but when the correlation coefficient is negative then it reduces the total VaR, assuming long positions in either case.
Correlation coefficient describes the direction and degree of relationship between two variables. Formula for ρ_{12} would then be:

$$\rho = \frac{\sum_{i=1}^{n}\left[(X_i^{\,1} - \bar{X}^{\,1})\cdot(X_i^{\,2} - \bar{X}^{\,2})\right]}{\sqrt{\sum_{i=1}^{n}(X_i^{\,1} - \bar{X}^{\,1})^2 \cdot \sum_{i=1}^{n}(X_i^{\,2} - \bar{X}^{\,2})^2}} \tag{3.4}$$

Due to the fact that the inputs needed, in order to calculate the VaR of a portfolio of assets becomes extremely large as the assets increases[iii], we can transform equation (3.3) into a linear form, in order to simplify our calculations. The below equation uses vectors of data instead of single ones and is exactly the same as equation (3.3):

$$VaR = \sqrt{\vec{V}R\vec{V}^{-1}} \tag{3.5}$$

where, V is the simple risk vector, R is the correlation matrix and $V^{-1}$is the transpose of the simple risk vector. The simple risk vector is just the multiplication of the position vector with the corresponding volatility vector, as shown below:

$$\begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_n \end{pmatrix} \times \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \vdots \\ \sigma_n \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \\ V_n \end{pmatrix} \tag{3.6}$$

Thus, the simplified matrix form that calculates the VaR of a portfolio of assets is of the following form:

$$\left[\begin{pmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \\ V_n \end{pmatrix} \times \begin{pmatrix} 1 & \rho_{1,2} & \rho_{1,3} & \cdots & \rho_{1,n} \\ \rho_{2,1} & 1 & \rho_{2,3} & \cdots & \rho_{2,n} \\ \rho_{3,1} & \rho_{3,2} & 1 & \cdots & \rho_{3,n} \\ \vdots & \vdots & \vdots & 1 & \vdots \\ \rho_{n,1} & \rho_{n,2} & \rho_{n,3} & \cdots & 1 \end{pmatrix} \times \begin{matrix} V_1 & V_2 & V_3 & \cdots & V_n \end{matrix}\right]^{\frac{1}{2}} \tag{3.7}$$

The benefits of the VaR – Covar approach are clear, as we deal with a simple form approach in which we only need to make a distribution assumption. It is easy to be understood, flexible and widely accepted. On the other hand, it has some worth mentioning flaws. The first has to do with the distribution assumption, as the asset returns do not follow an exact normal distribution but a fat tailed one in their majority. For example, if there are quite a few outliers in our sample, then the statistical distance of 1.65 standard deviations would not correspond to 5% (or 2.33 standard deviations on the 1%) of the sample but to greater percentage, making the actual VaR larger than the estimated one. Even if the normal distribution assumption holds, there might be miscalculations in the variances and / or covariances of the individual assets. This is true given that the estimation of these values is based on historic data and is susceptible to estimation errors. Also, there is one more problem that lies in the context of stationarity. Stationary variables are the ones that do not change from time to time due to external shocks. If the variables in question are not stationary, which is a fact most of the times, the calculation of VaR is not accurate. Finally, VaR needs convexity adjustments to account for the skewness[iv] in the distribution caused by the convexity of bonds and options. Although this adjustment is feasible, it still remains an approximation and does not account for the risks of non-vanilla derivatives.

3.3.Exponentially Weighted Moving Average

An important factor of popularity of VaR is undoubtedly the EWMA estimator revealed in 1996 from JP Morgan's RiskMetric Technical Document. The exponentially weighted moving average (EWMA) is a statistic for controlling that averages the data in a way that gives less and less weight to data as they are further removed in time and was firstly introduced by JP Morgan. The weights on past squared returns decline exponentially as we move backward in time, suggesting that most recent returns are more responsible for today's volatility. JP Morgan RiskMetrics' EWMA assumes conditional normality and as Nelson and Foster (1994) highlight in their research, EWMA is an optimal method assuming that returns are conditionally normally distributed. The forecasted volatility is modeled as an IGARCH (1, 1) and results from:

$$\sigma_{t+1}^2 = \lambda \cdot \sigma_t^2 + (1-\lambda) \cdot R_t^2 \tag{3.8}$$

where:

$$\sigma_t^2 = \frac{1}{\lambda} \cdot (1 - \lambda) \cdot \sum_{\tau=2}^{\infty} \lambda^{\tau-1} \cdot R_{t+1-\tau}^2 \tag{3.9}$$

The VaR, respectively to the other Var-Covar models, is then calculated by:

$$VaR = \sigma_{t+1} \cdot \Phi^{-1}(\alpha) \tag{3.10}$$

where a is the known confidence level of VaR.

The EWMA introduces lamda, which is called the smoothing parameter. JP Morgan sets λ=0,94 for daily variance forecasting and is used in this dissertation.

3.4.Historical Simulation

The Historical Simulation method is based on historic profit and loss data of the portfolio in question. This approach, although it has the advantage over the Var - Covar approach as it is based on portfolio specific past data, it cannot take into account unprecedented events and their impact on portfolio value. Historical simulation is just a non-parametric approach that calculates each day the profit and loss change based on actual observable prices. It does not make a normality assumption but rather it relies on the data embedded volatility and correlation measures. Next we present an example of the method. Assume that a portfolio consisting of X assets has presented a historical sensitivity in their corresponding return change of €10.000 per basis point. We move forward by estimating the profit and loss for each day, using the basis point change and our estimate of the sensitivity in the change of the portfolio value. Then, we rank our findings in a descending order and we use the 950th ranked observation value for the 95% VaR or the 990th ranked observation value for the 99% VaR.

The historical simulation method is the simplest method of the four methods presented. Its advantage over the VaR – Covar method is that it doesn't need the normality, constant correlations / covariances assumptions. The potential losses are calculated using historical returns of the risk factors in which the non-normality of their distribution is embedded, making the capture of extreme events ("Black Swans") more feasible.

3.5.Monte Carlo Simulation

The Monte Carlo simulation approach is a stochastic process where the use of a large number of iterations on several scenarios is being conducted (usually with the help of a computer with a spreadsheet processor) in order to generate a mid - outcome. This makes it not only a more flexible approach, when compared to the previous ones, but also a more realistic one which can yield more accurate findings. What the Monte Carlo approach does is that it generates a large number of randomly generated simulations which are aggregated to form the final outcome. Repeating the process for a significant amount of times can give a good indication of the actual output of the VaR model. In order to apply the Monte Carlo simulation we follow an algorithm similar to that of the Historical Simulation but instead of using historical observation it uses random generated data. Implementing the Monte Carlo simulation we need the following five steps and a strong processor:

a. The first step includes the determination of the length, T, of the data set and divide it with equal weights into a number of N increments, where $\Delta t = T/N$.
b. Under the second step we draw a random sample, using a random sample generator, and calculate the value at the end of the first increment.
c. Continue to apply step b until the end of the data set time horizon, using all the data increments.
d. Repeat steps b and c under a large number of M iterations and generate M possible outcomes for the selected time horizon.
e. Rank the M different outcomes in ascending order and identify the value at the selected confidence interval.

The advantage of the Monte Carlo simulation compared to the other methods described above, is that by simulating random paths can capture and model better non – linear payoff functions, especially in that of complex derivatives. In addition, Monte Carlo simulations can capture better tail (very rare) events and can provide more details of the occurrence of these tail events beyond VaR. On the other hand, Monte Carlo simulations require a great deal of computer power as the number of iterations goes larger and larger. For example, for a portfolio of 100 assets that we want to do 1.000 iterations we would need 100.000 different simulations, each of which increases the probability of model risk. As a result, an engine that is used to perform Monte Carlo simulations is usually very expensive. The sampling error is another drawback of Monte Carlo method, since, by analyzing samples and not the whole population in a procedure, statistics in the sample differ from statistics in the population.

3.6.The ARCH / GARCH models

The General Auto Regressive Conditional Heteroskedasticity (GARCH) model (Bollerslev, (1986)) is an econometrics model that is used to describe time series data for which it is believed that at any point in time the terms will have a characteristic variance. It is actually a mechanism that includes past variances in the explanation of future variances. Conditional stands for the independence on the observations of the immediate past. Autoregressive is the feedback mechanism that incorporates past observations into the present.

Garch is a generalized form of the ARCH model (Engle, 1982), which is a model that is used broadly in the finance industry to explain time varying volatility clustering. In time series, price changes tend to cluster together, presenting amplitudes of price changes. Mandelprot (1963), explaining the phenomenon, noted that: "*large changes tend to be followed by large changes and small changes tend to be followed by small changes*".

The basic assumption of Ordinary Least Squares (OLS)[v] method is that the expected value of the squared residuals is the same at any given point in time. This is the homoscedasticity assumption which is the focus of the ARCH and GARCH models. The problem of heteroskedasticity exists when the variance of the error terms is varying which is the case when the error terms are expected to differ in different period of time (i.e. larger in some periods comparing to some others). Although the estimation of the coefficients is unbiased, in such cases, the standard errors suffer from heteroskedasticity which makes the inference using confidence intervals look falsely precise. Also, the convenient OLS method applies equal weights in the observations when estimating volatility, which is not true as more recent events influence volatility more than older ones. The ARCH and GARCH models deal with these flaws, as their primary descriptive tool is to estimate the weights that will be applied in the data when forecasting volatility. ARCH and GARCH models treat heteroskedasticity as a variable to be modeled. In the next section of the dissertation, we will introduce the ARCH model which is the foundation of the GARCH model presented afterwards.

The ARCH (q) process is defined as the following equation shows:

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \cdot \varepsilon_{t-i}^2 \quad , \quad \alpha_0 > 0, \alpha_i \geq 0 \tag{3.11}$$

Where, σ_t is the variance at period t, a_0 is the constant and ε_t is the error term of period t. These error terms follow a process which includes a white noise process and a time – dependent standard deviation:

$$\varepsilon_t = \sigma_t \cdot z_t \tag{3.12}$$

Where, z_t is a white noise process. The estimation of ARCH(q), where q is the number of lagged values, is implemented using OLS as follows. A variable is white noise if there is:

– Constant mean

– No serial correlation (zero autocovariances, expect at lag zero)

– Homoskedasticity (constant variance)

– If variance =1 then it is iid (independent identically distributed) zero mean and finite variance random variable.

– If y_t is distributed as a standard normal, then the autocorrelation coefficients are normally distributed.

The first step is to estimate the best fitting Auto Regressive (AR) model using the following regression:

$$y_t = a_0 + a_1 \cdot y_{t-1} + \ldots + a_q \cdot y_{t-q} + \varepsilon_t = a_0 + \sum_{i=1}^{q} a_i \cdot y_{t-i} + \varepsilon_t \tag{3.13}$$

The estimates of the error term, ε_t, of the previous regression are squared and regressed against a constant and q lagged values, as shown below:

$$\hat{\varepsilon}_t^2 = \hat{a}_0 + \sum_{i=1}^{q} \hat{a}_i \cdot \hat{\varepsilon}_{t-1}^2 \tag{3.14}$$

The ARCH effect is widely examined in the literature and is applied to numerous fields of economics and finance. On the economics field, the persistence of ARCH effects and the asset volatility have led researchers to dig up the origin of this, both on the micro and the macro level. On the micro level, Lamoureux and Lastapes (1990) propose that the ARCH effects seem to cluster in the volumes of trading. When they inputted the trading volumes data into their variance equation of their GARCH(1,1) model, they found that lagged squared error terms are no longer significant.

The GARCH model of Bollerslev (1986), like the ARCH, is a weighted average of the past squared residuals but also it has declining weights that asymptotically tends zero. The product of these models is parsimonious and easy to estimate, yielding successful results. Especially, the GARCH model proposes that the Best Linear Unbiased Estimator (BLUE) of the future variance is a weighted average of the past variance, the current period predicted variance and the new information, in the current period, which is captured by the most recent squared residuals. On the macro level, Jagannathan (1991) found that nominal interest rates present statistical significance on the explanation of volatility. In addition, Jagannathan (1991) showed that a GARCH model with the nominal interest rate in it leads to a decrease in volatility persistence. As far as the interest rates are concerned, Shiller (1979) and Singleton (1980) have argued in favor of the high volatility of interest rates, as compared to the metrics of the rational expectations hypothesis. This means that the estimates of future interest rates, as described by the rational expectations hypothesis are biased.

