University of Stavanger

Faculty of science and technology

Master's thesis

Study program/specialization: Industrial economics/Risk management	Spring semester, 2011 Open
Author: **Jone Haugland**	…………………… (Author's signature)
Faculty supervisor: Frank Asche External supervisor: Øystein Håland (Statoil)	
Title: **Value-at-risk: A coherent measure of risk?** Norwegian title: Value-at-risk: Eit samanhangande vågnadsmål?	
ECTS: 30	
Keywords: - Value-at-risk (VaR) - Subadditivity - Historical simulation - Backtesting - Oil futures	Pages: 47 + 11 enclosed Sandnes, June 8th 2011

Value-at-risk: A coherent measure of risk?

By Jone Haugland

University of Stavanger

Department of industrial economics, risk management and planning

Abstract

Value-at-risk is an instrument which is widely used by financial institutions for calculating risk. It has been known since the late nineties that this tool lacks an important logical property: subadditivity. This can cause major errors, leading to systematic underestimation of risk when multiple portfolios are combined. It is known to be caused by coarseness in the return distribution and is thus most problematic when using historical simulation.

This thesis investigates the severity of subadditivity violations from historical simulation, using the oil markets as a source of data. In order to measure this, a variant of the standard backtest has been used. The amount of subadditivity is found to be strongly dependent on the correlation between the individual portfolios, but independent of the choice of confidence level and sample size. A negative correlation virtually eliminates non-subadditivity altogether.

Keywords: Value-at-risk, subadditivity, historical simulation, backtesting, oil futures.

Acknowledgements

This thesis concludes my Master of Science degree in Industrial economics at the University of Stavanger (UiS). It has been written during the period: January-June 2011.

The idea behind the theme was given to me by Statoil ASA, and I was given a desk at their offices at Forus to investigate and write. The data used throughout this thesis has also mainly been supplied by them. All calculations have been performed using Microsoft Excel 2010.

I would like to express my gratitude toward Øystein Håland who introduced me to the problem, as well as Jack Andre Djupvik, Andreas Hodne and Lars Dymbe, for always letting me feel welcome at Statoil and for providing encouragement and tips.

I would also like to thank Roy Endré Dahl at UiS who was also studying subadditivity at the same time. Both his articles and his concrete tips have proven invaluable for my work. I wish him the best of luck with his PhD.

Jone Haugland
Sandnes
June 2011.

Table of contents

1 Introduction

1.1 Backgound

The task of correctly quantifying financial risk has been an increasingly important issue for organizations. A wide array of risk measures exist and are being used in various industries, from simple calculations of expected losses and volatility to extremely intricate models that uses numerous parameters. The most widely used risk measure today, and the one preferred by Statoil, is called *value-at-risk* (usually abbreviated *VaR*).

While the quantile concept behind VaR is itself quite old, it was not until after the stock market crash of 1987 that it gained widespread use in finance. In the first years after that, it was mainly used in the derivatives market, but after the financial bank J.P. Morgan in 1994 freely released their methodology in the *RiskMetrics Technical Document*, it very quickly became the standard risk measure for companies and regulators alike. In fact, after the Basel II accord was implemented with its *Amendment to incorporate market risk*, the use of VaR has become the main tool for determining capital charges against market risk.

One of the main issues with VaR as a risk measure is that it is not subadditive (Artzner, 1999). When splitting up a portfolio into smaller parts, it can sometimes be seen that the sum of the risks for the smaller parts does not add up to the risk for the combined portfolio. This undermines the whole fundamental principle of modern portfolio theory that states that combining portfolios should by itself not create extra risk (Markowitz, 1959). This may create a number of unhealthy effects, the most obvious of these being an underestimation of the actual risk suffered when using simple estimation tools (Daníelsson, 2010). A secondary effect is that incentives may be distorted and lead traders to engage in risky trade, while the senior management believes that risk is reduced (Garcia, 2005).

This thesis will investigate the use of VaR in the oil market. The oil markets are of great importance in the global trade in that it still is the major source of energy in the world. In fact, the general consumption of oil, both by individuals and industries, are rising on a daily basis, even when new reserves are becoming increasingly costly to produce (Žiković, 2010). Also, the oil markets have become an important area for speculators. The price of oil often exhibits negative correlation with that of stocks, and therefore has a tendency of moving in the opposite direction of stocks. Because of this, oil futures have become a favored investment object for hedge funds (Hall, 2011).

1.2 Statoil and value-at-risk

Statoil ASA is one of the major oil and gas companies in the world. On average, its total daily trade exceeds 2 million barrels of crude oil and condensate. The products traded ranges from its core areas, oil and gas products, to methanol and lately even emission allowance contracts. Financial instruments used for this purpose includes futures, forwards and swaps, as well as standard cash trades. Options trading is only used to a lesser degree, and will not be examined in this paper.

The oil trading at Statoil serves the purpose of optimizing the profit from running its refineries at Mongstad and Kalundborg. Here, the crude oil and condensate is processed into commercial fuel products like petrol, aviation fuel, jet fuel, diesel, heating oil and heavy bunker oil. Other products created in the refining process could be solvents, lubricants, bitumen, wax and asphalt. Statoil supplies these refineries from a global crude oil market, while the refined products on the other hand are mainly sold on the European market.

The trading process at Statoil is rather decentralized, where the individual traders have a substantial amount of freedom in deciding which products to trade in and which trading tools to use in order to minimize the risk. The risk analysis desk sets risk limitations for the various traders, and the traders in term have to act inside these limitations. Since the late nineties, value-at-risk has been the preferred way of quantifying target risk. This way of organizing the trade is often called the rent-a-trader system and has the advantage of exploiting the expert knowledge of the individual traders to its fullest, while not letting them get bogged down in a complicated bureaucracy.

The method of actually calculating value-at-risk is constantly under debate. In the beginning, the delta-normal parametric approach was used, but as it often failed to account for extreme outliers, it has been replaced by historical simulation. More advanced methods like weighting schemes and Monte Carlo simulation have been tested, but have proved to be somewhat unwieldy and often less effective than one would expect.

1.3 Text outline

This first chapter has created a general overview to value-at-risk as a risk measure, its use in Statoil and introduced the issue of the lack of subadditivity.

The next chapter will focus on the general characteristics of oil trading, presenting the various financial instruments used in that regard. A discussion of the underlying mechanisms that drives the oil prices is also included. The principle of risk is presented and related specifically to the oil markets.

Chapter 3 gives a theoretical background to how value-at-risk tries to capture the risk and quantify it. The diversification principle is presented, first as a general idea in modern portfolio theory and then through the notion of subadditivity. VaR as a risk measure is consequently shown to lack subadditivity in certain cases. The later sections then discuss tail behavior, giving some attempts to quantify and model the tail distribution.

Chapter 4 gives an overview of the methodology used in historical simulation. Comparisons between that and other methods are included, along with a discussion of certain variables. Lastly, this section presents backtesting as a way of evaluating the performance of the simulation in the past.

Chapter 5 presents the data from the real products and attempts are made to create portfolios that generates interesting results. Two tests are performed to check how common subadditivity violations are. These are the single-day test, which checks at what confidence levels the portfolios are non-subadditive, and the backtest, which tracks the development of subadditivity violations over time.

Some concluding remarks, a bibliography and a collection of figures not shown in the text itself constitutes the last three chapters of this thesis.

1.4 Sources of information

Quite a lot has been written about value-at-risk the last fifteen years. In fact, entire books have been devoted to the subject. I drew most of the background information from three books: Jorion (2001), Dowd (1998) and Alexander (2008). The latter of these was probably the most useful, as it contained more recent models and was generally more thorough. There are newer editions available from both of the other two authors, but I was unfortunately unable to acquire those.

Some good information on how the oil market functions was collected from Scofield (2007) and Edwards (2010). Useful was also the more general approach by Hull (2006) in explaining the different financial tools used for trading. The most useful scientific papers discussing the issue of non-subadditivity would be Artzner et al. (1999), Daníelsson et al. (2010) and Žiković & Aktan (2010). Many of the ideas on how to structure this thesis were drawn from Dahl (2008). A full list of references is found at the end of this paper.

2 Petroleum trade and risk

This chapter will give a general overview of the oil markets and risks associated with the trading of oil derivatives. Chapter 2.1 gives a rough outline of the different types of risk. Chapter 2.2 presents an overview of how the markets for oil products function, highlighting different products and different trader roles. Chapter 2.3 explains the various types of financial contracts like futures, forwards and swaps, and how the oil traders use them in order to minimize their risk. Chapter 2.4 details the underlying mechanism behind the fluctuations in the oil prices and explains the fundamental difference between short and long term development.

2.1 What is risk?

Risk is a term that is widely used within many areas. It is difficult to pin down a precise definition of what risk is, but it is commonly related to uncertainty, danger or randomness. Risk could be said to be the possibility of suffering an undesirable outcome. But it is a rather vague term, and it may mean a number of different things to different people, so reducing it to a single number will always mean losing a substantial amount of information. Still, the human mind struggles when it is told to make good decisions if a lot of factors need to be taken into consideration, and to managers of companies these simplifications have been invaluable.

From a financial viewpoint, the undesirable outcomes are almost always in the form of monetary losses. The Basel committee recognizes that financial risk should be divided into three main areas:

- *Market risks* are related to changes in market prices. The value of a firm’s financial position changes continuously due to changes in value of its underlying components. The changes in the underlying components could be caused by changes in exchange rates between two currencies (often known as currency risk), changes in interest rate for lending money, changes in stock and bond prices (equity risk), or the changes in the price on specific goods (commodity risk)

- *Credit risks* are related to the possibility a trading partner not being able to fulfill his end of a contract and monetary losses caused by this. The most common example of this would be when a borrower is unable to pay back a debt he owes and defaults on it.

- *Operational risks* are related to monetary losses caused by failures or inadequacies in internal processes, people or systems, and losses caused by unforeseen external events. This is a broad umbrella category that would contain most of the risks not falling into the other two categories.

This paper is mainly concerned with the trade of petroleum products, which means that the market risk category, and more specifically the commodity risk category, is what applies the most. Even so, value-at-risk is a risk measure that is widely used for calculations of other types of risk as well. The results arrived at here could prove relevant both for other kinds of market risk and even in the fields of credit risk and operational risk.

2.2 Overview of oil markets

The trading of oil is dominated by the relationship between supplier and consumer countries. The major consumers are Europe, North-America and the industrializing countries of East Asia, while the major suppliers are generally less developed countries in the Middle East and South America. Norway is in a special position in that it is an industrialized country and at the same time a major exporter of oil.

Due to the historical instability of many of the major oil supplying countries, most of the oil refineries are located in the consumer countries rather than near to where it is produced. There is therefore a global trading of unrefined oil, commonly known as crude oil. Crude oil is transported long distances, generally by tankers, and bought by the refiners in the consumer countries. More refined oil products play a smaller role in the global trade and are usually traded locally. As a result, there exists a global market for crude oil, which is determined mostly by the global supply and demand, whereas the more refined products are more prone to local variations of supply and demand.

A variety of commercial products are created in the refinement process. Often they are divided into four categories:
1. Light distillates: Gasoline for automobiles and aeroplanes;
2. Middle distillates: Gas and diesel oils, jet and heating kerosenes
3. Heavy distillates: Marine bunker oil; crude oil used directly as fuel
4: Others: Naphtha; liquefied petroleum gas (LPG); solvents; petroleum coke; bitumen wax; lubricants.

As can be seen from Table 2.1, the demand for commercial oil products comes mainly from three regions: USA/Canada, Europe and Southeast Asia. The first dominates the markets for light distillates, while Southeast Asia does so for heavy distillates. What affects the particular demand in these regions is consequently what affects the global demand as a whole.

Crude oil comes in many varieties. Two common parameters to describe the quality of the crude are *density* and *sulfur content*. Lighter crudes have a higher proportion of simple hydrocarbons, and are more easily converted into commercial petroleum products like car fuel. Heavier crudes require in comparison more processing, and because of the higher viscosity they are also more taxing on processing equipment. Sulfur is an undesirable pollutant in crude oil, and requires the crude to undergo expensive cleansing. Crude oil with high or low sulfur content is known as sour and sweet crudes, respectively. The two commonly used benchmarks are Brent contracts which are used in the North Sea, and WTI contracts for oil produced in the USA. Other types of crude produced in these regions are described relative to these standards.

Edwards (2010) identifies the players engaged in the oil trade to fall into four groups: producers, consumers, refiners and marginal traders. The interaction between these participants is what drives the pricing mechanism of oil products, along with outside factors. The roles these participants play are in short:

- **Producers**: Get the crude oil up to the surface and sell it for profit that exceeds the operating costs.
- **Consumers**: Buy commercial petroleum products. This could be everything from a manufacturer to a private individual.
- **Refiners**: Buy and transport crude oil to refineries, where they convert it to commercial products. These are then sold for profit.
- **Marginal traders**: Buy processed or unprocessed products and try to resell the products at higher prices without altering the products themselves.

Because of the advantages of economies of scale, many of the large oil companies are involved in more than one role. Statoil is mainly known as an operator on offshore installations, and thus a producer, but is also engaged in refining, marginal trade, and even some consumption as well.

The production of oil has risen in a fairly linearly the since 1980, but many of the individual fields have reached its peak production years ago and is now declining, including most fields in the North Sea. The North Sea as a whole reached its peak in 1999.

	Light	Middle	Heavy	Others
USA and Canada	10 970	7 188	1 390	5 326
Latin America	1 227	1 872	704	973
Europe	3 758	7 572	1 882	3 203
Middle East	1 210	1 816	1 473	1 240
Fmr. Soviet Union	923	1 146	699	1 168
Africa	631	1 180	477	475
Southeast Asia	6 600	8 810	3 526	5 021

Table 2.1: *The daily consumption of refined oil products by geographical region in 2005 (in thousands bbl). Source: Schofield (2007)*

2.3 Financial contracts

Traders in the oil markets use similar instruments for deals as other commodity traders. The four most common types of trading contracts are *cash trades*, *futures*, *forwards* and *swaps*. Cash trades involve an exchange of physical crude oil for cash in the spot market. Futures and forwards are negotiated contracts to buy a certain amount of crude at a certain price some set time in the future. While futures are contracts that are bought and sold on an exchange, like the Oslo Stock Exchange, forwards are traded in an over-the-counter (OTC) marketplace.

Forwards are agreements directly between two parties, and they can often be quite customized. Any revision of the terms of the contract is often a complicated affair that might involve professional lawyers.