The mathematical formulation of a GARCH (p,q) model is described by the below formula:

$$\sigma_t^2 = a_0 + a_1 \cdot \varepsilon_{t-1}^2 + \ldots + a_q \varepsilon_{t-q}^2 + \beta_1 \cdot \sigma_{t-1}^2 + \ldots + \beta_p \cdot \sigma_{t-p}^2 = a_0 + \sum_{i=1}^{q} a_i \cdot \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \cdot \sigma_{t-i}^2$$

(3.15)

The above equation, which is an ARCH (q) process plus the term $\sum_{i=1}^{p} \beta_i \cdot \sigma_{t-i}^2$, is generated when an ARMA model is assumed for the variance of the residuals. In such case the model is a GARCH (p, q) model with p the order of the GARCH terms (σ^2)

and q the order of the ARCH terms (ε^2). In practice, when dealing with heteroskedasticity in empirical models, the most common method is the White's test. In the case of time series data, however, the best test is the test for ARCH and GARCH errors.

The estimation of the order of the GARCH terms, p, is quick and simple, as it needs just the three following steps. First, we need to estimate the best fitting AR(q) model following equation (3.13). Then, we need to estimate and plot the autocorrelation terms of the squared residuals by the following equation:

$$\rho = \frac{\sum_{t=i+1}^{\mathrm{T}} \left[\hat{\varepsilon}_t^2 - \hat{\sigma}_t^2 \;_i \cdot \; \hat{\varepsilon}_{t-1}^2 - \hat{\sigma}_{t-1}^2 \right]}{\sum_{i=1}^{T} \hat{\varepsilon}_t^2 - \hat{\sigma}_t^2 \;^2} \tag{3.16}$$

The asymptotic volatility of the autocorrelation function is $\frac{1}{\mathrm{T}}$, meaning that any values larger than this are an indication of GARCH errors. The estimation of the total number of lags can be performed using the Ljung – Box test and generally it is the number of lags that correspond to values less than $\frac{1}{\mathrm{T}}$, given the significance level. As we know, the Ljung – Box test follows the X^2 distribution with n degrees of freedom, given that the squared residuals are not correlated. Rejecting the null hypothesis of no GARCH errors indicates such errors in the conditional volatility.

In VaR method, in order to forecast one-day ahead variances, a GARCH (1,1) model is most preferable. Applied for a GARCH (1,1) model the equation (5.15) is transformed in:

$$\sigma_t^2 = a_0 + a_1 \cdot \varepsilon_{t-1}^2 + \beta_1 \cdot \sigma_{t-1}^2 \tag{3.17}$$

3.7.Backtesting

Backtesting of data is a method that can be applied to any data set in order to test the ability of a model to capture changes in its explanatory variables. It is a method most used in the science industry, mainly finance and capital markets but also in fields like weather and climate change forecasting and medicine tests. What the method of backtesting does, is that it tests how a model could have performed, if it has been used in a past data period. A very important backtesting function, used broadly in the banking and finance sector by the risk management department, is the backtesting of VaR models. The majority of these back tests compare the observed portfolio losses, for a given period, to the estimated VaR values. The common procedure of doing this is to calculate the number of times the observed portfolio losses fall outside the VaR estimations, given the confidence level.

A variety of backtesting procedures have been introduced since the 1990's as an effort to test the accuracy of VaR models. Although there are many different tests in the literature, the majorities of them focus on a specific transformation of the estimated VaR and realized profit and loss. If we define as H_t the event function which takes the values of 1, if the observed VaR value, $x_{t,t+1}$, exceeds the estimated one, VaR_t, and zero otherwise, we can form the following function set:

$$H_{t+1} = \begin{cases} 0, & if \;\; x_{t,t+1} \leq VaR_t \\ 1, & if \;\; x_{t,t+1} \geq VaR_t \end{cases} \tag{3.18}$$

The above set produces the event function H_{t+1} (e.g. {0,0,0,1,0,1 etc.}) which points the times where our VaR estimate was less than the actual outcome. Christoffersen (1998), proposes the closed form solution of the determination of the accuracy of VaR, which is the determination of whether the event function satisfies the following two properties:

1. The Unconditional Coverage Property states that the probability of a loss hit more than the estimated one, must be a fraction of 100% i.e. $P(H_{t+1}=1)=a$. If these losses hit more frequently than a x 100%, then the estimated VaR value understates systematically the risk level of the portfolio. On the other hand, if unexpected losses hit less frequently than a x 100%, then the VaR estimate is too conservative.

2. The Independence Property sets restrictions on the number of the occurrences where the observed VaR value is greater than the estimated one. In order for it to be valid, the elements of the event sequence must be independent from each other. Generally speaking, the Independence Property condition requires that the past data of VaR violations must not carry any information about the probability of future VaR violations. If past violations influence the occurrence of future violations, there is strong evidence of the inadequacy of the estimated VaR model.

The three measures of Christoffersen (2003), which are implemented in this dissertation, are explained and presented below:

- Firstly, we want to check if the fraction of violations of VaR is significantly different from promised fraction p, Unconditional Coverage Testing, named $LR_{uc,}$ provides us with that information and is produced by the equation:

$$LR_{uc} = -2 \cdot \ln \left[\frac{1-p^{\mathrm{T}_0} \cdot p^{\mathrm{T}_1}}{\left(1-\frac{\mathrm{T}_1}{\mathrm{T}}\right)^{\mathrm{T}_0} \cdot \left(\frac{\mathrm{T}_1}{\mathrm{T}}\right)^{\mathrm{T}_1}} \right] \sim \chi_1^2 \tag{3.19}$$

, where p represents the VaR coverage rate or the promised fraction, T_1 number of days that our risk model was not violated and T_0 that was violated.

- Second measure, Independence Testing, helps us to reject a VaR with clustered violations, so that the elements of our series will be independent from each other. In independence testing, the formula employed is:

$$LR_{ind} = -2 \cdot \ln \frac{\left(1-\frac{\mathrm{T}_1}{\mathrm{T}}\right)^{\mathrm{T}_0} \left(\frac{\mathrm{T}_1}{\mathrm{T}}\right)^{\mathrm{T}_1}}{(1-\pi_{01})^{\mathrm{T}_{00}} \pi_{01}^{\mathrm{T}_{01}} (1-\pi_{11})^{\mathrm{T}_{10}} \pi_{11}^{\mathrm{T}_{11}}} \sim \chi_1^2 \tag{3.20}$$

Where:

$$\pi_{01} = \frac{\mathrm{T}_{00}}{\mathrm{T}_{00} + \mathrm{T}_{01}} \tag{3.21}$$

$$\pi_{11} = \frac{\mathrm{T}_{11}}{\mathrm{T}_{10} + \mathrm{T}_{11}} \tag{3.22}$$

$T_{i,j} = 0,1$ is the number of observations with a j following an i.

- For conditional coverage, where we try to find simultaneously the if the VaR violations are independent and the average number of violations are correct, we use jointly the LR_{uc} and LR_{ind} tests. Therefore:

$$LR_{cc} = LR_{uc} + LR_{ind} \tag{3.23}$$

LR unconditional coverage and independent ratios are asymptotically chi-squared distributed with one degree of freedom, while LR conditional coverage follows a chi-squared distribution[vi] with two degrees of freedom.

One famous backtesting procedure is the Proportion of Failures (PoF) test that calculates the number of times of VaR violations in a given time span. Kupiec (1995) proposed this test statistic, which is captured by the following equation set for T observations:

$$PoF = 2 \cdot \log\left[\left(\frac{1-\hat{a}}{1-a}\right)^{T-H\ \ a} \cdot \left(\frac{\hat{a}}{a}\right)^{H\ \ a}\right] \tag{3.24}$$

$$\hat{a} = \frac{1}{T} H\ \ a \tag{3.25}$$

$$H\ \ a\ = \sum_{i=1}^{T} H_t\ \ a \tag{3.26}$$

The above test statistic estimates whether the proportion of VaR violations is statistically different from a x 100%, as explained before, testing the accuracy of the VaR estimation.

Another influential backtesting procedure is the Markov test presented by Christoffersen (1998). The Markov test examines the dependence of the probability of a VaR violation, at any given day, with the probability of a VaR violation, the previous day. The effectiveness and accuracy of the estimated VaR value depends on the independence of these likelihoods.

	$H_{t-1} = 0$	$H_{t-1} = 1$	
$H_t = 0$	N_1	N_2	$N_1 + N_2$
$H_t = 1$	N_3	N_4	$N_3 + N_4$
	$N_1 + N_3$	$N_2 + N_4$	N

Table 2 – Markov Test Matrix. The Markov test is conducted by creating a 2x2 contingency table that includes the number of VaR violations on adjacent days. If the VaR estimate is accurate, then the proportion of violations that occur after a previous one should equal the proportion of violations that occur on after an observation of no violation.

An updated version of the above Markov test is that by Christoffersen and Pelletier (2004), which states that the time period between VaR violations should be independent of the time period that has elapsed since the last violation, if VaR violations are completely independent from each other. Generally speaking, the likelihood of a VaR violation in the future should not depend on whether the last VaR violation occurred. The test, unlike the Markov test, cannot be constructed with the 2x2 matrix presented before; rather it requires the estimation of a model for the duration of the period between VaR violations. The estimation of the above model is conducted with the help of the maximum likelihood method and evidence shows that it has more power in determining VaR measures than the Markov test.

Although the mentioned tests, along with the broad independence tests list, provide a useful tool in the evaluation of the estimated VaR accuracy, they are prone to a major flaw which is the assertion that any valid VaR value will happen in a series of independent observations. In practice, there are countless ways in which the independence property may be violated. For example, the probability of a VaR violation at time t may not be depended on the probability of a VaR violation at time t-1 but instead at time t-10 or t-20 etc. To conclude, the verification of the accuracy of VaR is a task of high priority to every self - respecting risk management desk. In order for the verification to be proper, a risk manager needs a backtesting method. As described before, there are many different methods of backtesting VaR accuracy, each one with its pros and cons. Each risk manager should carefully implement backtesting methods in his VaR models in order to continuously evaluate risk.

4. DATASET

We obtain daily prices on three FFAs from Bloomberg, written on one TC2 and two TD3 routes and traded through Imarex during the period of 10/02/2005 to 20/12/2012. TD3 administrates a route in which very large tanker-ships of 250.000 metric tones, VLCC, carry crude oil from Eastern Gulf to Japan. TC2 route represents the transfer of defined petroleum contracts such as gasoline and diesel fuel accomplished by medium-range tankers, of 37.000 metric tones approximately, that sail from Europe Continent to USAC.

Because of the fact that each FFA has a certain expiration date, Imarex has to roll to the next shortest to maturity contract, few days before the expiration, in order to be able to provide Bloomberg with continuous prices.

We transform the three price series into log-return series with continuous compounding as follows:

$$y_t = \ln\left(\frac{P_{t+1}}{P_t}\right) \tag{4.1}$$

The price- to- turn transformation generally guarantees a stable data for modeling. This sense of stability is termed into econometrics as stationarity. Stationarity is the quality of a process in which the statistical parameters (mean, covariance) do not change over time. Mathematically:

$$\Delta y_t = a_i y_{t-1} + \sum_{i=1}^{P} \gamma_i \Delta y_{t-i} + \varepsilon_t \tag{4.2}$$

$$\Delta y_t = a_0 + a_i y_{t-1} + \sum_{i=1}^{P} \gamma_i \Delta y_{t-i} + \varepsilon_t \tag{4.3}$$

$$\Delta y_t = a_0 + a_i y_{t-1} + a_2 y_t + \sum_{i=1}^{P} \gamma_i \Delta y_{t-i} + \varepsilon_t \tag{4.4}$$

,where t stands the trend, ε_t stands for the white noise error and P for the lags of differences in order to assume that the estimated errors are not serially correlated.

The first form has constant and trend, the second has constant only and the third has no constant or trend.

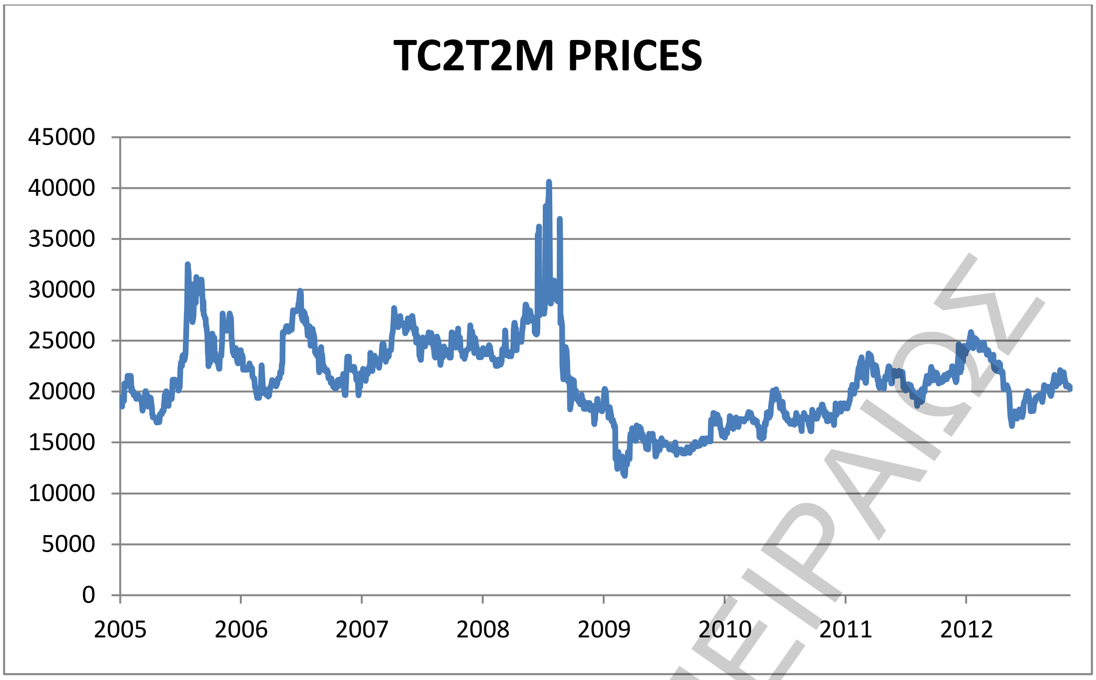
We can check for stationarity by checking for a unit root. The Augmented Dickey-Fuller test examines for unit root. We apply the ADF test without constant or trend to check for unit root and the null hypothesis is that there is a unit root.

As extracted from the results, our series are non-stationary in levels but stationary in first differences. This evidence approves the need of transforming our series in log-return series.

Descriptive Statistics for each FFA

Panel A: FFA (levels)			
	TD31M	**TC2**	**TD32M**
Mean	14.455,08	21.212,25	14.213,76
Median	13.395,00	21.090,00	13.210,00
Maximum	43.820,00	40.620,00	40.590,00
Minimum	7.630,00	11.710,00	7.750,00
Std. Dev.	4.767,88	4.111,82	4.171,27
Skewness	2,2473	0,4968	2,3014
Kurtosis	9,6826	3,8315	10,0516
Jarque-Bera	**5.372,51***	**139,05***	**5.873,71***
Probability	**0,0000***	**0,0000***	**0,0000***
Observations	1.988	1.988	1.988
ADF	-1,6501*	-0,4629*	-1,0688*
Panel B: FFA (differences)			
	TD31M	**TC2**	**TD32M**
Mean	-0,0001	0,0000	-0,0001
Median	0,0000	0,0000	0,0000
Maximum	0,3087	0,3066	0,3084
Minimum	-0,3935	-0,3242	-0,3645
Std. Dev.	0,0488	0,0273	0,0405
Skewness	-0,4754	-0,1500	-0,1920
Kurtosis	12,8589	45,1336	17,1203
Jarque-Bera	**8.122,03***	**146.982,80***	**16.519,41***
Probability	**0,0000***	**0,0000***	**0,0000***
Observations	1.987	1.987	1.987
ADF	-39,6322	-43,0590	-42,9819

Table 3a-Entries report the summary statistics for each one of the three FFAs observed on levels and differences. The Jarque-Bera test and Augmented Dickey-Fuller(ADF) are also demonstrated. One and two statistics denote rejection of the null hypothesis at the 1% and 5% respectively. Null hypothesis for Jarque-Bera is normality and for ADF is unit root (non-stationarity). The rejection of normality hypothesis is presented in the three histograms of returns below. The sample period is 10/02/2005-20/12/2012.



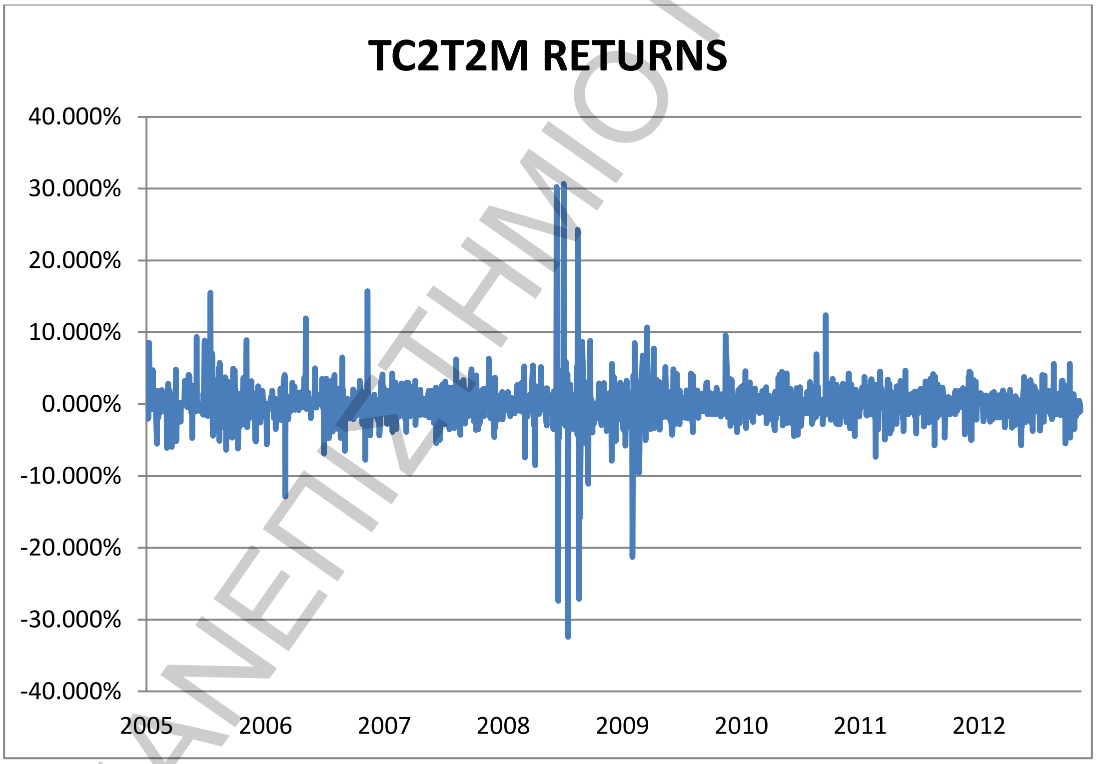


Figure 7- Schematic comparison of prices and returns of an FFA. The two graphs show vividly the stability of mean described above. When we get log-returns, the observations pose a 'quick return' on the zero mean.

Furthermore, we apply Engle ARCH test in order to search for heteroskedasticity (or volatility clustering as described above), all garch effects well captured by Garch models. The phenomenon of volatility clustering is observed in figure 9 as well.

	TD31M	TC2	TD32M
F-statistic	23.69*	19.63*	6.25*

Table 3b–Heteroskedasticity tests on data returns. Table reports Engle ARCH test for heteroskedasticity. The null hypothesis is homoskedasticity. One asterisk denotes rejection of null hypothesis at 1% significance level.

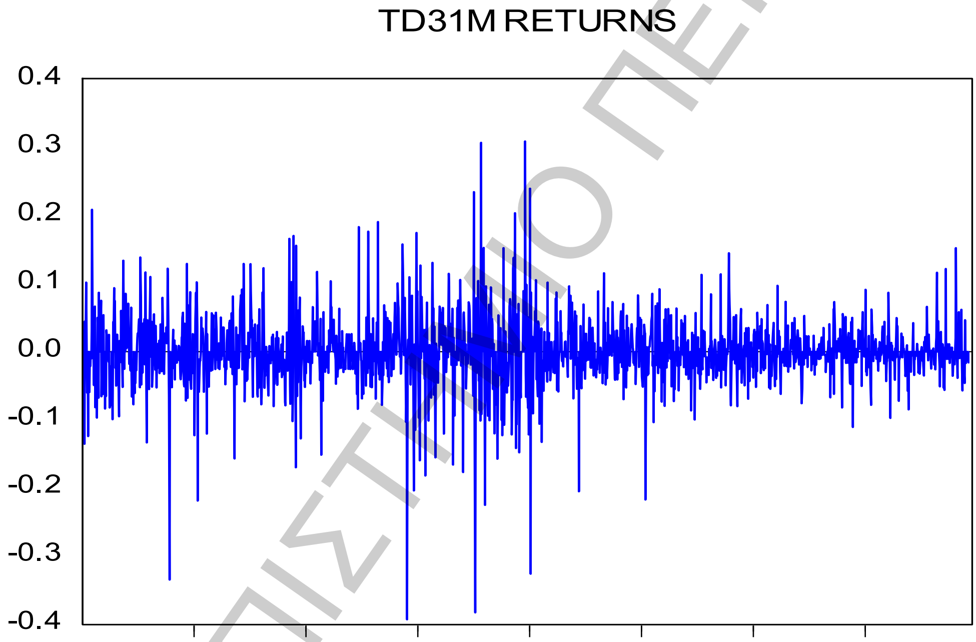


Figure 9-TD3 Returns. Volatility clustering is observed when the variance of returns is high for extended periods and then low for extended periods. The returns of our portfolio are exhibiting this exact point of view. It seems to have "virtual windows" of similar volatility extents.

The series are observed as a single derivative per route and as a portfolio composed of three equally weighted derivatives and the 95% and 99% VaR forecasts are generated for each one of them and for the portfolio as an entity. We apply the historical simulation, the Var-covar method, an Exponentially Weighted Moving Averages method and a GARCH (1, 1) model for 95% and 99% confidence interval.

5. METHODOLOGY

5.1.Historical Simulation in practice:

The single HS method is established in the following steps:

- Firstly we sub-divide the sample period in rolling windows of 100 and 250 observations. First rolling window consists of the first 100 observations. Second rolling window contains the second to the 101^{st} observation. We continue creating all the rolling windows according to the FIFO sequence, by letting the first observation out and adding the following of the last observation in the window.
- Secondly, returns are sorted in descending order in each rolling window.
- We select the loss that is equaled to the 5 percent of the time - when we apply the 95% confidence interval - and to the 1 percent of the time (for HS with 99% confidence interval. This is our Value at Risk using the 5% and 1% probability. Since we have a 100-rolling window this is the fifth and first worst observation respectively. For the 250-rolling window the observations are the average of the 12^{th} and 13^{th} return for the 95% c.i. and the average of second and third return for the 99% c.i. (using interpolation).

5.2.Variance-Covariance in practice:

When working on each FFA separately, the VaR is calculated as $1.65 \cdot \sqrt{\sigma^2}$, when implementing the 95% confidence interval, and $2.33 \cdot \sqrt{\sigma^2}$ when considering 99% confidence interval for the returns of each rolling window of the under-survey-period.

The most demanding task in using the Var-Covar method in a portfolio is the computation of the portfolio variance. Applying the aforementioned equation (3.3) for portfolio variances, the formula for three assets turns to:

$$VaR = \Phi^{-1} \ \alpha \ \sqrt{\sigma^2_{portfolio}} = \Phi^{-1} \ \alpha \ \sqrt{\chi_1^2\sigma_1^2 + \chi_2^2\sigma_2^2 + \chi_3^2\sigma_3^2 + 2\chi_1\chi_2\rho_{12}\sigma_1\sigma_2 + 2\chi_1\chi_3\rho_{13}\sigma_1\sigma_3 + 2\chi_2\chi_3\rho_{23}\sigma_2\sigma_3} \tag{5.1}$$

5.3.EWMA in practice:

Taking into consideration the rolling window-creation-process described above, we compute a VaR for every rolling window of 100 only observations. There was no need of applying the method in the 250-rolling window because after the 100th lag, there cumulated weight is almost 100%. The formula itself proves it as:

$$1-\lambda \ \cdot \sum_{\tau=1}^{100} \lambda^{\tau-1} = 0.998 \tag{5.2}$$

so that 99.8% of the weight has been included. Therefore, only 100 daily returns are needed to be stored in each rolling window in order to calculate the tomorrow's variance σ^2_{t+1}.