Futures, in comparison, are always standardized contracts made between one party and the exchange itself. Thus, the future contracts may themselves be traded freely in the market. To allow this liquidity the exchange requires the traders to deposit a *margin* to ensure the trader fulfills his obligations. This margin is typically between 5% and 10% of the value of the contract itself, and should be able to cover any changes in price over a one day period. Swaps are agreements to exchange cash flows in the future, the sizes of which are based on different price indices. Swaps are usually used to hedge against the price of a commodity.

The portfolio of a crude oil trader at a given time is called the trader's *position* in the market. Positions can be described as either *long* or *short*. If a trader takes a long position, he benefits from an increased price. The long position most typically means that the trader has ownership of an asset, like some given amount of oil. If a trader takes a short position, he benefit from a decline in price. This usually means the trader owes another party an obligation for an already set price. An example would be being obligated to deliver of some amount of oil to a refinery.

Because a party incurs considerable credit risk by entering into the trade of forwards and futures, measures are often done to reduce risk exposure. This is done through *hedging*. Hedging means taking a position that is opposite to an already existing position, thus negating it. For example, if a trader has entered a forward contract, agreeing to supply a refinery with a given amount of crude oil for a fixed price in the future, the trader faces large risks if the price of this crude rises. Also, this is not a very flexible position, and it may be difficult and costly to get out of. If the trader at the same time takes a long position in the futures market on a similar crude oil, the risk of price increase is virtually eliminated if the oil prices are correlated.

The prices of derivatives like forwards and futures can be described relatively to the spot price. If the market believes the price of oil will rise in the future, the price of the oil future will be larger than the current spot price. Similarly, if the market estimates the price to fall in the future, the future price will be smaller than the spot price. In the commodity market this is known as *contango* and *backwardation*, respectively. Figure 2.1 displays how the market can change from contango to backwardation.

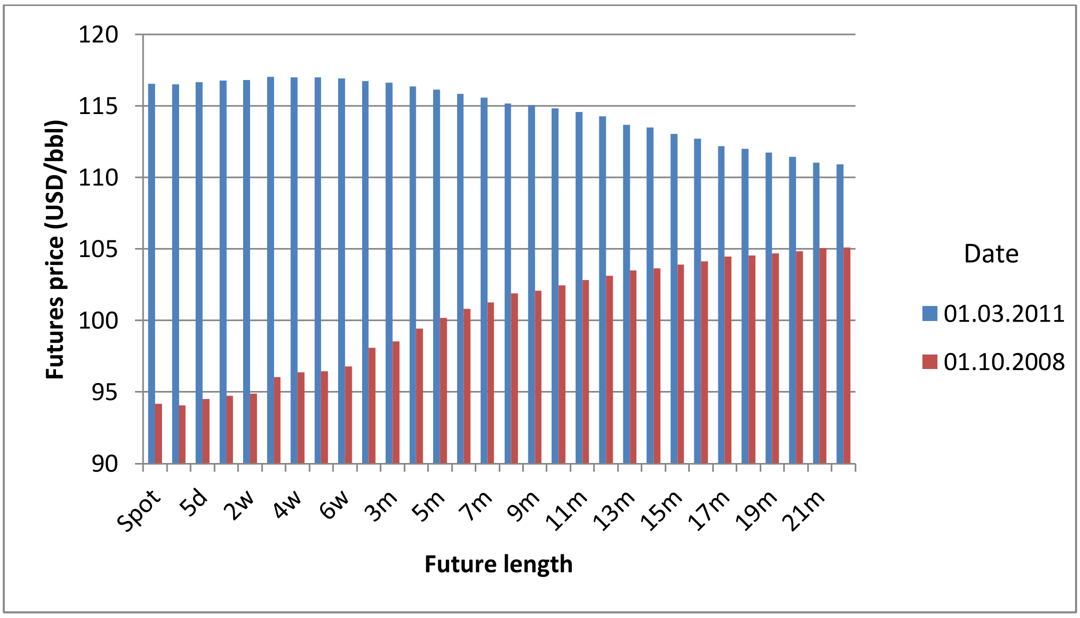

Figure 2.1*: The future prices of Brent crude oil traded on two different dates. The length of the future contract is measured in days (d), weeks (w) and months (m)*

2.4 Price drivers

The price of oil follows a highly irregular pattern, and fluctuates up and down over time. Not only does the price itself change, but the volatility of the price is also varying with time. This property is known as heteroskedasticity. The price can be relatively stable for long periods of time, only to be followed by violent changes that send ripples across the world. Classical economics explains this behavior through changes in supply and demand.

In the short-term demand and supply are quite inelastic. Whatever gasoline may cost at a certain time, cars cannot just switch to another fuel. Likewise, the fuel of ships, aeroplanes or factories cannot be easily changed in the short run. The supply of oil is also inelastic as the task of producing marginal amounts of oil is relatively inexpensive compared to the large initial investments made in drilling and assembling or transporting the platforms. Operating platforms is more or less the same whether it is producing at high or low percentages of full capacity. Thus, oil is pumped up roughly at a constant rate no matter the circumstances. Figure 2.2a shows how both the demand and supply curves must be steep, so that large changes in price only have a little impact on the quantities demanded or supplied.

Fundamental changes in the short-term are often known as *shocks* and will cause the curves themselves to move one way or the other. Figure 2.2b shows how a sudden reduction in the supply of oil causes a big increase in the oil price. Such a shock might be caused by something unexpected like a loss of output from rig closures or disruption of supply routes due to wars or natural catastrophes. Figure 2.2c shows how a sudden increase in demand shifts the curve to the right and causes the price to be increased substantially. Wars or natural catastrophes could be candidates for this as well, doubling the effect it would have on price.

The financial crisis of 2008 saw a major shock to demand, causing its curve to shift to the left. Consequently, the prices dropped from about 140 to 40 USD in just 6 months.

Even if they are inelastic in the short-term, the long-term supply and demand are very elastic. The demand for oil can often be seen to move in a cyclical pattern. Oil is an essential input in many industries. Thus, when the activity in the economy as a whole increases, it will make the demand for oil increase, and when it decreases, the demand for oil will also fall. A major source of increased demand is the growth of the economies of the newly industrialized counties (NICs), like China, India, Brazil and Turkey. An enormous increase in oil consumption in these countries has fueled the markets with extra demand.

The oil price is also dependent on the prices of substitute commodities. If there are cheaper alternatives available, the demand for oil will decrease. The most common substitutes for petroleum are most directly other non-renewable energy commodities like gas and coal, and more indirectly renewable energy commodities like solar and wind power or bio-fuel.

Governments may also influence the demand for oil through taxation or regulation. High taxes or extra regulation will cause the prices to increase and thus decrease the quantity demanded. Governments will create incentives with their taxation policies, as it could impose taxes on oil products while holding it low for substitutes like renewables.

Oil is a major source of heating, so demand will depend on climate. Most of the consumption of oil is done on the northern hemisphere and therefore demand tends to be higher during winter for this part of the world. According to Riley (2006), the prices have a tendency of moving in a cyclical pattern, where they are at their highest during the winter months. Especially harsh winters will drive the prices even further upward.

An interesting variable is the role speculation plays in the creation of demand itself. Speculators buy oil products in the hope that the prices will increase in the world market. It has been said that much of the recent short term spikes in oil prices can be attributed to the varying demand by investors, most notably hedge fund managers. Hamilton (2008) states that speculators have been increasingly engaged in the market for oil futures, purchasing surplus contracts in the hope of earning a profit when the contracts are ready to be fulfilled. According to a recent article by Hall & Rankin (2011), speculation may account for 70 % of all oil futures contracts bought, and whose owners never intend on taking control over the physical oil. Twenty years ago this only accounted for 30 %, so there is no doubt that speculation more and more has to be taken into consideration.

The limited supply of the world's hydrocarbon reserves is a long-term variable on the supply side. As new oilfields are discovered the size of proven reserves may increase, but the actual reserves in the earth's crust will diminish. By allocating economic resources to exploration, it will hopefully give the results of finding more economically viable fields, thus creating extra supply in the long run.

Since petroleum is a finite resource, the question of how much is left is always a debated question, and an important factor in the price of oil. New discoveries of oil reserves are continually made, although at a declining rate for most of the world. There are still large areas in many oil producing nations that remains unsurveyed in terms of potential future production. Much of these potential oil reserves is in control of state owned companies in authoritarian countries, like Saudi Aramco, and their estimates on the reserves left in a field is often not subject to the same amount of transparency as a western company would be. There is thus considerable uncertainty in determining the reserves left, and changes in the estimates, like a fully new discovery, will cause changes in the prices. Generally, when the estimates for reserves left go up, it will cause the price of oil to decrease.

Technology is also a major factor in long term supply. As new technology is invented, the ability to extract petroleum from the fields will increase. Today, the efficiency of this process is lower than one third on a world average. This means that two thirds of the oil is still left in the field when it is abandoned, which it will be once continued production no longer is economically profitable.

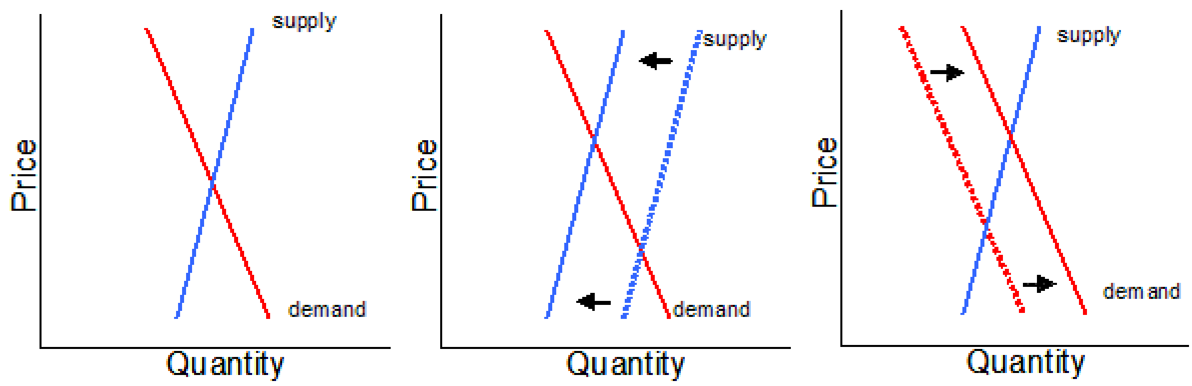

***Figure 2.2(a,b,c)**: Supply and demand curves for oil*

3 Value-at-risk and subadditivity

Value-at-risk has become the standard tool in determining how risky an investment is. Chapter 3.1 contains a general description of portfolio diversification as a way to reduce risk. Value-at-risk itself is defined in Chapter 3.2. Chapter 3.3 and 3.4 explain why applying the diversification principle on VaR directly sometimes can generate errors. A heavily tailed and disjointed return distribution is considered to be the source of such violations of subadditivity. Chapter 3.5 defines skew and kurtosis as tools to describe a certain distribution, and Chapter 3.6 contains a mathematical definition of heavy tailedness.

3.1 Reducing risk through diversification

The idea of diversifying investments, and *not placing all eggs in the same basket*, is intuitive and has to a certain degree always been practiced in economics. Still, it was not until the 1950s that it progressed beyond mere heuristics, and became a science of itself through the work of Harry Markowitz. Prior to this, theory and what was considered good business practice diverged. To create a good portfolio, theory said it should be built with those assets which offered the highest return and the lowest risk. The composition of the portfolio itself did not matter. Experienced investors knew that in real-life markets this was a far too simplistic view.

Markowitz (1959) details how portfolios could be constructed based on the total risk-return ratio, rather than merely focusing on the individual assets. A good composition could make the risks of different assets null each other out, thereby creating a less risky portfolio. He suggested that the value of an asset to an investor could best be evaluated by its mean, its standard deviation, and also to its correlation to other securities in the portfolio.

The correlation between two assets *X* and *Y*, with expected returns μ_X and μ_Y and standard deviations σ_X and σ_Y, can be expressed:

$$Corr(X,Y) = \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_X \sigma_Y}$$

The correlation should be a value between *-1* and *1*. A value of 1 means there is perfect correlation between the two assets, and they rise and fall in the same direction. A value of *-1* means perfect negative correlation, so the assets rise and fall in opposite directions. A value of *0* means there is no relationship between the assets, they rise and fall independently of each other.

If each asset is given a percentage weight w_X and w_Y, so that the sum of the weights will be *1*, the standard deviation of the total portfolio can be calculated as:

$$\sigma_{X+Y} = \sqrt{w_X^2 \sigma_X^2 + w_Y^2 \sigma_Y^2 + 2\rho_{X,Y} w_X w_Y \sigma_X \sigma_Y}$$

To demonstrate the reduced risk through diversification, consider a reasonably safe asset with $\sigma_X = 10\%$ and a more risky asset with $\sigma_Y = 20\%$, and zero correlation between them. If the portfolio consists only of the safe asset, portfolio standard deviation will be 10%, but if 20% of the weight instead had been on the risky asset, it would be 8,9%. In other words, by including some amount of the more risky asset, the total risk is actually reduced.

The diversification principle is widely accepted as almost a universal truth about how to view financial portfolios. It still has a few critics, however. Holton (2008) argues that the eggs-in-the-basket analogy not necessarily apply for financial risks. Instead he provides another example: Suggest someone is stranded on a desert island and finds three pools of water. He suspects one or more of them might be poisonous but do not know which one. He now has three choices: (1) randomly select one pool and drink exclusively from it, (2) randomly select two pools and drink from them, or (3) drink from all three pools. Common sense says he incurs the least risk by drinking only from one pool of water, yet portfolio theory says he should diversify and thus drink an equal amount from all three. This view on diversification is interesting but has not reached the mainstream and remains a fringe idea.

3.2 What is value-at-risk?

Value-at-risk is a single measure that tries to capture the risk of loss on a portfolio under normal fluctuations in the market. Perhaps the most common definition of this term is the one described by Jorion (2001):

"VaR summarizes the worst loss over a target horizon (T) with a given level of confidence (α)"

VaR therefore describes the lower quantile of the projected profit/loss-distribution (P&L). Using a historical sample size (*T*), the losses exceeding the VaR will only occur with a probability of $1 - \alpha$. Formally, this can be denoted:

$$VaR(\alpha) = \min \{x \mid F(x) \geq \alpha\}$$

where $F(x)$ is the cumulative probability function over the profit and loss x. Figure 3.1 gives an illustration of VaR.