5.4.GARCH(1,1) in practice:

The extreme kurtosis reported in table 3a supports the use of a GARCH model. We use a GARCH(1,1) model, the most preferable in bibliography of VaR, in order to get 1-day ahead predictions about volatility again. GARCH, being a time series technique that uses past variances to forecast future variances, is used in this dissertation in order to forecasted volatility be extracted. GARCH successfully captures types of heteroskedasticity such as volatility clustering and thick tailed returns.

In processing GARCH model we used larger rolling windows of 500 observations each, covering a period of two years, in order to produce more reliable estimates, assuming normal and t-student distribution. T-student distribution is used in order to relax the assumption that the conditional returns are normally distributed. T-student distribution is bell-shaped like normal but has heavier tails meaning that it is more likely to produce values far from the mean.

5.5.Backtesting in practice

Backtesting took place using the three measures of Christoffersen named LR_{uc}, LR_{ind}, and LR_{cc} for 1%, 5% and 10% significance level. The null hypothesis is that the VaR violations equal the coverage rate of VaR. Because of the fact that LR_{uc} and LR_{ind} are chi-squared distributed with one degree of freedom, the results are compared with the critical values of chi-squared distribution for one degree of freedom. These critical values are 2.71, 3.84, and 6.63 for 1%, 5% and 10% significance level respectively. On the other hand, the critical values of chi-distribution with two degrees of freedom are 4.61, 5.99 and 9.21 for the same respective significance levels. LR_{cc} results are compared with the latest as conditional coverage ratios follow chi-squared distribution with two degrees of freedom.

5.6. Alternative methodology

In practice, margins are computed using either VaR methology or other similar methods in which they multiply volatility with a factor depending on the riskiness of the asset. For example, NSE (National Stock Exchange of India) and BSE (Bombay Stock Exchange) divide companies in groups. In the safer portfolios, the two exchanges use the higher of: 3.5 times volatility and 7.5% of value for margining them.

Nos, the clearing house which used to cooperate with Imarex, uses SPAN (Standard portfolios Analysis of Risk) methology in the field of margins. SPAN is a margin system which provides assessments of portfolio risk by scanning through sixteen different price and volatility scenarios. Once the risk is calculated, SPAN credits or provides a margin reduction

Hedeegard (2011) in his research for margins finds that average percentage margin requirement is about 2.5 times volatility. We apply the finding in our three FFAs in contrast to the computed VaR margins calling it h-margin. We do so by computing the variance in each rolling window and multiply it with 2.5. The results are presented in the next section.

6. RESULTS AND DISCUSSION

Having a well-functioning margin system is a pre-requisite for any derivatives exchange. In this dissertation we have provided five VaR methods, and one standard way of computing margins, in order to provide the best method in setting margins in freight market. FFAs are OTC derivatives that are not traded in a mark-to-market daily clearing basis. We propose a use of margins in these type of derivatives, as losses harm participants equally and clearing protects exchanges of extreme return volatility periods, during which member may not be able to fulfill their commitments.
This dissertation tries to compose a valuation of freight risk through VaR with margining in shipping derivatives market, continuing similar research of freight market and margining systems.

Angelidis and Skiadopoulos (2007) implement various parametric and non-parametric VaR methods in various dry and wet cargoes and find that non-parametric methods perform better in shipping derivatives market.

According to the existing literature (Brooks et al (2005), Longin (1999), Cotter (2001) and a lot of others), non-parametric methods are known for their inability to produce sufficient accurate VaR forecasts. Pioneered by Tippet with the help of Fisher (1928), and supported by many scientists since then, traditional methods are found to be inferior to extreme value theory for estimating VaR, allowing the latter to emerge in the spotlight of financial risk management. Extreme value theory tries to assess the probability of events that are more extreme than any observed prior, dealing with extreme deviations from the median of probability functions. Jones and Perignon (2008) using margin data from Chicago Mercantile Exchanges, quantify default risk using extreme value theory as well.

Knott and Polenghi (2006) assessing margin coverage in future contacts, conclude that for their observed coverage levels, historical and t-distributions perform better in assuming appropriate margins.

The occupying with VaR models has resulted in a plethora of market risk measurement models highlighting the question of which is the most appropriate model for quantifying financial risk. In general, there is no model outperforming all alternative universally and under all circumstances. Moreover, results seem to depend on the market, time horizon examined and parameters assumed, and the way of backtesting.

In freight market, Historical Simulation seems to be more preferable than others. Dimitrakopoulos and Kavussanos (2009), in their freight markets research, agree with Angelidis and Skiadopoulos (2007) in the proper application of non-parametric methods. Taking into consideration the aforementioned findings in freight market, we are not surprised by the well-performed Historical Simulation method in this dissertation. Existing literature on energy commodities offers important implications for freight market due to common characteristics in-between the two fields of interest. Cabedo and Moya (2003) comparing HS, ARCH models and VaR-Covar methods, find that HS outperforms all others. In Angelidis, Benos (2006) and Sarma, Bekiros and Georgoutsos (2005), Kuester et al (2006) the Historical Simulation seems to rival parametric specifications.

Furthermore, Dimitrakopoulos and Kavussanos (2009) find that best performing models for quantifying daily exposures in the liquid and bulk sector was, along with the aforementioned Historical Simulation, the GARCH with normal innovations. This is coming to an agreement with our findings, since the GARCH model adopted in this dissertation, had a few and sometimes none rejection through the backtesting.

In this research, Value at Risk is met by means of portfolio and each FFA separately (TD3 route with expiration of one and two months and TC2 route with 2-month expiration). As far as the portfolio is concerned, Historical Simulation, either with HS-250 or HS-100, is the only method passing all the tests in both 95% and 99% confidence intervals by means of portfolio. GARCH method passes all the tests in the backtesting of 99% confidence interval for portfolio and gives accurate profit and loss forecasts. Furthermore, the well performing of GARCH method is approved in the backtesting of the three separate FFAs assuming t-student distribution. The more fat-tailed-oriented t-student distribution is approved to capture greater volatility excesses.

The Historical Simulation, relieved from normality hypothesis, does not suffer from tail-bias. The IMAREX, though, should take into consideration the trade-off between defense against default risk and "injury" of liquidity. The margins proposed through Historical Simulation could protect the exchange accurately from excessive price volatility, though this can cost abandoned derivatives positions. It is a conservative method as it commits margins in high levels when volatility exceeds and remains in the same levels for a long time before it changes abruptly again.

The Var-Covar method seems the most inaccurate. Because of the fact that conditional returns are not normally distributed, it is shown in the Jarque Bera Test, the extracted by Var-Covar method results understate the true VaR. Margins, if computed by this method, would be inappropriate for capturing the risks of freight derivatives.

EWMA performs better in 95% confidence interval when observing single portfolios of freight derivatives without forcing margins to high conservative levels.

In general comparison of 99% and 95% for each FFA, 99% confidence interval gives better results.

The alternative method (h-margin) with 2.5 times volatility is not backtested like the other methods. As it is shown in the graphs below, sometimes seems to cover losses that VaR methods cannot (figure 11) and some other times to overestimate risk (figure 12).

The graphs below show vividly the extracted results and the tables present the level of VaR and backtesting. The average proposed margin levels from the best method applied, GARCH, fluctuate between 5% and 12% .

	HS RW-100	HS RW-250	VAR-COVAR RW-100	VAR-COVAR RW-250	EWMA RW-100	GARCH(1,1) RW-500_n	GARCH(1,1) RW-500_t
AVERAGE VaR	-4,59%	-4,52%	-4,89%	-5,05%	-4,73%	-4,99%	-5,93%
MINIMUM VaR	-10,35%	-8,77%	-12,82%	-9,74%	-18,11%	-41,36%	-49,13%
MAXIMUM VaR	-2,27%	-2,84%	-2,42%	-2,76%	-1,81%	-2,67%	-3,17%
% OF EXCEPTIONS	4,50%	4,55%	3,87%	3,80%	3,87%	3,50%	2,49%
LRuc	1,01	0,77	5,50 **	5,72 **	5,50 **	7,87 ***	24,04 ***
LRind	1,26	1,59	3,06 *	2,19	0,55	3,84 **	1,94
LRcc	2,27	2,36	8,56 **	7,92 **	6,05 **	11,71 ***	25,98 ***

Table 4**: 95% VaR for portfolio and backtesting** . The three stars mean that the method does not pass the test in any of 1%, 5%, and 10% significance level. The two stars mean that the test is not passed in 5% and 10% significance level and the single star means that the test is not passed only in 10% significance level. Backtesting is conducted through the three tests of Christoffersen (LR_{uc}, LR_{ind}, and LR_{cc}) for 1%, 5% and 10% significance level. LR unconditional coverage and independent ratios are compared with the critical values of 6.63, 3.84 and 2.71 in 1%, 5% and 10% significance level respectively. The null hypothesis for LRuc test is that the average number VaR violations are correct. The null hypothesis tested for LRind is that the VaR violations are independent. For LR conditional coverage ratio, the respective critical values are 9.21, 5.99 and 4.61.The the null hypothesis test is that the average number of violations is correct and the VaR violations are independent. The average, minimum and maximum VaR are reported.

	HS RW-100	HS RW-250	VAR-COVAR RW-100	VAR-COVAR RW-250	EWMA RW-100	GARCH(1,1) RW-500_n	GARCH(1,1) RW-500_t
AVERAGE VaR	-10,80%	-10,15%	-6,91%	-7,13%	-6,68%	-7,05%	-7,81%
MINIMUM VaR	-34,06%	-23,91%	-18,11%	-13,75%	-25,57%	-58,40%	-64,66%
MAXIMUM VaR	-3,20%	-3,98%	-3,42%	-3,90%	-2,55%	-3,77%	-4,17%
% OF EXCEPTIONS	0,85%	0,86%	1,85%	1,67%	2,01%	1,28%	0,87%
LRuc	0,46	0,34	11,12 ***	6,55 **	15,14 ***	1,07	0,25
LRind	0,29	0,28	0,21	1,02	1,60	0,52	0,25
LRcc	0,76	0,62	11,33 ***	7,57 **	16,74 ***	1,58	0,50

Table 5: **99% VaR for portfolio and backtesting** . The three stars mean that the method does not pass the test in any of 1%, 5%, and 10% significance level. The two stars mean that the test is not passed in 5% and 10% significance level and the single star means that the test is not passed only in 10% significance level. Backtesting is conducted through the three tests of Christoffersen (LR_{uc}, LR_{ind}, and LR_{cc}) for 1%, 5% and 10% significance level. LR unconditional coverage and independent ratios are compared with the critical values of 6.63, 3.84 and 2.71 in 1%, 5% and 10% significance level respectively. The null hypothesis for LRuc test is that the average number VaR violations are correct. The null hypothesis tested for LRind is that the VaR violations are independent. For LR conditional coverage ratio, the respective critical values are 9.21, 5.99 and 4.61.The the null hypothesis test is that the average number of violations is correct and the VaR violations are independent. The average, minimum and maximum VaR are reported.

	FFAs	HS RW-100	HS RW-250	VAR-COVAR RW-100	VAR-COVAR RW-250	EWMA	GARCH(1,1) RW-500_n	GARCH(1,1) RW-500_t
AVERAGE VaR	TC2T2M	-3,53%	-3,38%	-4,03%	-4,22%	-3,86%	-4,23%	-5,03%
	TD3T1M	-7,10%	-7,05%	-7,57%	-7,77%	-7,31%	-7,77%	-9,22%
	TD3T2M	-5,76%	-5,80%	-6,11%	-6,33%	-5,86%	-6,02%	-7,16%
MINIMUM VaR	TC2T2M	-11,08%	-5,63%	-12,80%	-9,14%	-18,33%	-17,55%	-20,85%
	TD3T1M	-16,09%	-13,25%	-15,37%	-13,31%	-21,31%	-33,12%	-39,34%
	TD3T2M	-14,05%	-11,74%	-14,96%	-11,99%	-20,03%	-25,61%	-30,42%
MAXIMUM VaR	TC2T2M	-1,73%	-2,55%	-2,19%	-2,67%	-1,60%	-2,28%	-2,71%
	TD3T1M	-3,01%	-4,52%	-3,46%	-4,44%	-2,32%	-3,00%	-3,56%
	TD3T2M	-2,23%	-3,19%	-2,68%	-3,05%	-1,93%	-2,79%	-3,31%
% OF EXCEPTIONS	TC2T2M	5,25%	4,84%	3,66%	3,57%	4,35%	3,30%	2,42%
	TD3T1M	4,98%	4,84%	3,87%	3,63%	4,13%	3,43%	2,49%
	TD3T2M	4,45%	4,84%	4,40%	3,74%	4,35%	3,77%	2,56%
LRuc	TC2T2M	0,24	0,10	7,88 ***	8,28 ***	1,78	10,29 ***	25,51 ***
	TD3T1M	0,00	0,10	5,50 **	7,59 ***	3,16 *	8,63 ***	24,04 ***
	TD3T2M	1,24	0,10	1,50	6,31 **	1,78	5,19 **	22,62 ***
LRind	TC2T2M	0,74	0,44	1,34	0,10	2,98 *	0,16	1,19
	TD3T1M	7,19 ***	5,17 **	4,95 **	2,72 *	0,27	3,70 *	1,94
	TD3T2M	2,62	7,16 ***	2,79 *	1,00	0,66	0,85	2,05
LRcc	TC2T2M	0,98	0,53	9,22 ***	8,37 **	4,75 *	10,45 ***	26,70 ***
	TD3T1M	7,19 **	5,27 *	10,44 ***	10,31 ***	3,43	12,33 ***	25,98 ***
	TD3T2M	3,86	7,26 **	4,29	7,31 **	2,43	6,04 **	24,66 ***

Table 6: **95% VaR for each FFA and backtesting** . The three stars mean that the method does not pass the test in any of 1%, 5%, and 10% significance level. The two stars mean that the test is not passed in 5% and 10% significance level and the single star means that the test is not passed only in 10% significance level. Backtesting is conducted through the three tests of Christoffersen (LR_{uc}, LR_{ind}, and LR_{cc}) for 1%, 5% and 10% significance level. LR unconditional coverage and independent ratios are compared with the critical values of 6.63, 3.84 and 2.71 in 1%, 5% and 10% significance level respectively. The null hypothesis for LRuc test is that the average number VaR violations are correct. The null hypothesis tested for LRind is that the VaR violations are independent. For LR conditional coverage ratio, the respective critical values are 9.21, 5.99 and 4.61.The the null hypothesis test is that the average number of violations is correct and the VaR violations are independent. The average, minimum and maximum VaR are reported.

	FFAs	HS RW-100	HS RW-250	VAR-COVAR RW-100	VAR-COVAR RW-250	EWMA	GARCH(1,1) RW-500_n	GARCH(1,1) RW-500_t
AVERAGE VaR	TC2T2M	-8,82%	-8,88%	-5,69%	-5,95%	-5,45%	-5,98%	-6,62%
	TD3T1M	-17,80%	-15,50%	-10,69%	-10,97%	-10,32%	-10,97%	-12,14%
	TD3T2M	-13,20%	-12,90%	-8,63%	-8,94%	-8,28%	-8,51%	-9,42%
MINIMUM VaR	TC2T2M	-32,42%	-27,25%	-18,07%	-12,91%	-25,89%	-24,78%	-27,44%
	TD3T1M	-39,35%	-30,44%	-21,70%	-18,79%	-30,09%	-46,76%	-51,78%
	TD3T2M	-36,45%	-28,30%	-21,12%	-16,93%	-28,28%	-36,17%	-40,05%
MAXIMUM VaR	TC2T2M	-2,82%	-3,79%	-3,09%	-3,77%	-2,26%	-3,22%	-3,56%
	TD3T1M	-5,68%	-7,28%	-4,88%	-6,26%	-3,27%	-4,24%	-4,69%
	TD3T2M	-3,92%	-5,14%	-3,78%	-4,30%	-2,73%	-3,93%	-4,36%
% OF EXCEPTIONS	TC2T2M	0,90%	0,92%	1,59%	1,61%	1,85%	1,28%	1,01%
	TD3T1M	0,74%	0,75%	1,70%	1,61%	1,85%	1,55%	1,34%
	TD3T2M	1,11%	0,81%	1,96%	1,90%	2,12%	1,55%	1,14%
LRuc	TC2T2M	0,19	0,11	5,62 **	5,54 **	11,12 ***	1,07	0,00
	TD3T1M	1,39	1,22	7,63 ***	5,54 **	11,12 ***	3,85 **	1,61
	TD3T2M	0,23	0,71	13,74 ***	11,24 ***	18,08 ***	3,85 **	0,29
LRind	TC2T2M	2,16	0,32	0,48	0,55	1,94	1,40	2,20
	TD3T1M	0,22	0,21	1,14	0,55	1,94	0,75	0,57
	TD3T2M	1,45	0,24	4,30 **	0,23	0,07	0,75	0,42
LRcc	TC2T2M	2,36	0,43	6,11 **	6,09 **	13,07 ***	2,47	2,20
	TD3T1M	1,62	1,43	8,77 **	6,09 **	13,07 ***	4,60	2,19
	TD3T2M	1,69	0,95	18,04 ***	11,47 ***	18,15 ***	4,60	0,71

Table 7: **99% VaR for each FFA and backtesting** . The three stars mean that the method does not pass the test in any of 1%, 5%, and 10% significance level. The two stars mean that the test is not passed in 5% and 10% significance level and the single star means that the test is not passed only in 10% significance level. Backtesting is conducted through the three tests of Christoffersen (LR_{uc}, LR_{ind}, and LR_{cc}) for 1%, 5% and 10% significance level. LR unconditional coverage and independent ratios are compared with the critical values of 6.63, 3.84 and 2.71 in 1%, 5% and 10% significance level respectively. The null hypothesis for LRuc test is that the average number VaR violations are correct. The null hypothesis tested for LRind is that the VaR violations are independent. For LR conditional coverage ratio, the respective critical values are 9.21, 5.99 and 4.61.The the null hypothesis test is that the average number of violations is correct and the VaR violations are independent. The average, minimum and maximum VaR are reported.

	MARGINS RW-100	MARGINS RW-250
AVERAGE VaR	-7,41%	-7,65%
MINIMUM VaR	-19,43%	-14,75%
MAXIMUM VaR	-3,67%	-4,18%

Table 8- Hedeegard portfolio margins. Alternative computing of portfolio margins through 2.5 times volatility .

	FFAs	MARGINS RW-100	MARGINS RW-250
AVERAGE VAR	TC2T2M	-6,11%	-6,39%
	TD3T1M	-11,47%	-11,77%
	TD3T2M	-9,26%	-9,59%
MINIMUM VAR	TC2T2M	-19,39%	-13,86%
	TD3T1M	-23,28%	-20,16%
	TD3T2M	-22,66%	-18,16%
MAXIMUM VAR	TC2T2M	-3,31%	-4,04%
	TD3T1M	-5,24%	-6,72%
	TD3T2M	-4,06%	-4,62%

Table 9- Hedeegard margins per FFA. Alternative computing of separate FFA margins through 2.5 times volatility .

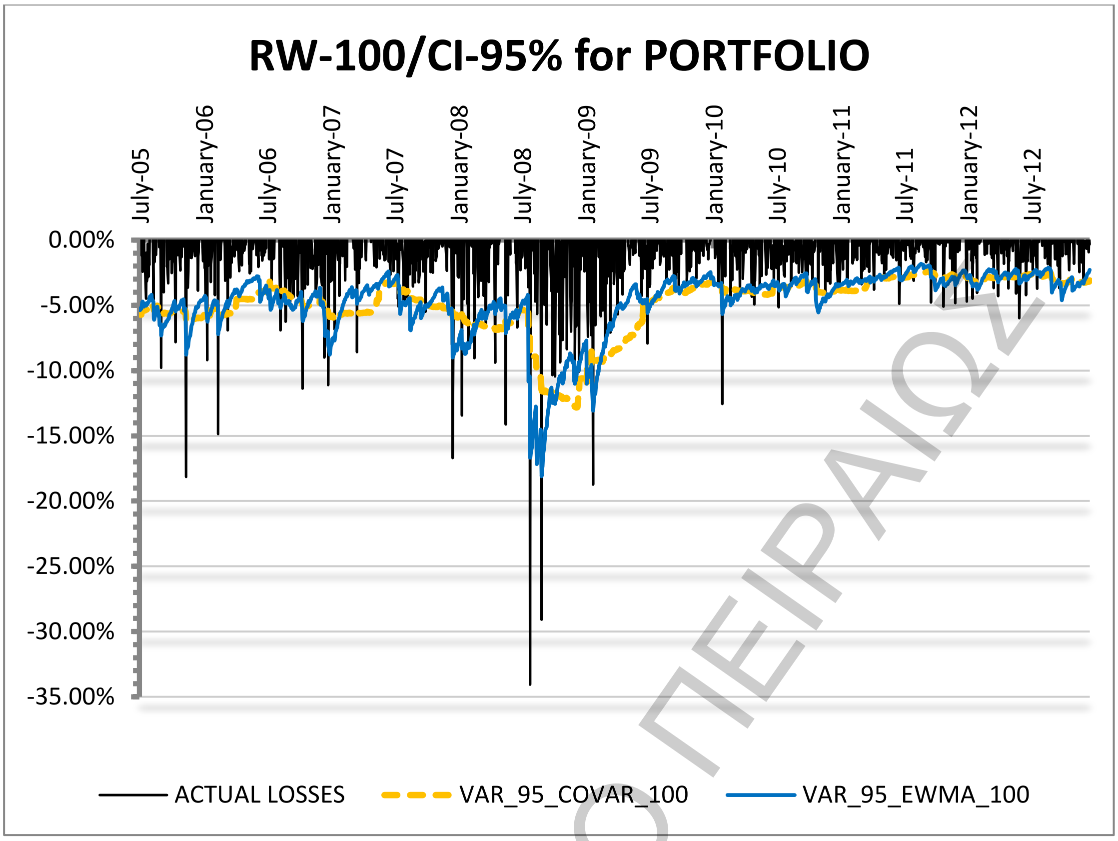


Figure 9-The 95% one day Var-Covar 100, the EWMA-100 and the actual losses over the backtesting period.

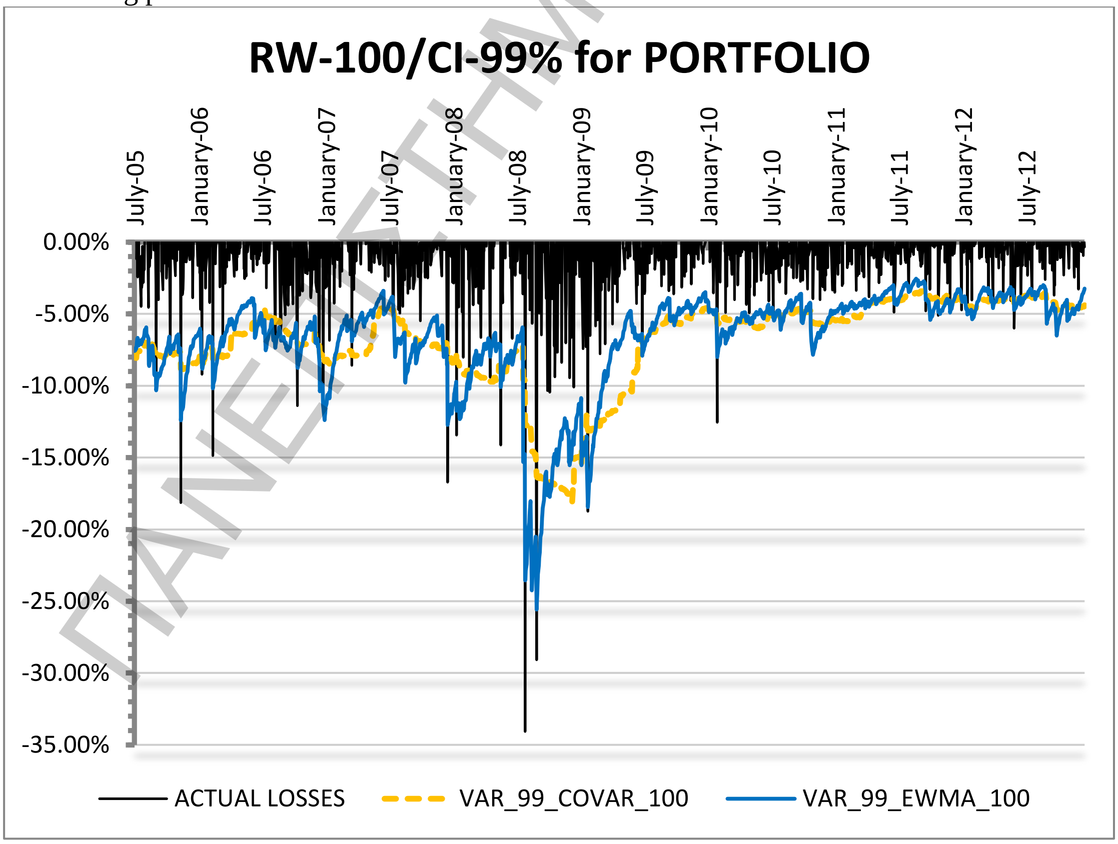


Figure 10-The 99% one day Var-Covar 100, the EWMA-100 and the actual losses over the backtesting period.