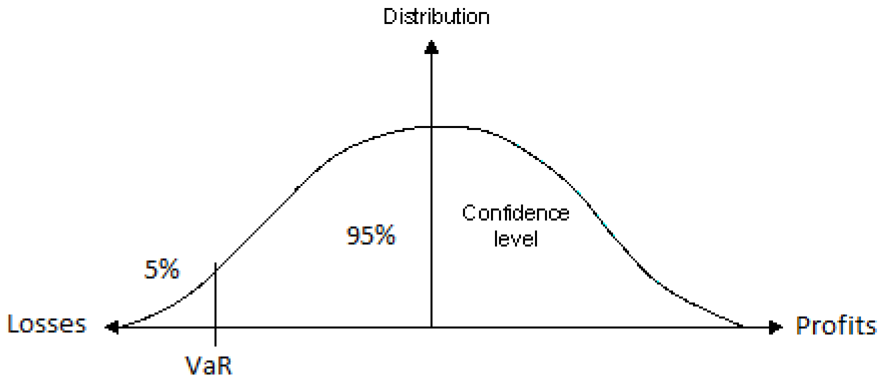

***Figure 3.1** – VaR on a P&L distribution at 95% confidence level.*

The drawback of this measure is that it does not give information about the losses that occur with a probability of less than $1 - \alpha$. VaR is incapable of distinguishing between situations where losses in the tail region are only a bit worse and situations where tail losses are overwhelming. Zikovic (2008) and many others argue that in practice this has a tendency to make risk estimates overly optimistic.

3.3 Coherent measures of risk

In 1997 Artzner et al. published a widely influential article called *"Thinking Coherently"*. There they defined that in order for a risk metric (r) to be coherent, as they called it, for any risks A and B (i.e. two random, independent monetary losses), it would need to satisfy these four logical properties:

1. Monotonicity: If $A \geq B$, then $r(A) \geq r(B)$
2. Positive homogeneity: for $a \geq 0$, $a\, r(A) = r(aA)$
3. Translation invariance: $r(A + a) = r(A) - a$
4. Subadditivity: $r(A + B) \leq r(A) + r(B)$

The first property, monotonicity, states that if risk A is always larger than risk B, the risk measure should be able to reflect that under all circumstances. Homogeneity is about proportional scaling. The risk of 10 shares of some stock should be exactly 10 times the risk of a single share. The translation invariance is about adding cash to a portfolio to act as insurance against losses. If some amount of cash in a portfolio is invested in a safe asset *(a)*, like a bank account, the portfolio risk should be reduced by this amount. Finally, subadditivity is the principle of reduced risk by diversification. The risk measure of a combined portfolio should never exceed the sum of its components. A merger should never in itself create extra risk.

In their follow-up paper in 1999 called *"Coherent measures of risk"* Artzner et al. prove that value-at-risk satisfies the first three properties, but not the last one about subadditivity. They were thus able to claim that VaR was not a coherent measure of risk. Other methods of measuring risk can be shown to be coherent, and have been adopted by many risk managers. Expected shortfall (ES) is perhaps the most notable of these, and numerous academic papers are advocating it over VaR (In addition to Artzner et al. (1999), see Hull (2006) or Embrechts et al. for more info about ES). Nevertheless, value-at-risk continues to this day to be the risk metric of choice, mostly due to its comparative simplicity and ease-of-use.

There are various ways to demonstrate lack of subadditivity. A good example of this was presented by Daníelsson et al. (2010) and is only a slight alteration of Artzner's original proof. Consider a portfolio of two identical but independent assets, X_1 and X_2. These are assumed to follow the standard normal distribution most of the time ($\mu=0, \sigma=1$), but both are prone to random shocks, which occur with a probability of 0,04. A shock will generate a loss of 10. This can be denoted:

$$\{X_1, X_2\} = x + \beta, \quad x \sim \mathcal{N}(0,1), \qquad \beta = \begin{cases} 0, & with\ prob.\ 0{,}96 \\ -10, & with\ prob.\ 0{,}04 \end{cases}$$

Expected loss on these assets will be: $E(X_1) = E(X_2) = 0 + E(\beta) = 0 \cdot 0{,}96 - 10 \cdot 0{,}04 = \underline{-0{,}4}$. The variance is calculated: $var(X_1) = var(X_2) = var(x+\beta) = 1 + var(\beta) = 1 + 10^2 \cdot 0{,}04 = \underline{5{,}0}$, which translates to a volatility (st. deviation) of about 2,24. Since only 4 % of the values will be centered around -10, calculating the value-at-risk at a 95% confidence level on total will be identical to calculating the (95+4)%-level for the distribution centered around 0. Thus, one gets: $\mathrm{VaR}(X_1) = \mathrm{VaR}(X_2) = z_{0,95+0,04} = 2{,}326$.

VaR for the combined portfolio would lie somewhere between -9 and -10. This will necessarily mean:

$$VaR(X_1 + X_2) > VaR(X_1) + VaR(X_2) = 2 \cdot 2{,}326 = 4{,}672$$

As seen, the portfolio VaR is greater than the sum of the individual VaRs. This example is relevant for areas where there are possibilities of very severe events. The probability of 0,04 for a shock occurring was chosen because it is only slightly smaller than the significance level for the individual risks. VaR would therefore be included in the main distribution. For the combined portfolio, however, the probability of *at least one shock* is 2 * 0,04 = 0,08. In this case, VaR would be situated somewhere in the "shock-distribution" centered around -10. Figure 3.2 illustrates how the VaR is moved when the portfolio is combined.

The term *superadditivity*, as a word describing the opposite of subadditivity, has not yet gained widespread use among risk analysts. In this thesis, it will sometimes be applied to instances when VaR lacks subadditivity, along with the terms *non-subadditive* and *subadditivity violation*.

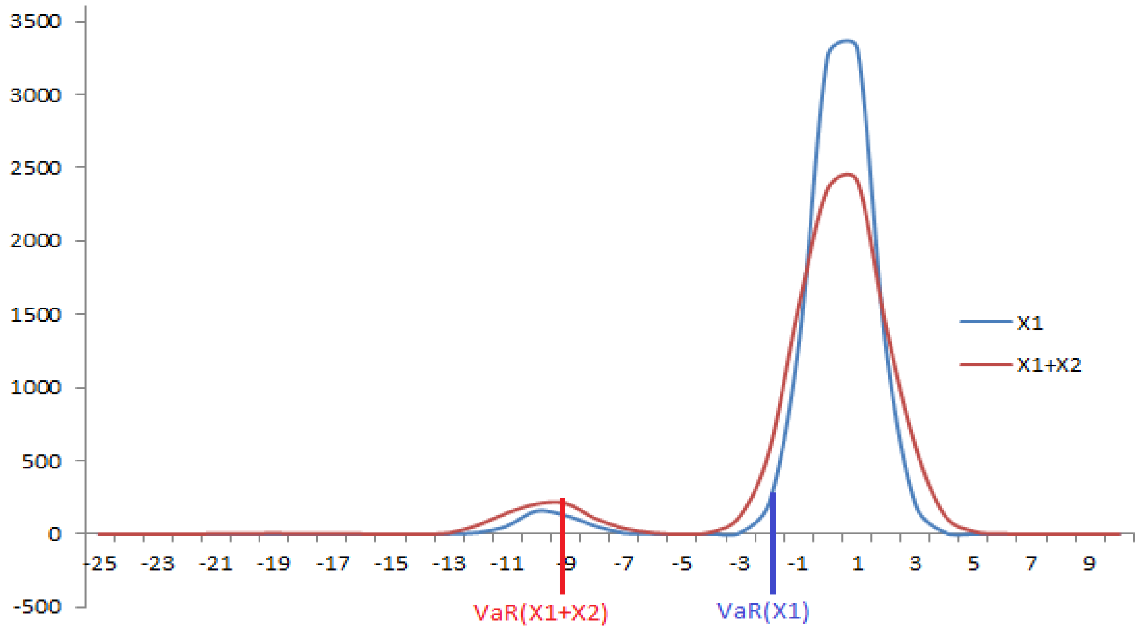

Figure 3.2: *VaR for a single asset (X1) and for a combined portfolio (X1+X2).*

3.4 The need for coherence in practice

Some might claim that the need for coherence under all circumstances is not really necessary. Daníelsson et al. (2010) argues that subadditivity is only expected to be violated under two quite rare circumstances. These are: (1) Assets, whose probability distribution has a super-fat tail, and (2) when the confidence level is in the interior of the return distribution. VaR is otherwise expected to be subadditive, and could thus be assumed to be coherent for most risk managers

In practice, it is only (1) that is relevant for most companies because the confidence level is always set fairly close to 1, and is therefore in the tail region. The historical return follows a distribution that has a fatter tail than the normal distribution, but most of the time this cannot be classified as super-fat tail (see section 3.6). It is still the case that subadditivity for VaR is violated under some circumstances, and it is often under these that the need for a good risk measure is the greatest.

The advantage of having a subadditive risk measure is that by adding the individual risks together, the sum of these will always overestimate the risk for the combined portfolio. This would then be used as a conservative estimate of the combined risk, and one could always be certain that the real combined risk was equal or smaller than this number. In an organization where decision making is decentralized, for example where individual commodity traders have a large amount of personal freedom to engage in trading, the sum of the reported individual risks could then be used as a maximum for the total risk the organization faces. Since value-at-risk is not a subadditive risk measure, merely adding the VaR-numbers together gives nothing of real value and could actually be truly misleading if used indiscriminately.

Traders might be tempted to break up their portfolio into separate accounts for each risk incurred. The traders are ordered to fulfill margin requirements, but since the individual risks could add up to less than the total risk, they might actually be able to reduce their margin requirements by splitting their portfolios into parts. Similarly, this might give incentives for banks to break up in order to reduce their regulatory capital requirements. The total capital requirements of all the smaller banks could be less than the total capital requirement for the large bank, while in reality the risks are generally greater for small banks. Garcia et al. (2005) states that subadditivity is a required property in connection with the aggregation of risks across desks, business units, accounts, or subsidiary companies. Also, they argue that it ought to be a greater concern to regulators where firms might be motivated to break up into affiliates to satisfy capital requirements.

3.5 Skewness and leptokurtosis

The normal distribution is often used to describe a number of different phenomena. Its use in risk estimation is limited due to its tendency to neglect extreme outcomes. Asset return distributions often follow something that resembles the normal distribution, but which is skewed to one side and contain leptokurtosis, or heavy-tailedness. A number of papers have investigated the distribution of historical financial data and the conclusion has been that skewness and leptokurtosis are general characteristics of financial data.

Papers specifically discussing the loss distribution of energy derivatives, on the other hand, are far less numerous. Giot and Laurent (2003) investigate returns on WTI-crude and find that it displays negative skewness and leptokurtosis. The parametric approach of RiskMetrics was found to be very ill suited for simulating this loss distribution, especially at higher confidence levels.

A similar simulation on Brent benchmark crude oil gives the same results. In the histogram in Figure 3.3, the blue curve denotes the fitted normal distribution and the red columns show the direct historical data of P&L the in the period: October 2002 – December 2008. As seen, the actual data is skewed slightly to the right, thus having a negative skew. The actual data also exhibit occasional extreme losses and extreme gains that are not captured by the normal distribution.

Both skewness and kurtosis can be expressed in mathematical terms, using the concept of moments. The n-th mathematical moment M_n is defined as:

$$M_n = \frac{E(X-\overline{X})^n}{\sigma^n}$$

The moment is normalized, meaning it is a dimensionless number.

The first moment is equal to the mean value and the second moment is equal to the variance. The third moment denotes the skewness, and the kurtosis could be described using n = 4. Using adjustments for a limited sample size N, the skewness could be calculated:

$$Skew = \frac{E(X-\overline{X})^3}{\sigma^3} = \frac{N}{(N-1)(N-2)}\sum_{i=1}^{N}\frac{(X_i-\overline{X})^3}{\sigma^3}$$

If skewness is a positive number, the distribution is skewed to the right, while negative skewness indicates skew to the left. A pure normal distribution, or one otherwise symmetrically distributed around a mean value, would have a skew = 0.

Similarly, the forth moment, the kurtosis, could be calculated:

$$Kurtosis = \frac{E(X-\overline{X})^4}{\sigma^4} = \left(\frac{N^2+2N+3}{(N-1)(N-2)(N-3)}\sum_{i=1}^{N}\frac{(X_i-\overline{X})^4}{\sigma^4}\right) - \frac{3(N-1)(N-3)}{N(N-2)(N-3)}$$

If the kurtosis is a positive number, the distribution exhibits fatter tails than the normal distribution. On the other hand, a negative number means the distribution decreases more quickly with deviations from the mean. A second effect of the fourth moment is that an increase will cause the top of the distribution will become more pointy, whereas a decrease will make it rounder.

Skew and kurtosis is easily calculated in Excel using the SKEW and KURT functions, respectively.

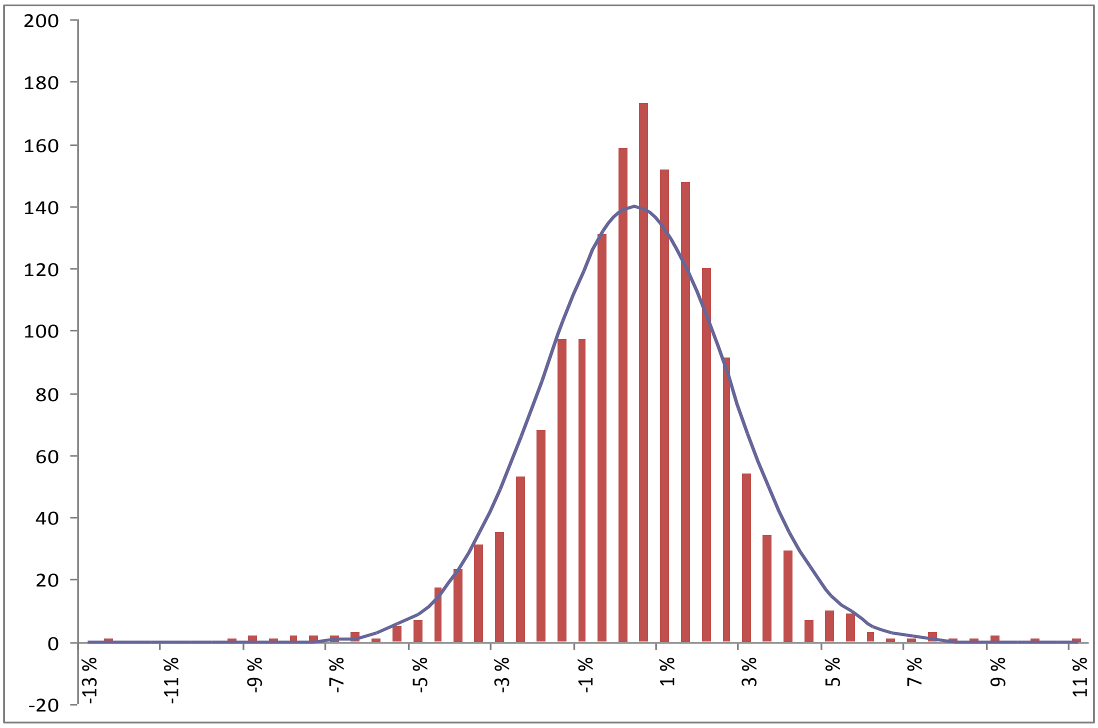

Figure 3.3: *Real data compared to the normal distribution. The real data shows skew to the right and leptokurtosis compared with the normal distribution. The fourth moment is calculated to be 2,346.*

3.6 Fat and super-fat tails

The normal distribution is often used to describe a number of different phenomena. Its use in risk estimation is limited due to its tendency to neglect extreme outcomes. Asset return distributions often follow something that resembles the normal distribution, but which has leptokurtosis, or heavy-tailedness.

The common definition of leptokurtosis is when the kurtosis is higher than the normal kurtosis, that is when the fourth moment as calculated with the formula in the previous chapter, is larger than 0. This is a simple and workable definition, but in many cases perhaps too simple. The fourth moment only captures the mass in the center of the distribution compared with that in the tails. If the tail is not continuous and smooth, but truncated like in the example in Figure 3.2, the relative mass of the tail may be small, and the fourth moment will only show a slight increase in kurtosis. As seen, in these cases a combined portfolio may display lack of subadditivity.

Daníelsson et al (2010) introduces an alternative definition of fat tailedness. Their definition states that a distribution is fat tailed if the cumulative distribution is varying regularly when the time period approaches infinity. Formally, this can be described:

$$\lim_{t\to\infty}\frac{F(-tx)}{F(-t)} = x^{-\beta} \quad => \quad F(-x) = x^{-\beta}g(x), \quad for\ x > 0$$

Here, $\beta > 0$ is a constant known as the *tail index* of the distribution. In order to make the implication above true, $g(x)$ must be a function that is slowly varying which converges to 1 at infinity: $g(tx)/g(t) \to 1$ *if* $t\to\infty$. An example of such a function would be a logarithm. The density function can be obtained by deriving the cumulative function. In the extreme tail region x is sufficiently large that this could be approximated as:

$$f(-x) = \frac{dF}{dx} \approx \beta g(x) \cdot x^{-\beta-1}$$

This means that the density declines at the power rate $x^{-\beta-1}$ in the tail region. This creates a considerably fatter tail than the normal distribution. The fatness of the tail depends on the tail index, where larger β would make the function decline more rapidly.

The tail index is often related to the moments, and the function could be described by how many *bounded* moments it has. The n-th moment is said to be bounded if $M_n > \beta$. The moments larger than the tail index are known as *unbounded*. The mean is known to be bounded for most financial assets, and often the variance is bounded as well. Under these circumstances, the tail is elongated compared to the normal distribution but it is still a smooth declining graph,

For $0 < \beta < 1$, however, there are no bounded moments. In the extreme tail regions, the density function do not follow a smooth graph, but the data falling here may be inter-dispersed by very rare events that results in extremely large losses. These are sometimes known as super-fat tails.

VaR can be shown to be subadditive for all stable distributions when the first moment is bounded. Ibragimov (2005) displayed that this is true even for extremely volatile assets, even up to the variance approaching infinity, as long as $\beta>1$. If the distribution is not too heavy-tailed, most economic models are quite robust toward heavy-tailedness assumptions. However, when the distribution instead has long tailed density, these implications are reversed.

Markets that grant loss distributions with an unbounded mean are not very common. Markets for weather insurance contracts are characterized by having long periods of relative calm. But once a weather catastrophe occurs, the holders of such contracts could suffer extreme losses. Even in these markets, the tail index is known to be hovering just around 1. This has led many analysts to defend the use of value-at-risk as the standard risk measure against those who criticize it for its universal lack of subadditivity.

The commodity markets have not been particularly well researched in terms of its tail index. One of the classic texts in tail behavior is Mandelbrot (1963), which analyzed that of historical cotton prices. Here, the tail index is estimated to be approximately *1,7.* Given the properties of oil markets, as discussed in Chapter 2.2, one might suspect that especially the market for oil products would have an even lower tail index.

4 Historical simulation

This chapter gives an overview of the historical simulation, which is the most common method of calculating value-at-risk. Chapter 4.1 describes the actual method of performing the simulation, and Chapter 4.2 discusses why historical simulation has become so popular along with some important weaknesses. Backtesting is described in Chapter 4.3 as a way of checking how well the simulation has performed in the past.

4.1 General

Value-at-risk has since its infancy gained widespread recognition as the primary tool for measuring market risk, even considering its known lack of universal subadditivity. But even if there is agreement on the use of VaR, the matter of actually calculating VaR itself is very much under debate. Three main categories of methods have arisen since the early 1990s: The parametric method, the historical method and the use of Monte Carlo simulation, with a wide array of different approaches within each category. Each of these methods offers certain strengths and weaknesses, and the difficulty of reaching a consensus on which one is the best is due to the fact that there are differences in what the goal of VaR is. For one, the choice of method may depend on the number and type of assets that is to be examined, and secondly to the objective which VaR is supposed to accomplish. The risk manager will have to weigh and compare the tradeoffs between the various methods for each individual case.

Historical simulation is based on direct empirical data and the quantile of the sample. To do this simulation, data of the historical daily returns on a portfolio of n risk factors must be collected over an observation period T. The relative weight of risk factor i in the portfolio is denoted w_i, while the return on the risk factor i over the period t is denoted $R_{i,t}$. The portfolio return (or more accurately P&L) over period t can thus be calculated as:

$$R_t^p = \sum_{i=1}^{n} w_i R_{i,t}$$

Here, observations are made from period $t=0$ to $t=T$, so a total number of T portfolio returns are acquired. T is the size of the sample and could be any number. The standard sample size used at Statoil is T=1000, approximately four year of trading days. Portfolio P&L can then be plotted into a histogram and the daily α-VaR can be read off as the (1- α)-th lower quantile of the distribution. For example, if T=1000 and α=0,95, then VaR will be 50th largest daily loss. In Microsoft Excel there is a function called SMALL which returns the n-th smallest number from an array of numbers. This has been used to simulate long positions, while a similar function called LARGE is used for short positions.

The daily return on the individual risk factors is calculated using the logarithmic difference between the price on day t and day t-1. Historical prices are either open source, collected in-house or attained from companies specialized in collecting data, such as Platts. The formula for log-returns is:

$$R_t = ln\left(\frac{V_t}{V_{t-1}}\right)$$

The two main alternative methods to historical simulation is parametric VaR and Monte Carlo simulation. Calculating VaR parametrically requires the analyst to assume the return distribution to be stated with a particular mathematical function. The most common of these would be the normal distribution where the expected value and standard deviation would be the parameters used. Monte Carlo simulation refers to any method that randomly generates trials, but by itself does not tell anything about the underlying methodology. Components of these could be incorporated into the historical simulation such as through volatility adjustment or weighting schemes, further discussed in Chapter 5.6.

4.2 Strengths and weaknesses

Historical simulation is a simple method that is easily understood not only by the risk managers themselves, but by non-specialists as well, making it attractive for senior management in many organizations that may not have in-depth knowledge of the process. This method does not require calculation of correlations between the individual assets. For complicated positions with numerous risk factors, where estimating correlations otherwise would have to be made using variance-covariance matrices, historical simulation eases the computational workload considerably. Using less complicated estimation techniques also reduces the chances of incurring model risk, the risk of using an inadequate or inappropriate model to measuring risk.

HS does not require any assumption on the distribution of the profit and loss, such as normal, student-t or any other. It does not even need to assume that the returns are independently distributed, making the tendency of volatile returns to cluster together much less of a problem. Fat tails, which are difficult to estimate parametrically, are also well reflected using direct data sets.

There are still limitations and problems associated with historical simulation. The most severe of these is arguably the need for complete reliance on one particular data set. The HS approach thus assumes that the future will be like the past, that the particular data set gives an accurate picture on the future risks. When markets go through longer periods of relative calm, this assumption generally holds quite well. But since the quiet periods often are abruptly followed by periods of violent fluctuations, the historical approach can be rather slow to adapt to this and the VaR estimates tend to be too low for quite some time. On the other hand, if markets have been volatile in the past, but is about to go into a calmer period, the estimate remains too high.

The choice of sample size can be a crucial one for creating an effective historical model. Figure 4.1 and 4.2 shows how different the historical VaR may be for different sample periods. Not only is there an absolute difference between the three graphs, but the calculated VaRs sometimes seem to move in opposite directions. Figure 4.3 compares the VaRs calculated using the historical simulation and ones calculated through the delta-normal parametric method. As expected because of heavy-tailedness, historical simulation overall gives a larger VaR than the parametric. Using a larger sample size reduces the difference between the two methods.

The lack of consensus on the choice of sample size is due to a tradeoff involved between accuracy and flexibility. To account well for losses incurred in the tail region of the distribution, a large sample size is necessary. Otherwise, there will not be enough observations to make a continuous graph. The distribution will instead be disjointed. As seen in Figure 3.3, the data in the tail region can be very disjointed, even when the data goes back six years. It would be quite illogical to conclude that this exact distribution would continue in the future. The higher the confidence level, the longer observation period is needed.

On the other hand, since the futures markets are shown to be heteroskedastic and undergo systematic changes in volatility, a long observation period will make it rely too heavily on returns in the distant past. As seen in Figure 4.1, longer sample sizes makes the model slow to react to changes in volatility and it also tends to lag longer on high VaR after the worst fluctuations have ceased. This is even more clearly illustrated in Figure 4.4, where data from the relatively volatile recent period have been examined. Only the very small sampled VaR is able to capture the increased volatility during late 2008 and early 2009 in a satisfying way, albeit with a certain lag. Using sample sizes of 250 and 500, the charted VaR shows very little resemblance to the volatility actually experienced. The rise and fall of the historical VaR is therefore often rather arbitrary, and even more so with increased sample size.

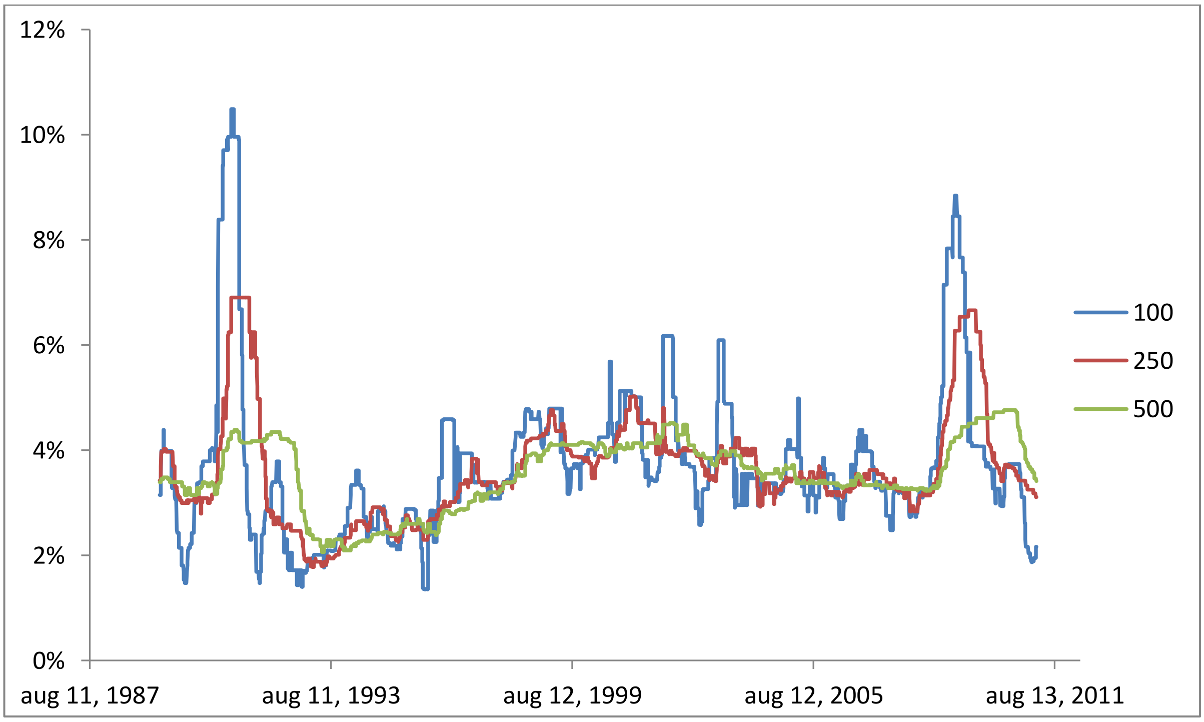

***Figure 4.1**: Historical 95%-VaR for Brent crude, using three small sample sizes.*

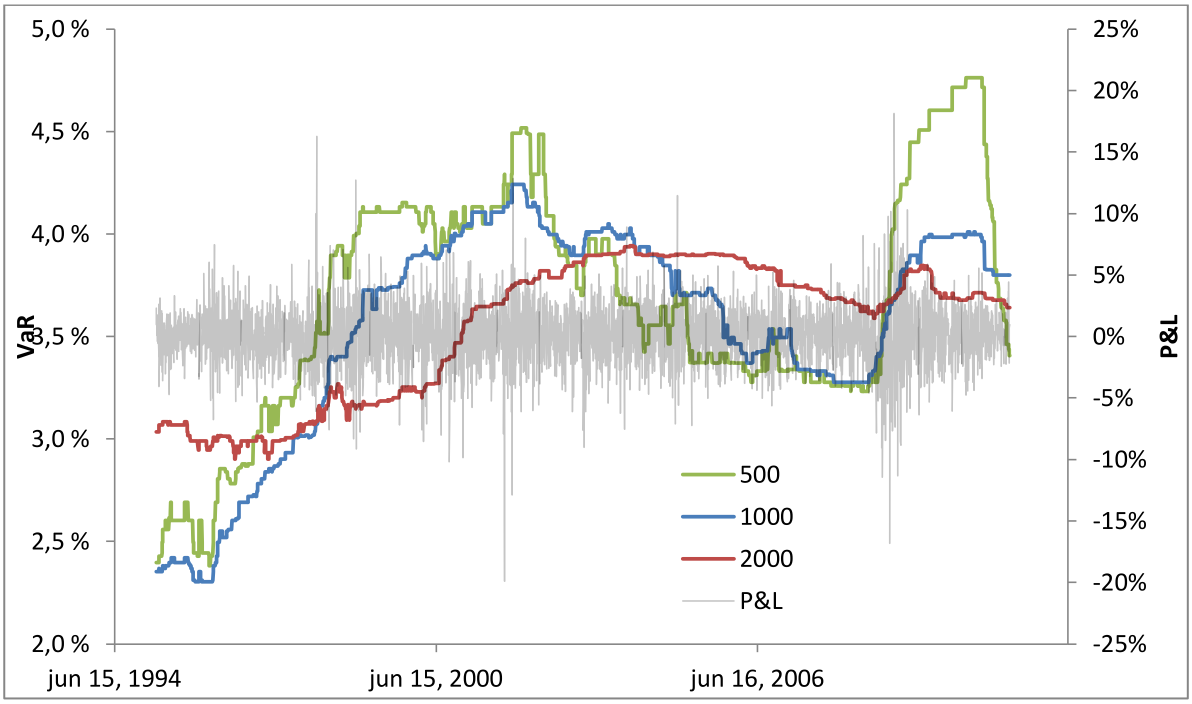

Figure 4.2: *Historical 95%-VaR for Brent crude, using three large sample sizes. A comparison with the actual P&L is also included in grey.*