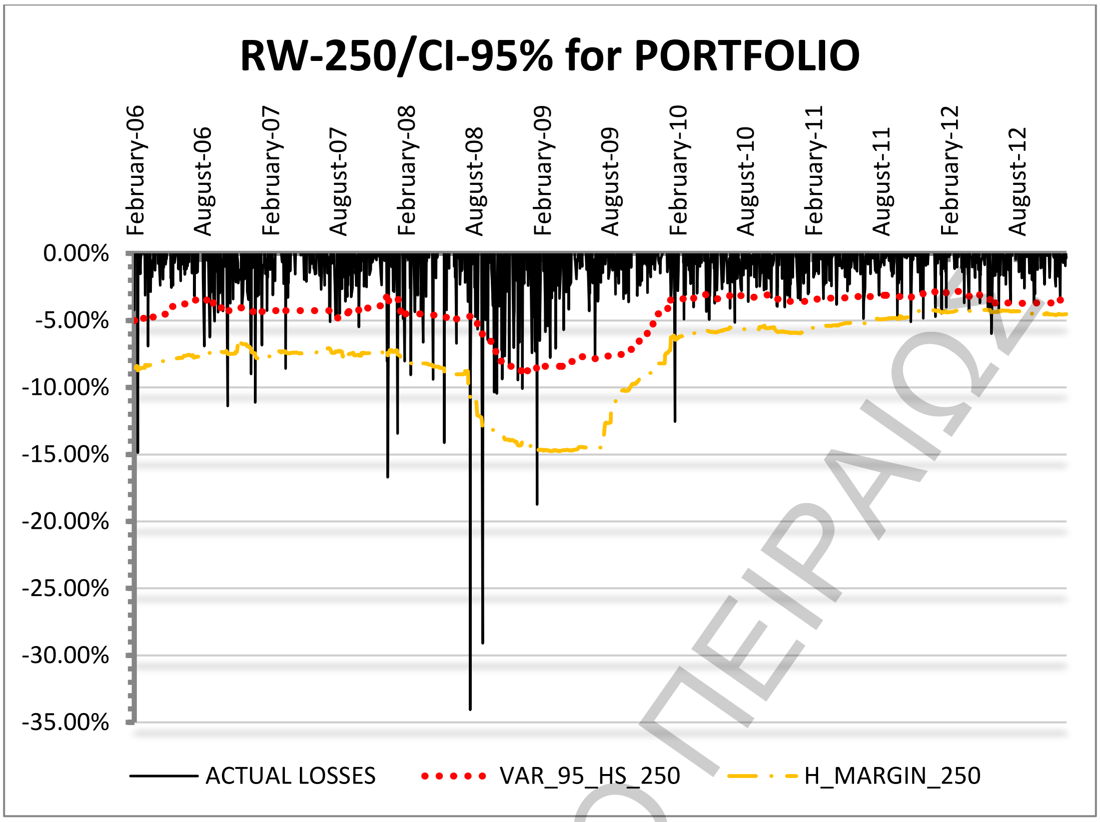


Figure 11-The 95% one day HS-250 VaR, the alternative method (h-margin-250) and the actual losses over the backtesting period.

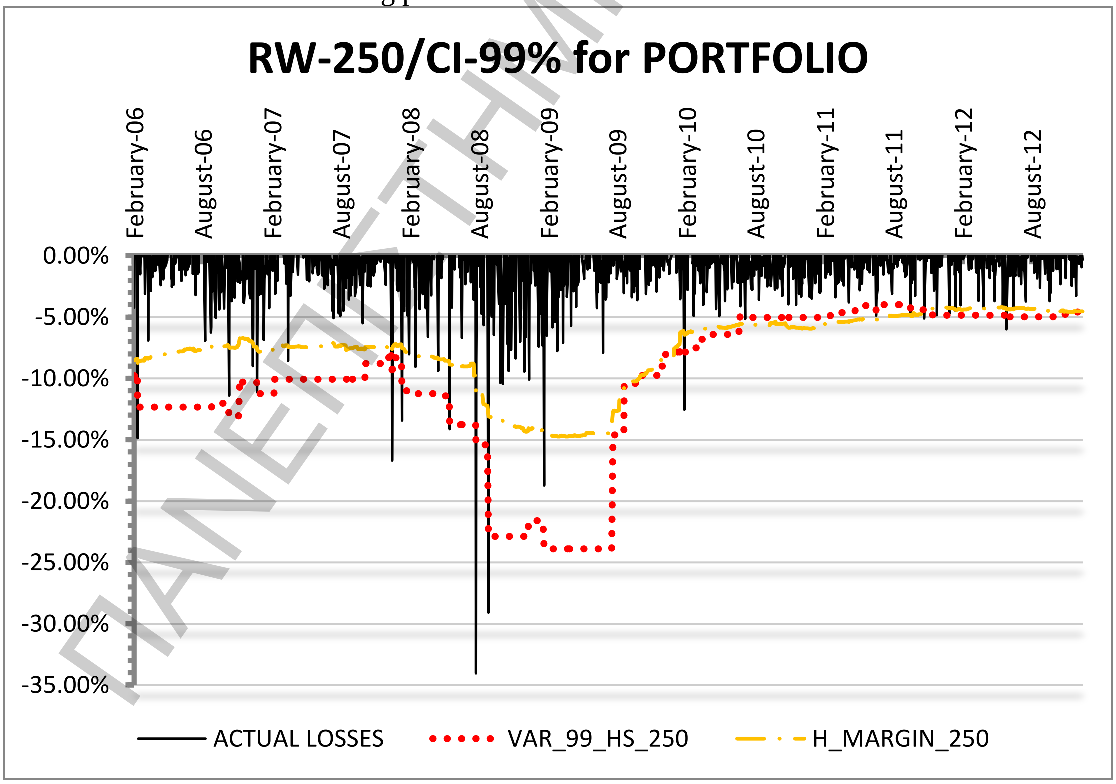


Figure 12-The 99% one day Var-Covar 250-VaR, the alternative method (h-margin-250) and the actual losses over the backtesting period.

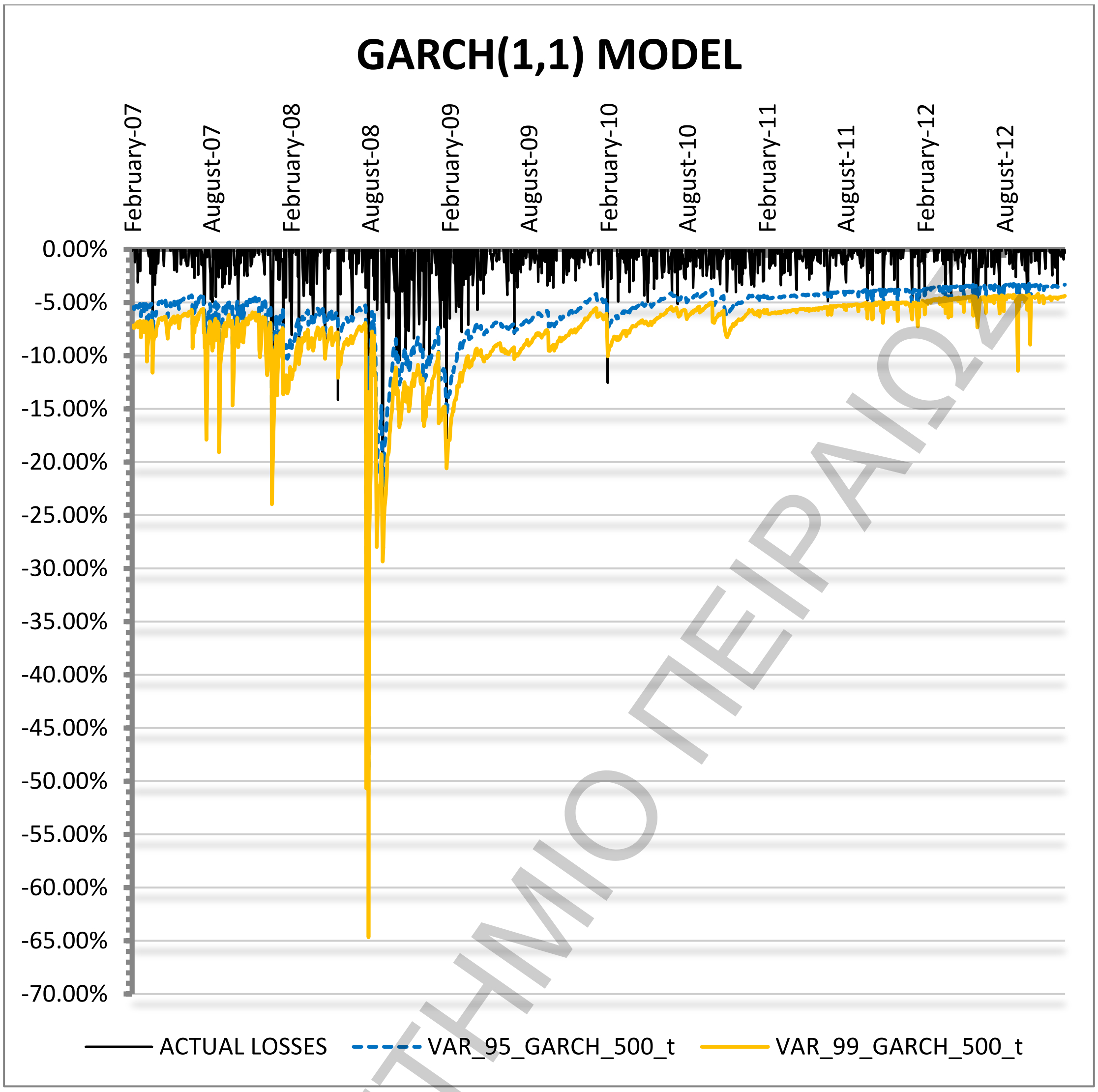


Figure 13-The 95% Garch-500-t-student-VaR, the 99% Garch-500-t-student-VaR and the actual losses over the backtesting period.

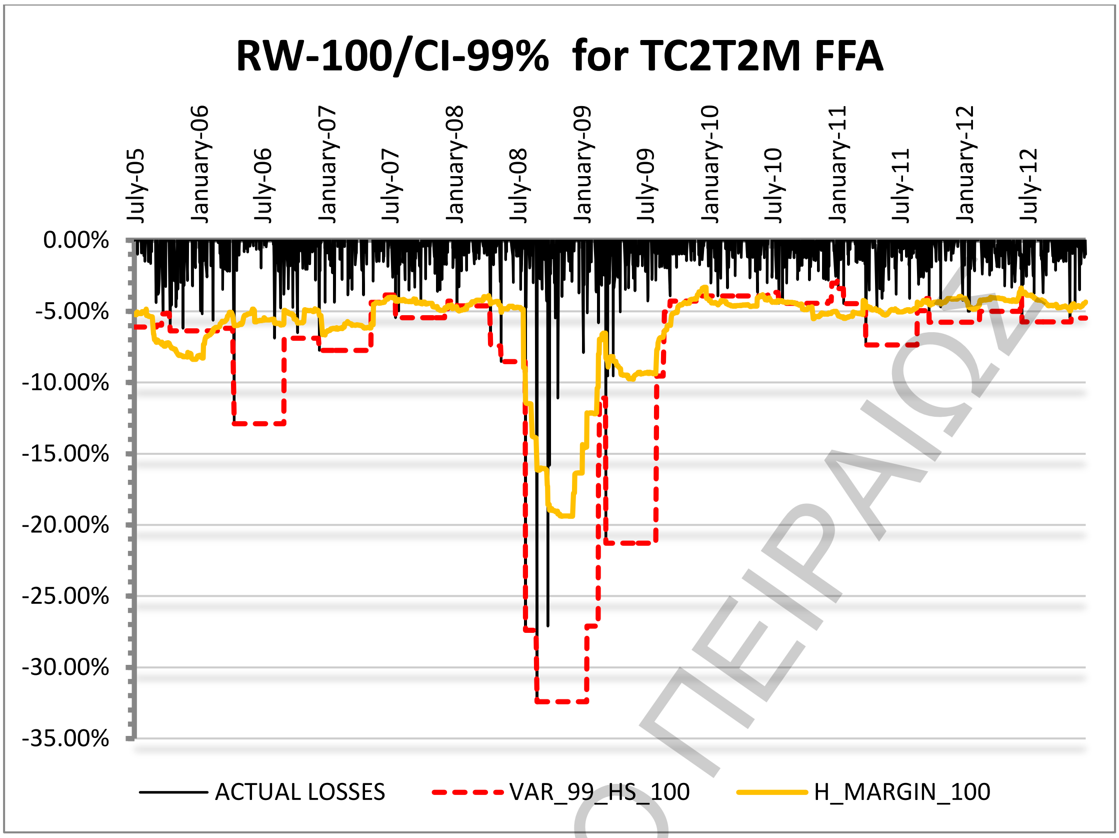


Figure 14-The 99% one day HS-100 VaR, the alternative method (h-margin-100) and the actual losses over the backtesting period.