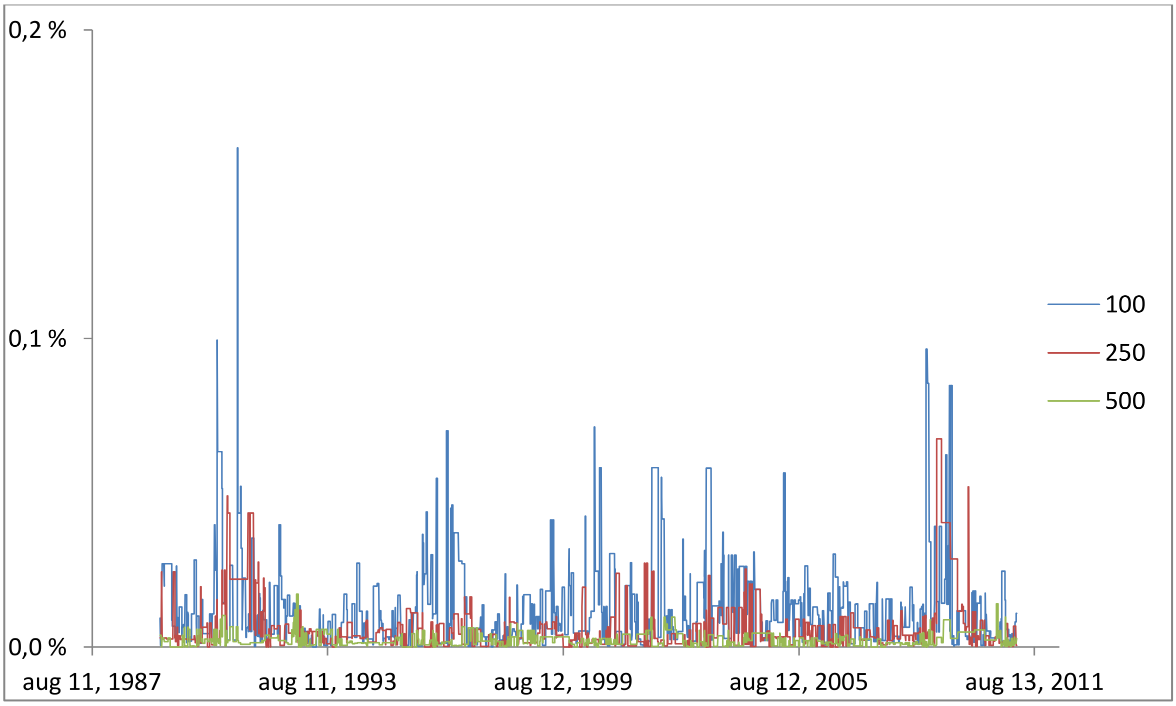

Figure 4.3: *Difference between historical 95%-VaR and delta normal 95%-VaR for sample sizes 100, 250 and 500 on daily return on Brent crude.*

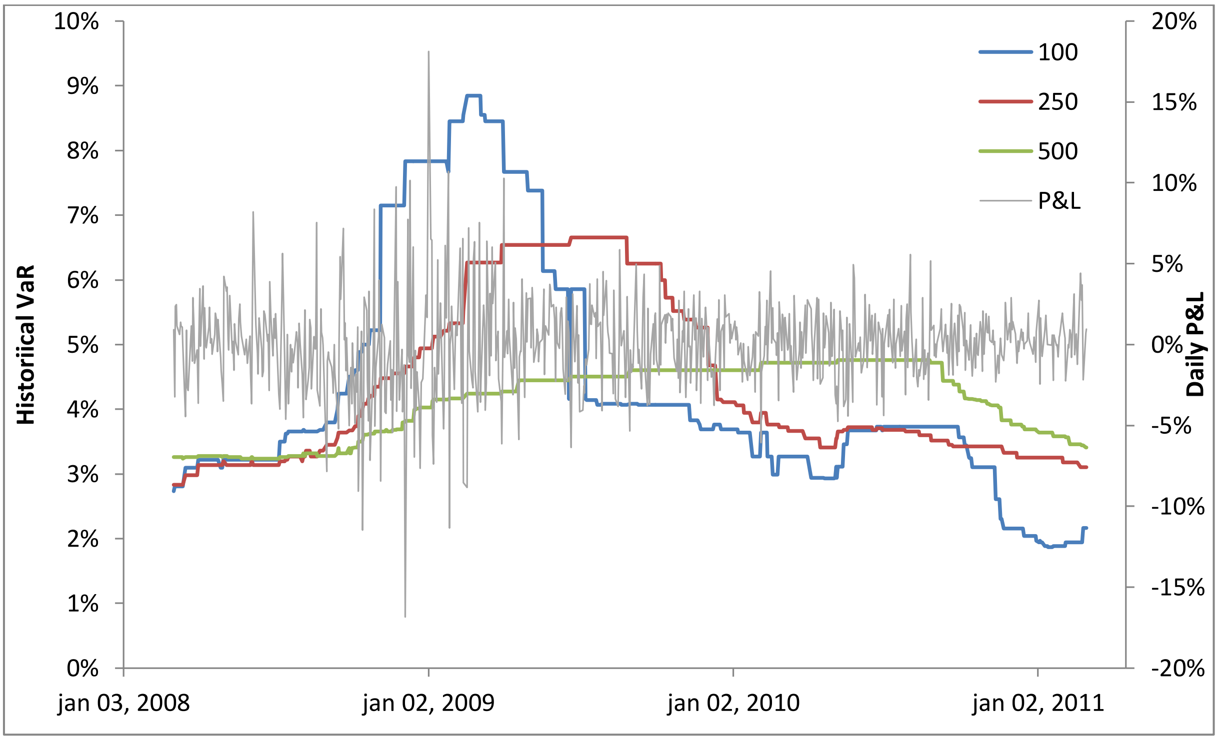

Figure 4.4: *Development of historical VaR the last three years using periods 100, 250 and 500. In grey are the percentile daily P&L for Brent crude shown.*

4.3 Backtesting

The preferred method of testing whether VaR is able to predict the real risk is to look back and see how well it has performed in the past. If the confidence level is 95%, one would expect that losses exceeding the VaR-number to occur roughly 5% of the days. If the more than 5% of the losses exceed the VaR-number, then the model has underestimated the risk. If, on the other hand, losses exceeding VaR occur less than 5%, the model is too conservative and overestimates the risk. The former causes the trader to keep too little capital in reserve to be able to cover losses, while the latter causes too much capital to be kept and thus may result in ineffective use of resources.

Backtesting is performed using a rolling windows approach. The estimation sample is rolled over the whole data period but the length of the window is kept constant, the first beginning at the earliest data in the time series of historical returns. Backtesting involves assuming that the loss distribution is a Bernoulli variable. A Bernoulli variable may only take the form of one of two values, where 'success' is labeled 1 and 'failure' is labeled 0. In this case, success means that the loss today has exceeded yesterday's VaR. An indicator function that displays this can be defined as:

$$I_{t+1,\alpha} = \begin{cases} 1, & when\ L_{t+1} > VaR_{t,\alpha} \\ 0, & when\ L_{t+1} \leq VaR_{t,\alpha} \end{cases}$$

If the VaR number is correct, then the probability of exceeding the VaR should be:

$$P\left(I_{t+1,\alpha} = 1\right) = 1 - \alpha$$

Further, the Bernoulli process could be extended to a binomial distribution. Defining *S* as the sum of all the violations of the VaR-number with a sample size *n:*

$$S = f(n, \alpha) = \sum_{t=1}^{n} I_{t+1,\alpha}$$

then the expected number and its standard deviation can be described as:

$$E(S) = n\alpha, \qquad SD(S) = \sqrt{n\alpha(1-\alpha)}$$

Assuming that the central limit theorem applies when *n* is a large number, the distribution for *S* is approximately normal. A *1-θ* confidence interval can then be constructed from these data:

$$(n\alpha - z_\theta\sqrt{n\alpha(1-\alpha} \, , n\alpha + z_\theta\sqrt{n\alpha(1-\alpha})$$

A hypothesis test is then performed with a null hypothesis H_0: $S = n\alpha$. This will be rejected if the number of violations falls outside the confidence interval. The normality assumption does not necessarily hold true, especially considering that a sufficiently large *n* not always will be available. Whether or not a hypothesis is rejected is also heavily dependent on the choice of significance level, which in itself is rather arbitrary. For this purpose a θ = 0,05 will be used.

Table 4.1 shows a backtest performed on a portfolio of 100% Brent crude, the VaR development of which was shown in Figures 4.1 and 4.2. As can be seen, performing backtests on VaR with a small VaR period is practically impossible. The number of exceedances is too small to gain any valuable information and whether or not the backtest succeeds is too dependent on single dates.

	VaR period (days)								
n (days)	**100**	**250**	**500**	**1000**	**2000**	**E(S)**	**SD(S)**	**Min**	**Max**
100	3	0	0	0	0	5	2,18	0,7	9,3
250	9	6	0	3	5	12,5	3,45	5,7	19,3
500	14	7	5	12	16	25	4,87	15,4	34,6
1000	43	46	49	57	56	50	6,89	36,5	63,5
2000	72	93	91	94	84	100	9,75	80,9	119,1
5000	203	250	245	263	246	250	15,41	219,8	280,2

Table 4.1: *Backtest for the n latest calculated historical VaRs at March 1st 2011. Different VaR periods have been used to illustrate how well it performs. To the right are shown the expected number of exceedance and the 95% confidence interval. Where the backtest fails has been marked red.*

The same is true for small sample sizes. At *n=100*, the number of exceedances is actually zero for periods over 250 days. The backtest seems to fail at sample sizes below 1000. To account for that, the confidence interval is relatively wider when compared to long backtests, but the normal distribution is not fully able to capture that. A fatter tailed distribution might be able to account better for that, but most likely small sampled backtesting ought to be avoided altogether.

While the Table 4.2 only shows the backtest for a single day, one can keep rolling the backtest itself. Figure 4.5 shows *S* as a function of time where the backtest window is the last 1000 trading days at a given time. The performance of this middle range backtests shows that the lower the VaR period, the less frequently *S* goes beyond its theoretical bounds. For the graph for a VaR period of 2000 days the confidence interval was actually exceeded constantly for more than four years in the period between October 1997 and April 2002. The performance of the 500 VaR backtest on the other hand was remarkably good, keeping well inside the confidence interval even during volatile periods.

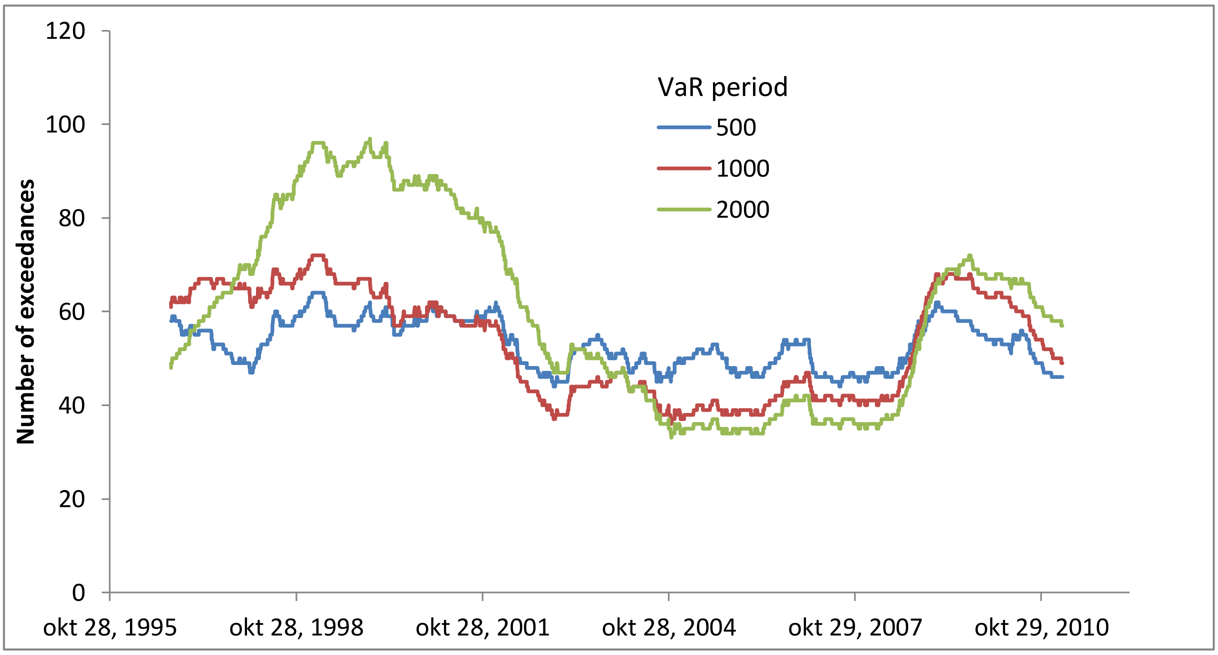

Figure 4.5: *Rolling backtest with a sample size of 1000 at 95% confidence level for the fluctuating prices of Brent crude. The minimum and maximum allowed by the confidence interval in Table 5.2 is S=36 and S=63*

5 Data analysis and results

The thesis uses data from the oil markets in performing historical simulation. Chapter 5.1 discusses the data in detail, explaining how a real position has been simplified to be easier to analyze. Chapter 5.2 introduces eight portfolio combinations to be checked for subadditivity. Single-day tests are explained in chapter 5.3 while their resulting figures are shown in appendix A. This has been stretched over a testing period in chapter 5.4, where a backtest has been constructed to see how often a particular portfolio displays subadditivity over time. Chapter 5.5 discusses the relationship between correlation and subadditivity, while chapter 5.6 offers some ways to reduce the amount of superadditivity through parametrical adjustments.

5.1 Empirical data

Real data from a financial portfolio at Statoil has been used as data for this analysis. As the various traders at Statoil engage in trading, the collective portfolio for the company as a whole becomes very complicated. It holds a wide array of different types of products, each with a different maturity date.

To analyze a complete position such as this is almost impossible, so some simplifications need to be done. Certain products tend to dominate trading. For Statoil, these are Brent and WTI crudes. In addition to these, a highly variable position in Middle Eastern heavy crudes seems to have a substantial impact on the portfolio. These are mainly Dubai oils, and come in a more rigid contract form like a forward or swap. The most important of the refined products is gasoline, which itself is mostly traded though futures contracts.

The length of these contracts can vary between a single day and up to four years in the future. Still, the trade is highly dominated by contracts with a maturity date which is three months or less from now.

On the request of Statoil, the actual values of the positions have not been included. The positions relative to the total position is shown in Table 5.1. The historical prices, on the other hand, are accurate. These have been collected in-house starting from April 1st 2003. The registered price data on a particular exchange is known as a ticker. The tickers included in this analysis are

A. IPE-BRT-FUT1
B. PLA-BRT-DTD1
C. NYM-WTI-FUT
D. PLA-DUB-SWP
E. NYM-UNL-FUT

The first abbreviation tells by which institution the prices have been collected. IPE is an old, but still used code for the Intercontinental Exchange (ICE) in London. NYM is the code that indicates the product to be traded on the New York Mercantile Exchange (NYMEX). Products that are traded over-the-counter do not have official data like this. Rather, a company known as Platts specializes in collecting data for OTC-products. PLA then indicates that data is from Platts.

The second code is the physical product itself. BRT = Brent crude, WTI = Western Texas Intermediate, DUB = Dubai, UNL = Unleaded gasoline.

The third code points to which financial instrument that is used. FUT means future, SWP means swap, while DTD indicates that the crude is *dated*. The last term only applies to Brent oil and it means that the shipment has received a date on which oil will be loaded onto a tanker. If the cargo does not have a designated loading date, it is known as a *paper barrel*. This is used mainly for hedging purposes or speculation. If the last code contains a 1 at the end, the product is traded on the European markets, whereas no number means the product is traded elsewhere.

The price data consists of contracts with a maturity date within three months. Price data for spot price, 1 month, 2 months and 3 months contracts are used. There is in fact much more detailed price data available, stating the contract length in weeks or even days, but in order to make things simple, only the monthly prices are presented. The approximation used is that contracts with a maturity date within month uses the spot price, ones with a maturity time between one and two months uses the one month futures price, and so on.

From now on the contracts will be abbreviated with codes stating products and maturity lengths. For example, the two month swap contract on product D, Dubai crude, will be called D.2. Unfortunately, for product C price data was only available for up two months, so there the three month future had to be omitted.

		Maturity month					
	Product	**March**	**April**	**May**	**June**	**Sum up to 3 months**	**Sum all months**
A	IPE-BRT-FUT1	2,61 %	14,351 %	6,916 %	2,906 %	26,783 %	33,405 %
B	PLA-BRT-DTD1	1,32 %	20,76 %	3,37 %	1,91 %	27,36 %	31,09 %
C	NYM-WTI-FUT	2,14 %	2,03 %	2,52 %	1,86 %	8,55 %	8,99 %
D	PLA-DUB-SWP	1,11 %	2,96 %	0,49 %	0,18 %	4,73 %	4,73 %
E	NYM-UNL-FUT	2,77 %	0,55 %	1,25 %	0,49 %	5,06 %	5,38 %
	Sum these products	9,94 %	40,65 %	14,54 %	7,35 %		
	Sum all products	13,05 %	44,24 %	17,80 %	8,47 %	**Total:**	**72,48 %**

Table 5.1*: The relative sizes of the positions at March 3th 2011. Simplifying to five products and maturity length of less than 3 months, one is able to capture 72 % of the total position. Short positions are in red, but are in this particular case calculated as positive numbers.*

Much of the price data is highly correlated, as could be expected. Figure 5.1 shows how the prices for different contract lengths move in very similar fashion. Both the short term and the long term development of the spot price are emulated closely by the futures prices. For this particular product, there seems to have been contango during most of the period.

The correlation matrix in Table 5.2 measures how similar the prices moves. The correlation factors between two products are calculated according to the formula defined in Chapter 3.2. The correlation between all of the crude product combinations goes below 0,98. Product E on the other hand is a commercial product and sold on a market that differs some from the crude market, as discussed in Chapter 2.3. It is therefore not surprising to observe a considerably lower correlation factor when comparing this product with the others. The correlation factor for these data seems approximately to lie in the interval [0,96 , 0,98].

Figure 5.2 reveals that the long term development is similar both for crude and refined products, but the short term fluctuations may be very different. The refined product seems to be more prone to bubbles and bursts and exhibits a higher degree of volatility clustering. The correlation is heavily dependent on the amount of data used to calculate it. From smaller samples the correlation would often be considerably less. The long-term correlation between crude oil and refined products was noted in a study by Asche et al. (2003), and was shown to be very high for most combinations of light and middle distillates, as well as naphta. The correlation between crude oil and heavy distillates was on the other hand found to be relatively low.

All of the products tested exhibit leptokurtosis compared to the normal distribution. Somewhat surprising was the how much this differed from product to product. As shown in Table 5.3, prices for Brent contracts seem to be the considerably less volatile than the other two crudes, having a kurtosis of roughly one third compared to that of WTI and Dubai crude. The standard deviation reflects this rather poorly, as it is only slightly lower. This leads to the conclusion that it would be a bad choice of risk measure, and thus that VaR based purely on an assumption on normality would be equally ineffective.

For the two Brent contracts and the gasoline, the kurtosis is higher for when the maturity length is short. This is most dramatically seen for gasoline, where the spot price exhibits a kurtosis of 8,5 the two month futures price only has a kurtosis of 2,6. This is probably due to the traders' knowledge of the price behavior, and the tendency for bubbles and bursts. For WTI and Dubai crudes this is reversed, and short term contracts seem to have the lowest kurtosis.

WTI and Dubai prices also have the highest skewedness. They are all skewed quite heavily toward the profit side compared with the symmetrical normal distribution, except for the two month WTI future where it is suddenly skewed clearly toward the loss side. Most the other data only show a low skew.

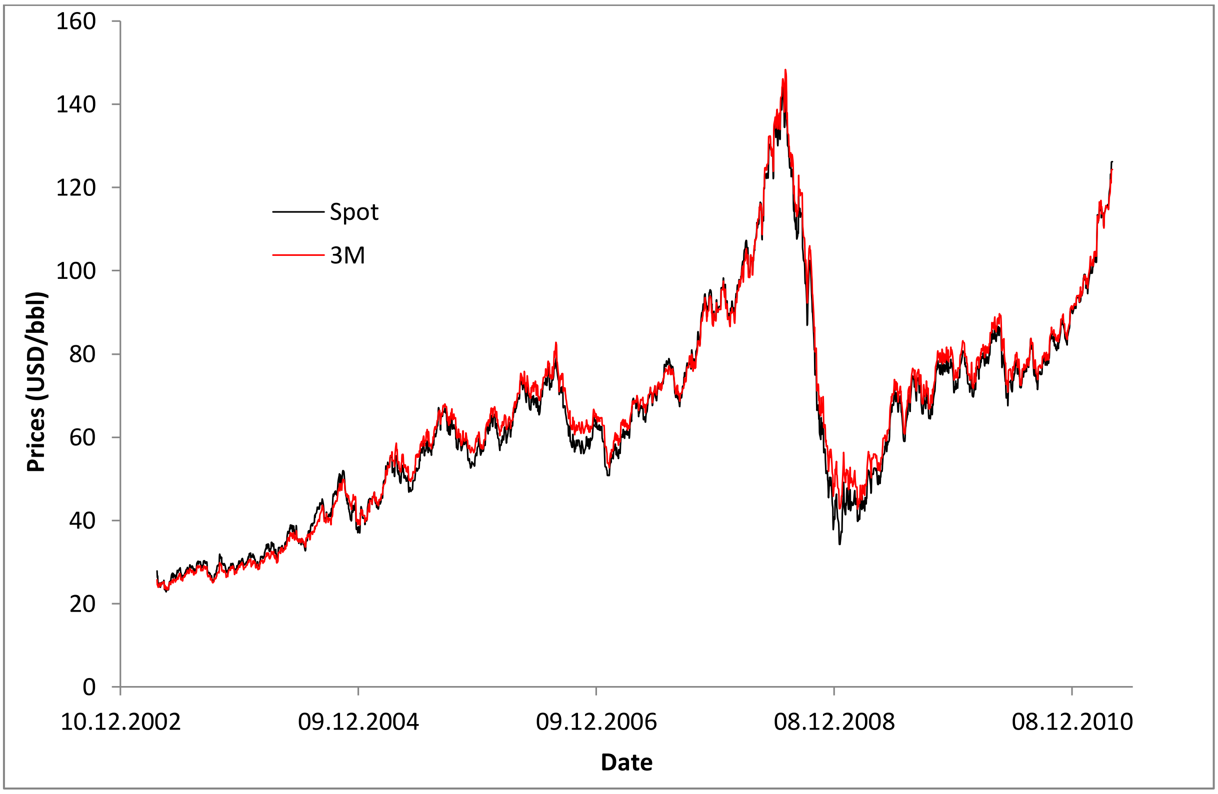

Figure 5.1*: Development of the prices of the product IPE-BRT-FUT1. Shown is the spot price along with futures contracts with length 3 month. Correlation factor = 0,9966*

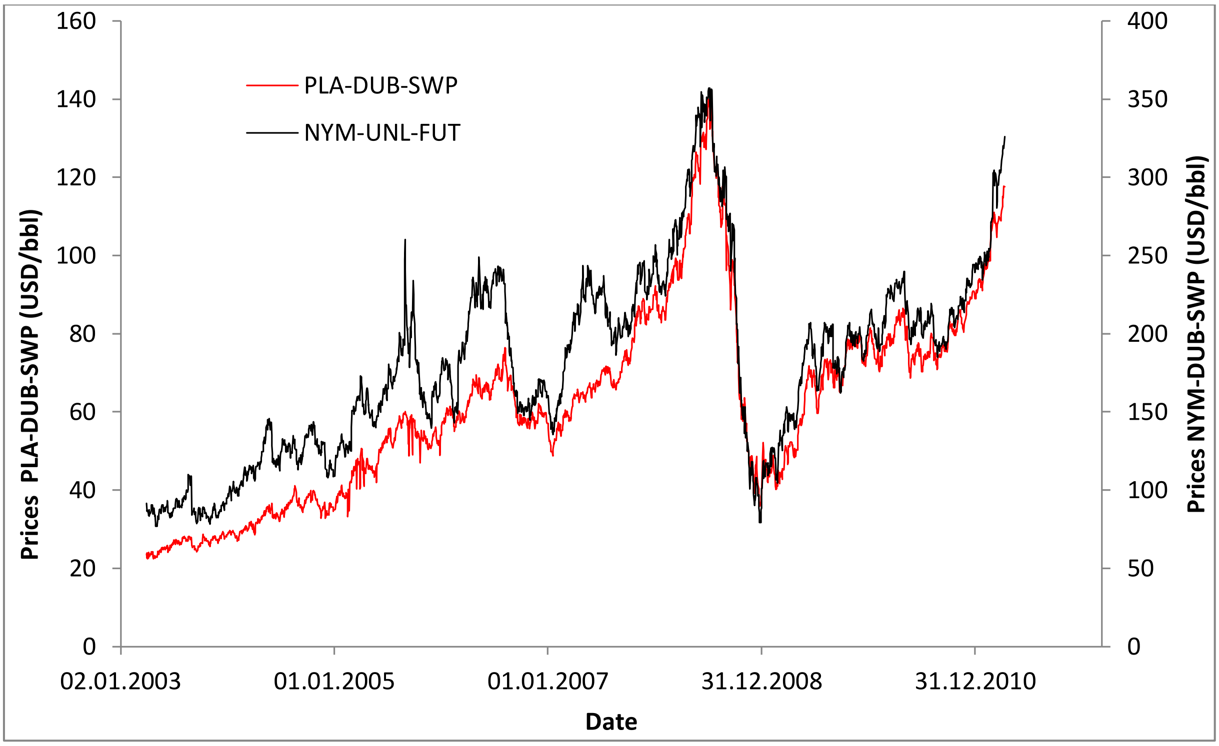

Figure 5.2: *Comparison between the spot prices for Dubai crude swaps and unleaded gasoline. Correlation factor = 0,9612*