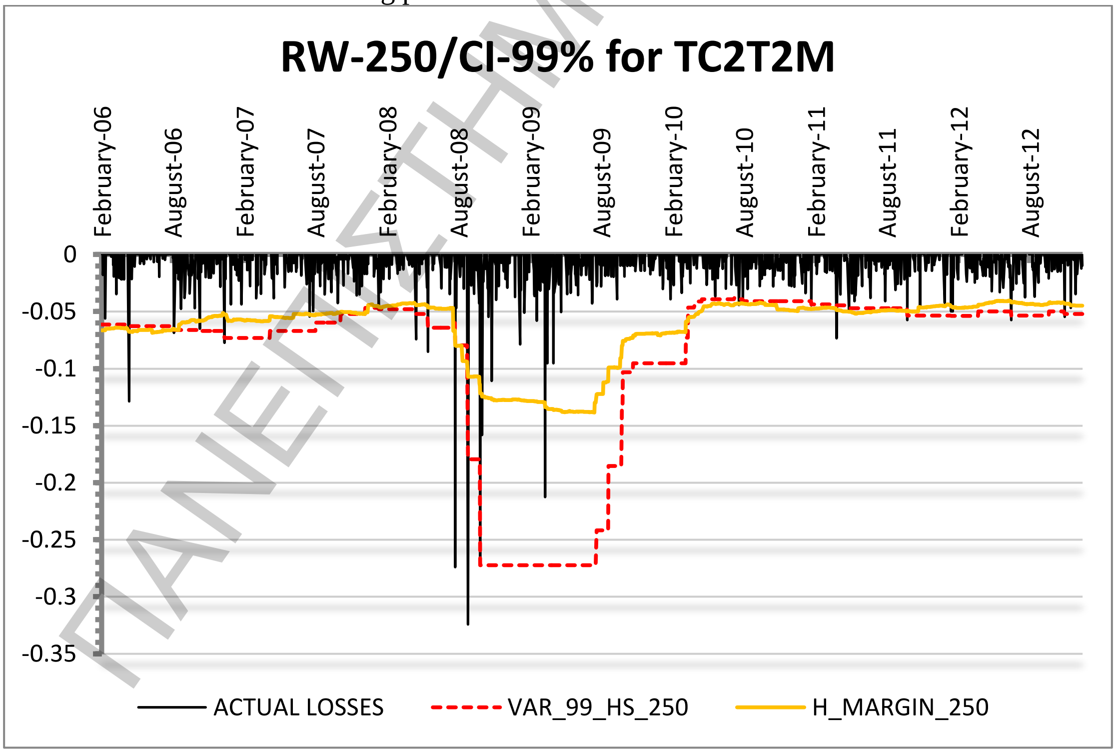


Figure 15-The 99% one day HS-250 VaR, the alternative method (h-margin-100) and the actual losses over the backtesting period.

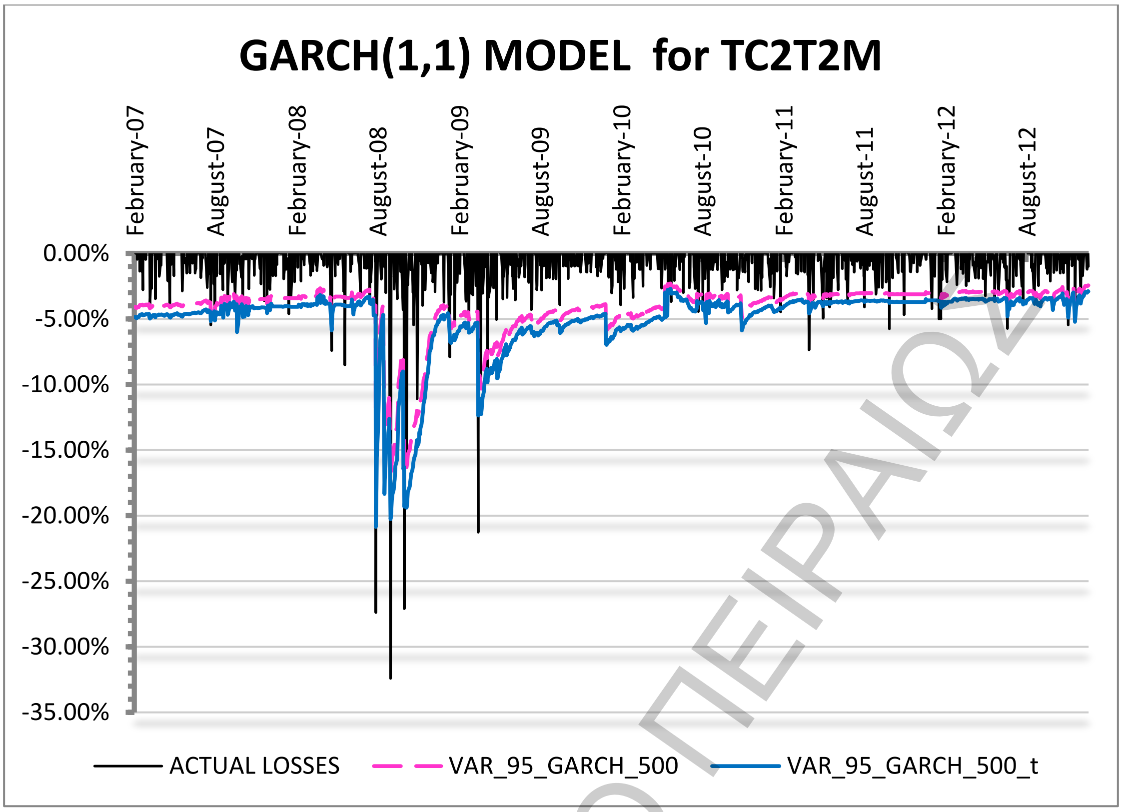


Figure 16-The 95% Garch-500-t-student-VaR, the 95% Garch-500-normal-VaR and the actual losses over the backtesting period.

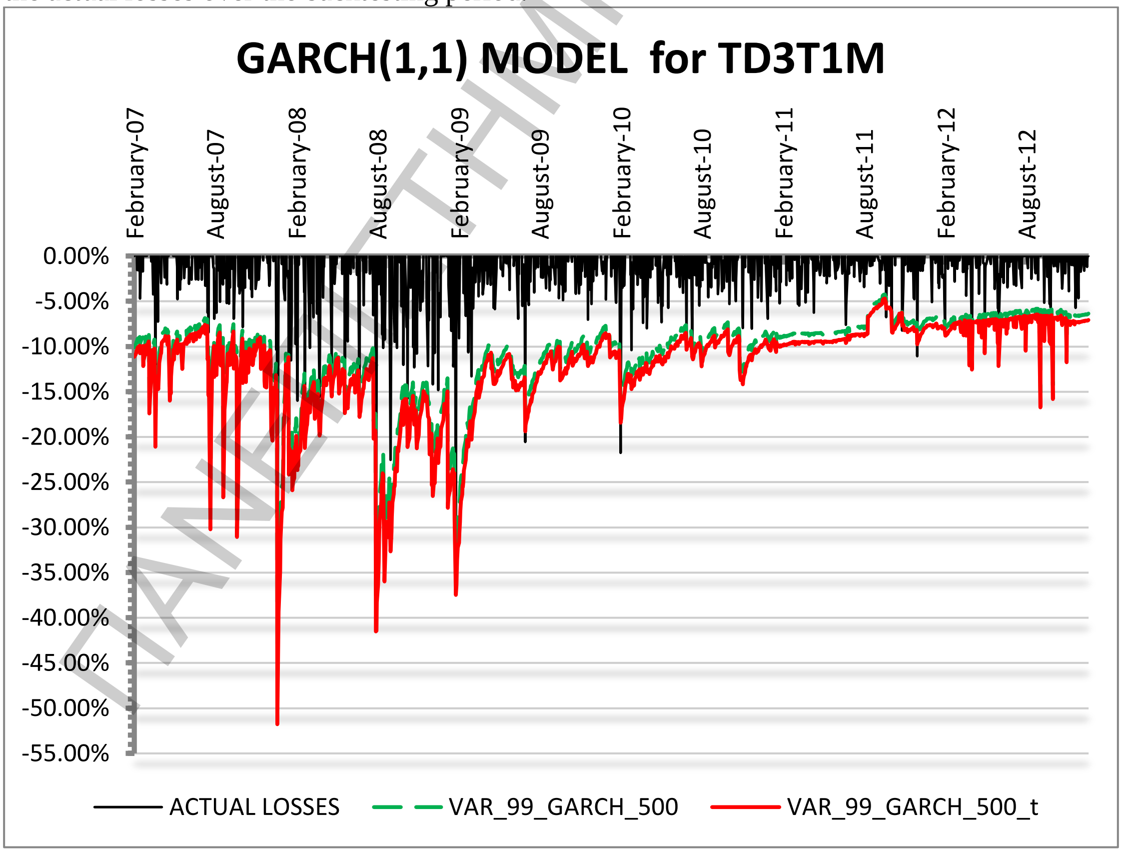


Figure 17-The 99% Garch-500-t-student-VaR, the 99% Garch-500-normal-VaR and the actual losses over the backtesting period.

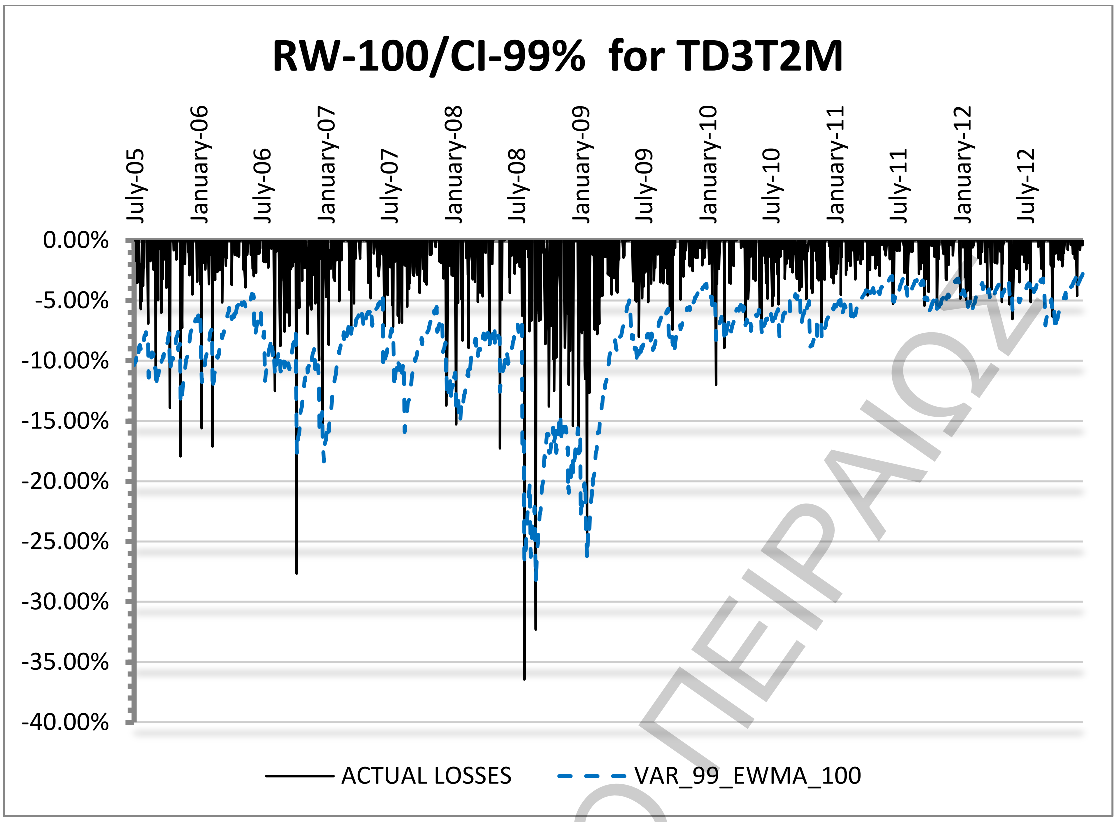


Figure 18-The 99% one day EWMA-100-VaR and the actual losses over the backtesting period.