		IPE-BRT-FUT1				PLA-BRT-DTD1				NYM-WTI-FUT			PLA-DUB-SWP				NYM-UNL-FUT			
		Spot	**1M**	**2M**	**3M**	**Spot**	**1M**	**2M**	**3M**	**Spot**	**1M**	**2M**	**Spot**	**1M**	**2M**	**3M**	**Spot**	**1M**	**2M**	**3M**
IPE-BRT-FUT1	**Spot**	1	0,9990	0,9979	0,9966	0,9999	0,9997	0,9995	0,9984	0,9927	0,9937	0,9940	0,9946	0,9942	0,9938	0,9930	0,9707	0,9729	0,9779	0,9816
	1M	0,9990	1	0,9997	0,9991	0,9987	0,9996	0,9998	0,9998	0,9913	0,9933	0,9959	0,9960	0,9957	0,9956	0,9951	0,9687	0,9710	0,9767	0,9813
	2M	0,9979	0,9997	1	0,9998	0,9975	0,9988	0,9992	0,9999	0,9897	0,9921	0,9961	0,9962	0,9959	0,9960	0,9959	0,9670	0,9693	0,9753	0,9803
	3M	0,9966	0,9991	0,9998	1	0,9961	0,9977	0,9982	0,9996	0,9879	0,9907	0,9958	0,9962	0,9959	0,9962	0,9963	0,9652	0,9675	0,9737	0,9791
PLA-BRT-DTD1	**Spot**	0,9999	0,9987	0,9975	0,9961	1	0,9996	0,9994	0,9980	0,9927	0,9936	0,9936	0,9941	0,9938	0,9933	0,9924	0,9711	0,9733	0,9782	0,9817
	1M	0,9997	0,9996	0,9988	0,9977	0,9996	1	1,0000	0,9992	0,9924	0,9939	0,9951	0,9952	0,9948	0,9945	0,9938	0,9702	0,9725	0,9778	0,9819
	2M	0,9995	0,9998	0,9992	0,9982	0,9994	1,0000	1	0,9995	0,9921	0,9938	0,9955	0,9955	0,9952	0,9949	0,9943	0,9698	0,9720	0,9775	0,9818
	3M	0,9984	0,9998	0,9999	0,9996	0,9980	0,9992	0,9995	1	0,9903	0,9925	0,9961	0,9964	0,9961	0,9961	0,9959	0,9678	0,9700	0,9759	0,9808
NYM-WTI-FUT	**Spot**	0,9927	0,9913	0,9897	0,9879	0,9927	0,9924	0,9921	0,9903	1	0,9997	0,9949	0,9856	0,9854	0,9848	0,9837	0,9641	0,9661	0,9697	0,9726
	1M	0,9937	0,9933	0,9921	0,9907	0,9936	0,9939	0,9938	0,9925	0,9997	1	0,9968	0,9878	0,9876	0,9872	0,9863	0,9647	0,9667	0,9709	0,9742
	2M	0,9940	0,9959	0,9961	0,9958	0,9936	0,9951	0,9955	0,9961	0,9949	0,9968	1	0,9917	0,9915	0,9916	0,9915	0,9641	0,9663	0,9719	0,9768
PLA-DUB-SWP	**Spot**	0,9946	0,9960	0,9962	0,9962	0,9941	0,9952	0,9955	0,9964	0,9856	0,9878	0,9917	1	0,9996	0,9996	0,9993	0,9612	0,9634	0,9689	0,9738
	1M	0,9942	0,9957	0,9959	0,9959	0,9938	0,9948	0,9952	0,9961	0,9854	0,9876	0,9915	0,9996	1	0,9999	0,9996	0,9615	0,9636	0,9691	0,9740
	2M	0,9938	0,9956	0,9960	0,9962	0,9933	0,9945	0,9949	0,9961	0,9848	0,9872	0,9916	0,9996	0,9999	1	0,9999	0,9609	0,9631	0,9688	0,9738
	3M	0,9930	0,9951	0,9959	0,9963	0,9924	0,9938	0,9943	0,9959	0,9837	0,9863	0,9915	0,9993	0,9996	0,9999	1	0,9598	0,9620	0,9679	0,9732
NYM-UNL-FUT	**Spot**	0,9707	0,9687	0,9670	0,9652	0,9711	0,9702	0,9698	0,9678	0,9641	0,9647	0,9641	0,9612	0,9615	0,9609	0,9598	1	0,9995	0,9956	0,9884
	1M	0,9729	0,9710	0,9693	0,9675	0,9733	0,9725	0,9720	0,9700	0,9661	0,9667	0,9663	0,9634	0,9636	0,9631	0,9620	0,9995	1	0,9973	0,9906
	2M	0,9779	0,9767	0,9753	0,9737	0,9782	0,9778	0,9775	0,9759	0,9697	0,9709	0,9719	0,9689	0,9691	0,9688	0,9679	0,9956	0,9973	1	0,9973
	3M	0,9816	0,9813	0,9803	0,9791	0,9817	0,9819	0,9818	0,9808	0,9726	0,9742	0,9768	0,9738	0,9740	0,9738	0,9732	0,9884	0,9906	0,9973	1

Table 5.2: *Correlation between the prices of the different products*

#	Product	Maturity length	Mean	Median	Min	Max	St. dev.	Skew	Kurtosis
1	IPE-BRT-FUT1	Spot	0,072 %	0,034 %	-12,533 %	11,851 %	2,232 %	-0,1286	2,9094
		1M	0,074 %	0,019 %	-11,223 %	12,305 %	2,076 %	-0,0125	2,7515
		2M	0,075 %	0,020 %	-10,600 %	11,307 %	2,018 %	-0,0214	2,5362
		3M	0,075 %	0,025 %	-10,138 %	10,514 %	1,972 %	-0,0292	2,3891
2	PLA-BRT-DTD1	Spot	0,072 %	0,010 %	-12,816 %	13,719 %	2,231 %	-0,0734	3,1300
		1M	0,073 %	0,022 %	-11,994 %	12,902 %	2,152 %	-0,0047	3,1857
		2M	0,073 %	0,028 %	-11,679 %	12,651 %	2,125 %	0,0106	3,1651
		3M	0,074 %	0,015 %	-10,805 %	11,451 %	2,028 %	-0,0064	2,8460
3	NYM-WTI-FUT	Spot	0,061 %	0,000 %	-21,155 %	24,077 %	2,604 %	0,23382	9,3879
		1M	0,062 %	0,000 %	-20,704 %	23,402 %	2,389 %	0,11799	9,2309
		2M	0,066 %	0,000 %	-24,793 %	21,541 %	2,357 %	-0,1201	14,384
		3M	-	-	-	-	-	-	-
4	PLA-DUB-SWP	Spot	0,077 %	0,000 %	-17,003 %	18,478 %	2,597 %	0,12897	10,232
		1M	0,076 %	0,000 %	-17,003 %	18,478 %	2,597 %	0,12990	10,236
		2M	0,077 %	0,000 %	-17,555 %	19,031 %	2,585 %	0,11641	11,344
		3M	0,078 %	0,000 %	-17,863 %	19,693 %	2,587 %	0,10567	11,880
5	NYM-UNL-FUT	Spot	0,061 %	0,000 %	-24,979 %	25,013 %	2,754 %	-0,0248	8,5219
		1M	0,061 %	0,000 %	-16,072 %	19,913 %	2,589 %	-0,0070	4,0361
		2M	0,061 %	0,000 %	-11,227 %	13,441 %	2,380 %	-0,0398	2,5840
		3M	0,062 %	0,000 %	-11,009 %	11,591 %	2,252 %	-0,0537	2,5565

Table 5.3: *Analysis of the historical daily profits and losses for the different contracts. The entire sample is used (N=2092).*

5.2 Portfolio selection

Putting together portfolios from the assets presented in Table 5.3 is an action that presents a number of challenges. An infinite number of different portfolios could theoretically be created and thus an infinite number of results would be possible to acquire.

In my work I have attempted numerous portfolio compositions and most of them results in very chaotic and highly sensitive results. Two opposing goals soon presented themselves. On one hand, it is preferable to keep things as simple as possible. A portfolio with only a few different contracts would have a low number of variables. It should then be easier to track which of these are impacting the portfolio as a whole.

On the other hand, simple portfolios are perhaps not too relevant for large corporations with as complex a position as Statoil's. Often complexity can cause novel effects to occur, that do not apply when only a few variables are taken into account. The argument that subadditivity violations only apply on the level of individual traders, but are neutralized at the managerial level when the traders are part of a greater organization, is an equally important aspect.

Test #	Portfolio A	Portfolio B
1	*$1000 of A.0*	*$1000 of A.1*
2	*$1000 of D.0*	*$1000 of D.2*
3	*$1000 of B.0*	*$1000 of E.1*
4	*$1000 each of D.0 + D.3 + E.2*	*$1000 each of C.0 + D.2 + E.2*
5	*$1000 of A.0.*	*-$1000 of A.1*
6	*-$1000 of A.0*	*-$1000 of A.1*

***Table 5.4**: Six simple portfolio combinations used in the tests for subadditivity. Negative numbers mean short positions*

In the end the test results presented here are performed on six positions that should be considered simple, and two that could be considered complex. In creating the complex portfolios, the actual position of Statoil at March 3th 2011 has been a guiding principle, but is not actually reproduced. Tables 5.4 and 5.5 presents the positions used in checking for subadditivity.

The first test combines two very correlated portfolios containing a single asset each (corr.=0,999). Both of these assets are common and are traded with the most flexible instrument, the futures contract. This should be representative as the most standardized of all portfolio combinations.

Test 2 is very similar to the first test, but the product it holds is less common and the contract type is less flexible. The historical prices for these contracts also display a high amount of kurtosis, compared with those in the first test. Test 3 combines two portfolios of very different products, one is a crude oil and the other is a refined product. The correlation here is noticeably less. Test 4 creates a middle ground between a simple and a complex portfolio. Test 5 makes a long and short position neutralize each other. Test 6 combines two short positions.

The portfolios in tests 7 and 8 are based loosely on the relative sizes of the real position presented in Table 5.1. Test 7 places all the Brent contracts in portfolio A while all the contracts for other products are placed in portfolio B. The last test places all the long positions in one portfolio and all the short ones in the other.

		IPE-BRT-FUT1				PLA-BRT-DTD1				NYM-WTI-FUT			PLA-DUB-SWP				NYM-UNL-FUT			
#		**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
7	**A**	30	160	80	30	10	230	40	20											
	B									20	20	30	10	30	5	0	30	5	10	5
8	**A**				30		230		20			30	10	30	5	0				
	B	30	160	80		10		40		20	20						30	5	10	5

***Table 5.5**: The complex portfolio combinations used in the tests for subadditivity. Black marks long positions and red marks short positions.*

5.3 Single-day subadditivity testing

The criterion for subadditivity can be written as: $r(A) + r(B) - r(A + B) \geq 0$. If the risk measure is value-at-risk, ideally, the sum of the historical VaRs for the individual portfolios minus the VaR for the combined portfolio should always be larger than zero. When this criterion is not satisfied, subadditivity is violated. This should be true for all confidence levels and all time horizons. Table 5.6 shows this criterion employed in practice at a single day on the positions described in the previous chapter. The tests are performed at confidence levels 95% and 99% and with historical samples going back 250, 500 and 1000 days, respectively. As seen, most of the time the subadditivity criterion is satisfied, but in a small number of cases the VaR the combined portfolio is larger than the sum of the individual VaRs. In this particular case, the superadditivity is present at 3 out of 48 instances.

Because of the nature of historical simulation, the tail region of the loss distribution will always be coarse and discontinuous. Each day there will almost certainly be some parts of the tail where subadditivity is not satisfied. Figure 5.3 illustrates this by combining two simple portfolios. Portfolio A contains $1000 of Product A with maturity date less than one month (using the spot price), and portfolio B contains $1000 of the same product but uses the one month futures price. The graphs display the difference between the sum of the individual VaRs and that for the combined portfolio. At some confidence levels this value is below zero. Note that the results are not dependent on the absolute sizes of the positions, only the weight ratio between them.

There does not appear to be any clear pattern as to at which confidence levels subadditivity is most likely to be violated. They are just as likely to occur at lower levels like $\alpha = 95\%$ as higher levels like $\alpha = 99\%$. This is unexpected as the distribution tends to get more disjointed the farther out in the tail regions the observations are made, and one would have expected that to cause additional violations.

A larger sample size makes the fluctuations in the function be reduced, as should be expected, but the amount of subadditivity violations seems to be about the same.

More tests have been performed on other simple portfolios. The results of these are shown in Figures A.1-A.7 in the appendix. Here subadditivity violations occur in at least some confidence levels for Tests 2, 4, 6 and 7, while the others are subadditive at all confidence levels. For Tests 3 and 8, the resulting additivity plot creates something that resembles continuous graphs that decrease as the confidence level is lowered.

#		HistVaR at α=95%				HistVaR at α=99%			
		A	B	A+B	Diff	A	B	A+B	Diff
1	**250**	25,96	25,41	51,10	0,27	41,16	41,50	82,28	0,38
	500	31,21	29,93	60,52	0,62	47,68	44,24	84,52	7,40
	1000	37,98	36,30	74,00	0,29	74,87	67,69	140,09	2,48
2	**250**	23,99	24,27	37,23	11,03	39,66	40,53	64,38	15,81
	500	31,29	31,77	41,01	22,05	42,30	41,56	67,18	16,68
	1000	36,51	36,75	67,33	5,94	68,11	66,16	134,27	0,00
3	**250**	26,14	28,90	43,00	12,04	39,89	46,67	69,28	17,28
	500	29,82	31,23	46,86	14,19	43,77	48,51	74,14	18,14
	1000	38,61	43,48	67,99	14,10	73,72	81,32	118,38	36,66
4	**250**	64,13	62,85	126,34	0,64	91,88	100,52	192,40	0,00
	500	65,48	67,41	129,54	3,35	106,86	111,63	215,88	2,61
	1000	91,76	99,78	195,69	-4,15	157,46	173,48	326,50	4,45
5	**250**	25,96	26,99	5,59	47,36	41,16	39,94	18,31	62,79
	500	31,21	30,71	5,91	56,02	47,68	44,30	21,86	70,13
	1000	37,98	33,92	8,95	62,95	74,87	69,04	22,97	120,94
6	**250**	27,02	26,99	52,12	1,88	43,90	39,94	83,84	0,00
	500	31,61	30,71	59,94	2,39	44,59	44,30	87,81	1,08
	1000	35,66	33,92	67,47	2,11	76,61	69,04	148,18	-2,53
7	**250**	15,03	3,36	18,66	-0,27	23,96	5,35	28,15	1,15
	500	18,14	3,58	21,67	0,05	26,10	5,94	29,52	2,51
	1000	22,36	5,18	26,64	0,91	40,73	9,00	48,34	1,38
8	**250**	8,27	10,17	2,10	16,34	13,10	16,18	4,38	24,90
	500	9,85	11,13	2,37	18,60	13,94	16,43	4,03	26,35
	1000	12,69	12,92	3,01	22,61	23,25	27,06	5,40	44,91

Table 5.6: *Overview of the individual subadditivity tests on the date March 3th 2011. The column labeled "Diff" subtracts the A+B column from the sum of the A and B columns. Subadditivity is violated where this gives a negative number.*