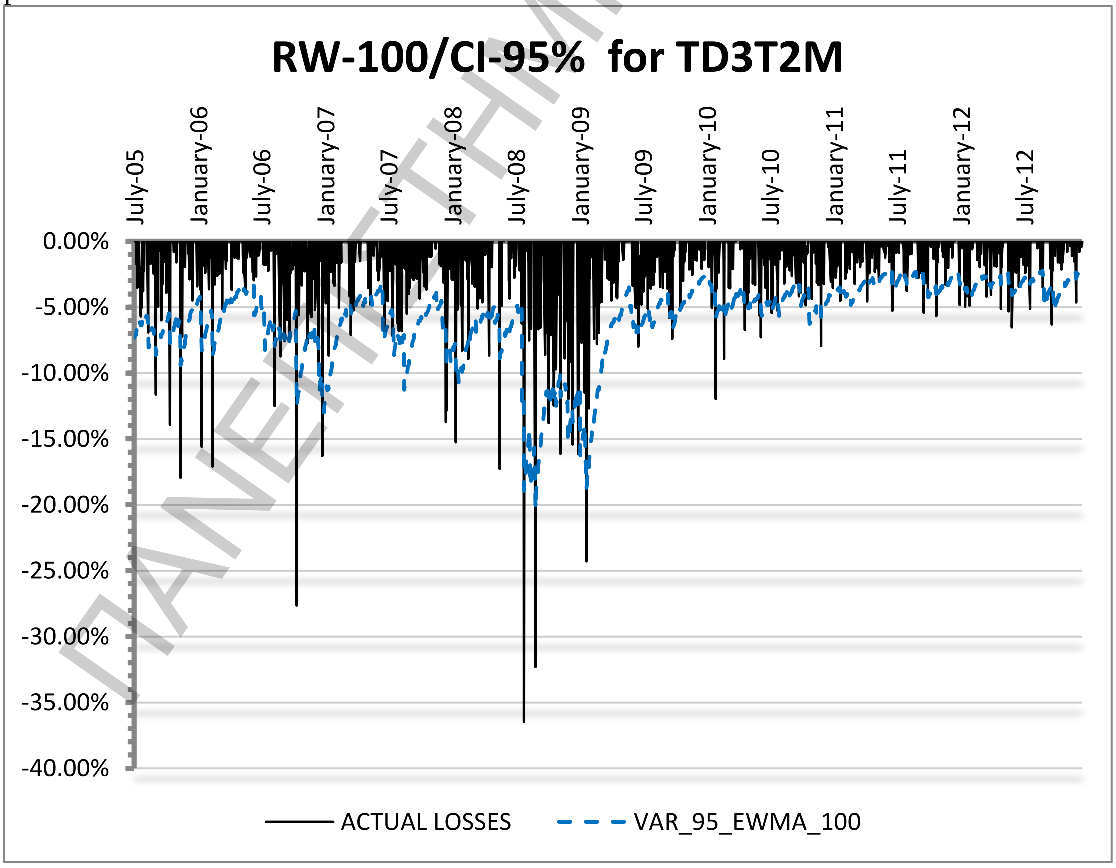


Figure 19-The 95% one day EWMA-100-VaR and the actual losses over the backtesting period.

7. CONCLUSIONS

Maritime transport is the most essential sector of national trade. Margining is of outmost importance in the volatile world of freight market and should be measured and set accurately. This dissertation approached freight margining system by a Value at Risk perspective. We applied a number of parametric and non-parametric methods in the field of FFAs in order to decide which is the most accurate in setting margins in freight derivatives market. As a general opinion, we can see that different volatility models achieve accurate VaR estimates with Historical Simulation and GARCH with t-student assumptions dominating among them. Var-Covar method experiences a wide range of rejections which results that it cannot cope with large movements in freight prices. This weakness in this particular VaR model can be attributed to the observed non-normality of freight profit and loss. An important conclusion emerging through our empirical analysis, which verifies similar research in the same field, is that in the race of margining through VaR estimation, HS comes first in terms of backtesting but after GARCH in conservatism terms.

---

[i] One use of "Black Monday" term is the Monday October,1987, when stock markets around the world crashed shedding a huge value in a very short time. The crash began in Hong Kong and spread west to Europe, hitting the USA after other markets had already declined by a significant margin.

[ii] These extreme events were named "Black Swans" by the Wall Street trader and author, Nasim Taleb.

[iii] For example, a portfolio consisting of 100 assets needs 49.500 covariances estimations for the calculation of VaR

[iv] Skewness is the measure of asymmetry distribution of a random variable

[v] Ordinary least squared method is a method for estimating the unknown parameter in a linear regression model.

[vi] Chi-squared distribution is with k degrees of freedom is the distribution of the sum of the k independent normal random variables. The chi-squared distribution is widely used in the common chi-squared tests for goodness of fit of an observed distribution to a theoretical one.

Acharya V.V., Lochstoer L.A. and Ramadorai T. (2011) "Limits to arbitrage and hedging. Evidence from commodity markets.", National Bureau of Economic Research.

Angelidis T. and Benos A. (2006) "Adjusted Value at Risk different regimes, skewness and kurtosis.", Applied Financial Economics, vol.16, p.p.835-851.

Angelidis T. and Skiadopoulos G. (2008) "Measuring The Market Risk of freight Rates: A Value at Risk Approach.", International Journal of Theoretical and Applied Finance, vol.11, No.5, p.p.447-469.

Bates D. and Craine R. (1998) "Valuing the futures market clearing house's default exposure during the 1987 Crash.", Journal of Money, Credit and Banking, vol.31, p.p.248-272.

Beenstock and Vergottis (1989) "An econometric model for dry cargo freight and shipping.", Applied Economics, vol.21, p.p.339-356.

Beenstock and Vergottis (1989) "An econometric model of the World Tanker Market.", Journal of Transport Economics and Policy, vol.23, p.p.263-280.

Bekiros S. and Georgoutsos D.(2005)"Estimation of Value At Risk by extreme value and conventional methods: a comparative evaluation of their predictive performance. Journal of International financial Markets, Institutions and Money, vol. 15, pp. 209-228

Bollerslev T. (1986) "Generalized Autoregressive Conditional Heteroskedasticity.", Journal of Econometrics, vol.31, p.p.307-327.

Branch A.E. (1998) "Elements of shipping", USA.

Brooks C, Clare A.D. Molle D.J.W and Persand G. (2005) "A comparison of extreme value theory approaches for determining value at risk.", Journal of Empirical Finance, vol.12, p.p.339-352.

Cabedo D. and Moya I. (2005) "Estimating Oil price 'Value at Risk' using the historical simulation approach.", Energy Economics, vol.25, p.p.239-253.

Christoffersen P. (1998) "Evaluating Interval Forecasts.", International Economic Review, vol.39, p.p.841-862.

Christoffersen P. (2003) "Elements of Financial Risk Management", Oxford.

Christoffersen P. and Pelletier D. (2004) "Backtesting Value at Risk: A Duration Based Approach.", Journal of Financial Econometrics, vol.2, p.p.84-108.

Cotter J. (2001) "Margin Exceedences for European Stock Index Futures using Extreme Value Theory.", Journal of Banking and Finance, vol.25, p.p.475-502.

Edwards F. R. (1983) "The Clearing Associating Future markets: Guarantor and Regulator.", Journal of futures markets, vol.4, p.p.369-392.

Engle R. (1982) "Autoregressive Conditional Heteroskedasticity with estimates of United Kingdom Inflation.", Econometrica, vol.50, p.p.987-1008.

Figlewski S. (1984) "Margins setting for Stock Index Futures and Options,", Journal of futures markets, vol.4(3), p.p.385-416.

Fishe R., Goldberg L., Gosnell T. and Sinha S. (1990) "Margin Requirements in Future Markets: Their Relationship to Price Volatility", Journal of futures markets, vol.10, No.5, p.p.541-554.

Gârleanu N. and Pedersen L.H (2011) "Margin-based asset pricing and deviations from the law of one price.", Review of Financial Studies, vol.24, p.p.1980-2022.

Gay G.D., Hunter W.C. and Kolb R.W.(1986) "A comparative Analysis of futures Contract Margins, Journal of futures markets, vol.6(2), p.p.307-324.

Goulas L., Skiadopoulos G. (2012)"Are freight futures markets efficient? Evidence from Imarex", International Journal of Forecasting, vol., pp. 644-659

Hardouvelis G. (1988) "Margin Requirements and Stock Market Margin Requirements and Volatility", Journal of financial Services Research, vol.3, p.p.139-151.

Hardouvelis G. and Dongcheol K. (1995) "Margin Requirements, Price Fluctuations and Market Participation in Metal Futures", Journal of Money, Credit and Banking, vol.27, No.3, p.p.659-671.

Hardouvelis G. and Dongcheol K. (1996) "Price Volatility and Future Markets", Journal of futures markets, vol.16, No.1, p.p.81-111.

Jagannathan R. (1991) "Implications of security market data for models of dynamic economics.", Journal of Political Economy, vol.99, p.p.225-262.

Jones R.A. and Perignon C. (2010) "Derivatives Clearing, Default Risk and Insurance.", Working Paper, HEC Paris.

Kavussanos M.G. and Alizadeh A.H. (2002) "The expectations hypothesis of the term structure and Risk Premiums in dry bulk Shipping freight markets.", Journal of Transport Economics and Policy, vol.36, p.p.267-304.

Kavussanos M.G. and Dimitrakopoulos D.N. (2007) "Measuring freight risk in the tanker shipping sector.", Working Paper, Athens University of Economics and Business.

Kavussanos M.G. and Nomikos N.K. (2000) "Hedging in the freight futures market.", Journal of Derivatives, vol.8, p.p.41-58.

Kavussanos M.G. and Visvikis J.D.(2006) "Shipping freight derivatives: A survey of recent evidence", Maritime Policy Management, vol.33, p.p.233-255.

Knott R. and Polenghi M. (2006), Working Paper No.287, Bank of England.

Kupiec P. (1991) "Prudential margin policy in a futures-style settlement system", Journal of futures markets, vol.13, Issue 4, p.p.389-408.

Kupiec P. (1995) "Techniques for verifying the accuracy of risk management models", Journal of Derivatives, vol.3, p.p.73-84.

Kupiec P. and Sharpe S. (1991) "Animal Spirits, Margin Requirements and Stock Price Volatility", Journal of Finance, vol.46, No.2, p.p.717-732.

Lamoureux C.G. and Lastrapes W.D. (1990) "Heteroskedasticity in Stock Return Data: Volume versus GARCH effects.", Journal of Finance, vol.1, p.p.221-229.

Longin F.M (1999) "Optimal margin level in futures markets: Extreme Price Movements." Journal of futures markets, vol.19, p.p.127-152.

Luu, Venus Y.H., Cheng T.C. Edwin (Eds) "Shipping and Transport logistics.", Springer, New York.

MandelBrot B.B. (1963) "The variation of certain speculative prices.", Journal of Business XXXVI, p.p.392-417.

Nelson D. and Foster D. (1994) "Asymptotic filtering theory for univariate ARCH models.", Econometrica, vol.62, p.p.1-41.

Rasmussen T. S. and Tversland E. (2007) "Managing Risk with freight futures form IMAREX. Testing hedging Effectiveness and the Unbiasedness hypothesis."

Shiller R.J. (1979) "The Volatility of Long-Term Interest Rates and Expectations Models of the Term-Structure.", Journal of Political Economy, vol.87, p.p.1190-1219.

Singleton K.J.(1980) "Expectations Models of the Term-Structure and Implied Variance Bounds.", Journal of Political Economy, vol.88, p.p.1159-1176.

SPAN (1995) "Standard Portfolio Analysis of Risk.", Chicago Mercantile Exchange.

Stopford M. (2009) "Maritime Economics", New York, 3rd Edition.

Unctad (2012) "Review of Maritime Transport.", United Nations Conference on Trade and Development, Geneva