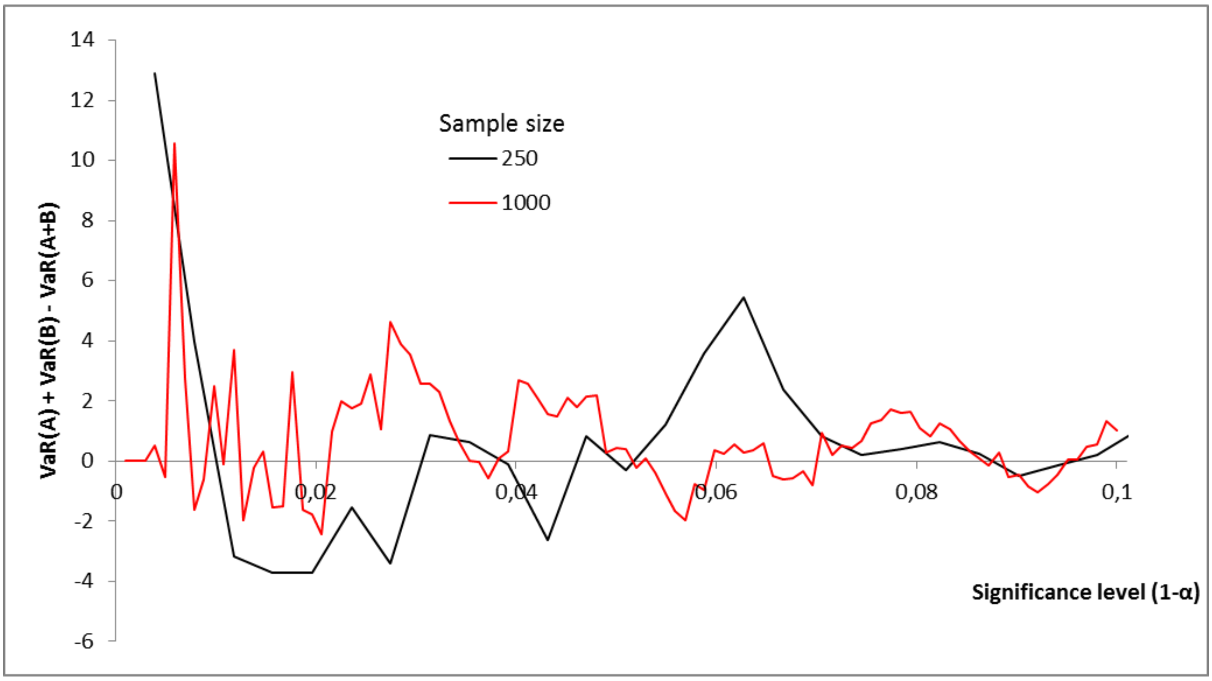

Figure 5.3*: Single-day test: Portfolio A: $1000 of A.0.   Portfolio B: $1000 of A.1*

5.4 Subadditivity backtesting

Backtesting akin to that discussed in Chapter 4.5 is also possible when considering subadditivity violations. An indicator function must be adopted with a Bernoulli variable such that it returns 1 when subadditivity is violated and 0 when it is satisfied.

$$I_{t+1,\alpha} = \begin{cases} 1, & VaR(A) + VaR(B) - VaR(A+B) < 0 \\ 0, & VaR(A) + VaR(B) - VaR(A+B) \geq 0 \end{cases}$$

This backtest will have to be used at set confidence levels. The ones chosen for this purpose are α = 95% and α = 99%. The sum of these variables can then be divided with the total number of days to acquire a percentage of how often the subadditivity principle is violated.

Table 5.7 shows a summary of the backtests performed on the portfolios mentioned in Chapter 5.2. Short backtesting periods were shown to be very inaccurate for these kinds of historical prices, as seen in Chapter 4.3. Also, price data stretching back 2000 days or more was not available for all products. Therefore, 1000 days was chosen as the sole backtesting period for this purpose.

In this summary are shown the amount of time that subadditivity is violated. This would translate to how often the graphs in Figure 5.4 and 5.5 are below zero. As seen, violations of the subadditivity axiom can be quite substantial for many of the portfolios. There are multiple cases where value-at-risk is in a superadditive state over half the time.

#	P&L corr.	α	Sample size 250	500	1000
1	0,952	0,95	33,2 %	14,5 %	28,4 %
		0,99	1,3 %	48,5 %	4,4 %
2	0,885	0,95	18,5 %	21,6 %	24,1 %
		0,99	10,2 %	51,4 %	47,5 %
3	0,354	0,95	0,0 %	0,0 %	0,0 %
		0,99	0,0 %	0,0 %	0,0 %
4	0,933	0,95	17,8 %	33,8 %	62,9 %
		0,99	33,5 %	52,5 %	1,5 %
5	-0,952	0,95	0,0 %	0,0 %	0,0 %
		0,99	0,0 %	0,0 %	0,0 %
6	0,952	0,95	15,7 %	19,3 %	0,0 %
		0,99	11,5 %	52,3 %	57,6 %
7	0,805	0,95	41,8 %	33,4 %	3,2 %
		0,99	3,7 %	34,9 %	2,9 %
8	-0,973	0,95	0,0 %	0,0 %	0,0 %
		0,99	0,0 %	0,0 %	0,0 %

Table 5.7: *Subadditivity backtest with a testing period of 1000 days. The correlation coefficient on the historical P&L is also shown for comparison.*

No general tendency can be traced to explain how either the choice of confidence level or the choice of sample size is impacting the amount of superadditivity. That subadditivity violations seem to be independent of confidence level is very surprising. Given that it is generally thought to be caused by coarseness in the tail-region of historical return distributions, one would expect it to be more prevalent at higher confidence levels. The limited amount of tests performed could be a likely reason for this. This would mean that there is in fact a relationship between the two, but it is dwarfed by the large randomness factor of the tests.

Figure 5.4 below shows the development of the first subadditivity test over time at 95% and Figure 5.5 does so at 99%. Similar graphs for the other tests are shown in Figures B.1-B.14 in the appendix, using the same portfolios as in the previous chapter. The graphs sketched with sample size 500 often falls somewhere between those using 250 and 1000, respectively. To keep the figures reader friendly, those data are not shown. This is not always true, however. As seen in Figure 5.5, the 99%-VaR is kept stable after the most volatile periods of financial crisis in 2008. While it has stabilized in subadditive state for sample sizes 250 and 1000, for 500 it has stabilized in a marginally non-subadditive state. This has a major impact on the results shown in Table 5.7.

Tests 3, 5 and 8 are the only ones which have no violations at all. As seen in their respective figures in Appendix B, subadditivity is clearly satisfied during the entire period. This is true for all tried confidence levels and sample sizes. Interestingly, it is only for these tests the correlation between the two portfolios is low. The correlation for Test 5 is negative, and VaR is even more subadditive here than for Test 3. That combining two portfolios of with a low correlation coefficient reduces risk, and ideally also the VaR number, is widely known and was discussed in Chapter 3.1. But from the tests here, a low correlation also seems to increase the amount of subadditivity as well, all the way to practically eliminating the problem at negative correlation.

Given the nature of the historical simulation, first of all the rolling window approach, one of the properties of this kind of backtesting is that violations of subadditivity tend to cluster together. Rather than being single day occurrences, subadditivity often stays violated for quite some time. Larger sample size makes the function dependent on data further in the past, while new data has relatively less of an impact. Thus, subadditivity clustering becomes more prominent with longer samples. In the example in Figure 5.4, for sample size = 1000, value-at-risk is continuously non-subadditive from September 7th 2007 to August 11th 2008. After September 4th 2008, subadditivity is not violated again until the end of the period.

This kind of backtesting suffers from a limitation compared to the standard backtesting technique discussed in Chapter 4.5. For the VaR-number itself, a certain amount of violations are expected. In fact, the whole principle of value-at-risk is that it is expected to be violated with a certain probability 1-α, and therefore a hypothesis test can be constructed from this to check its performance. This is not true for subadditivity testing. Here, violations are not expected at all and can potentially induce additional risks in itself. The ideal number of violations is therefore zero, but there is no marker over whether this portfolio ‘succeeds’ or ‘fails’ the subadditivity test. The only indication is that the number should be as low as possible.

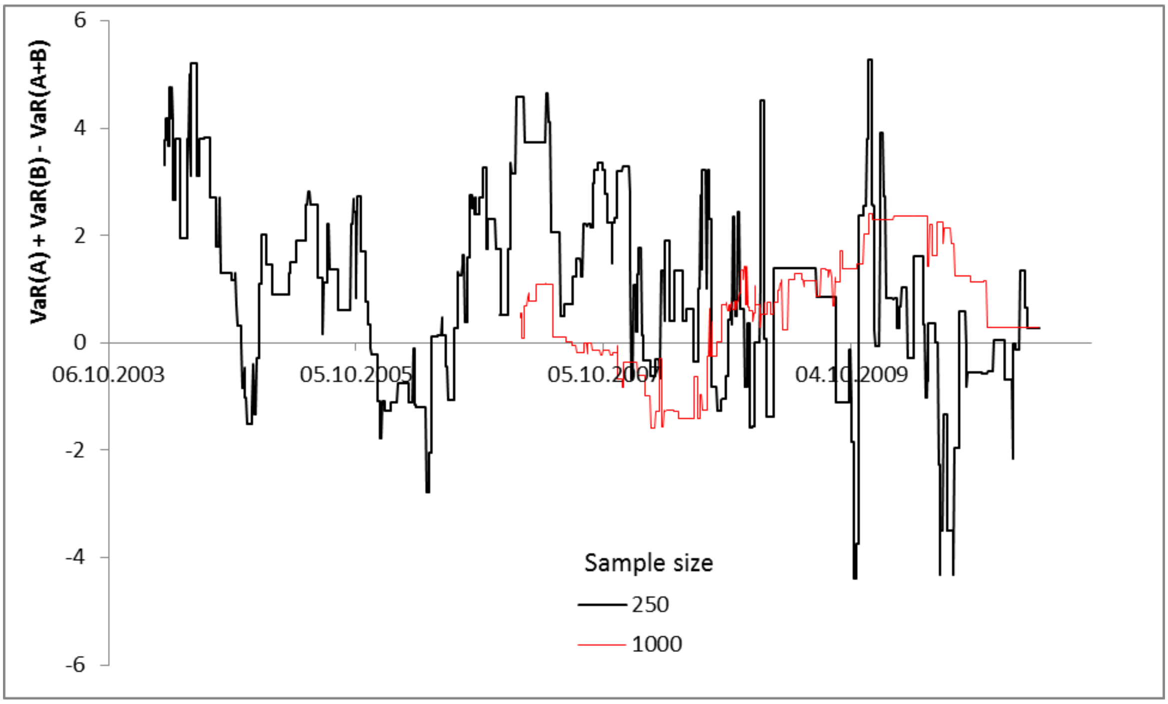

Figure 5.4: *Subadditivity backtest at α = 95% with a portfolio A: \$1000 of A.0 and a portfolio B: \$1000 of A.1*

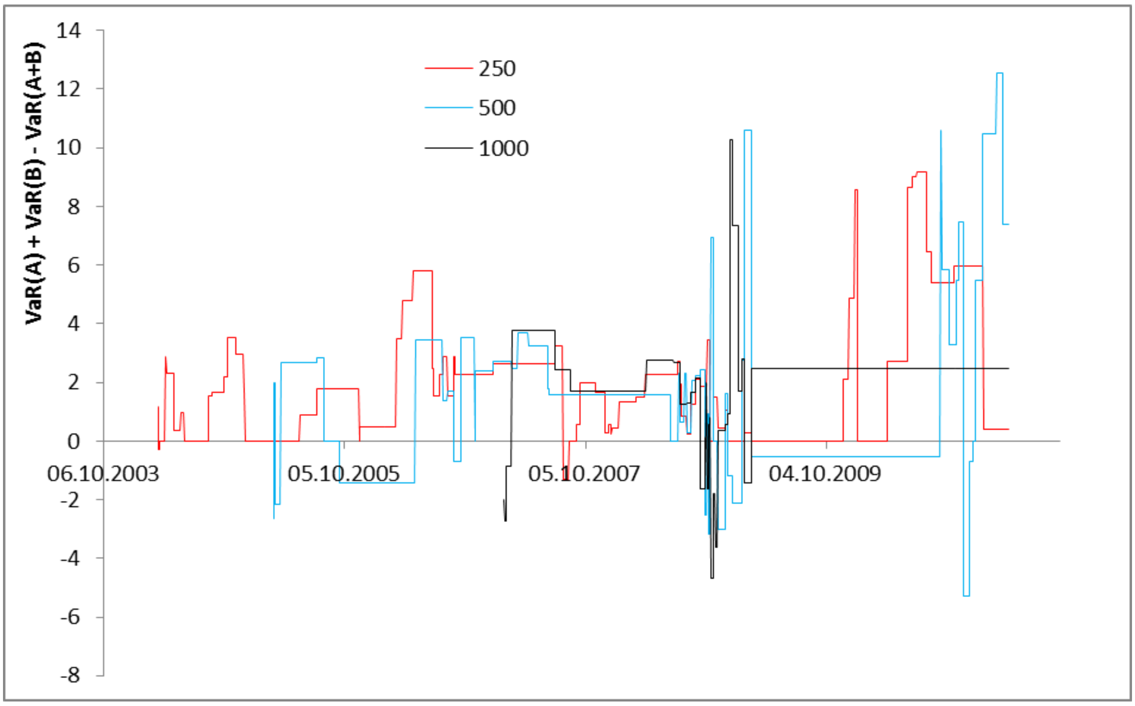

Figure 5.4: *Subadditivity backtest at α = 99% with a portfolio A: \$1000 of A.0 and a portfolio B: \$1000 of A.1*

5.5 Follow-up testing

The previous chapters discussed portfolios which were supposed to be representative for certain types of portfolios. It is also possible to try to create portfolios which should generate certain results.

Since low correlation seems to increase the likelihood of VaR being subadditive, an interesting point would be to see how low the correlation coefficient could be reduced, and still have some degree of superadditivity. Through the scientifically unsound method of just trying out a large amount of different portfolios more or less at random, some instances of superadditivity were found where the correlation coefficient was fairly close zero.

The lowest of these was for quite a simple combination where the first portfolio contained $1000 worth of spot prices in Brent crude. The second portfolio was contained a long position $1000 one-month Brent futures which was hedged by $1000 short position of three-month Dated Brent futures. This calculated to a correlation between the portfolios of 0,039. Figure 5.5 shows the development of the subadditivity test. As can be seen, the 250 day-period VaR at 95% confidence level dips just below zero during the autumn of 2008, incidentally the volatile period caused by the financial crash. In this situation, subadditivity is violated 1,0% of the total time period.

Pushing this even more to the extreme, Figure 5.6 shows the backtest of the same portfolios but where the position of Dated Brent has been increased to 1009$. Correlation in this case has dropped just below zero (at -0,002), while a tiny fraction of superadditivity still remains on one single day.

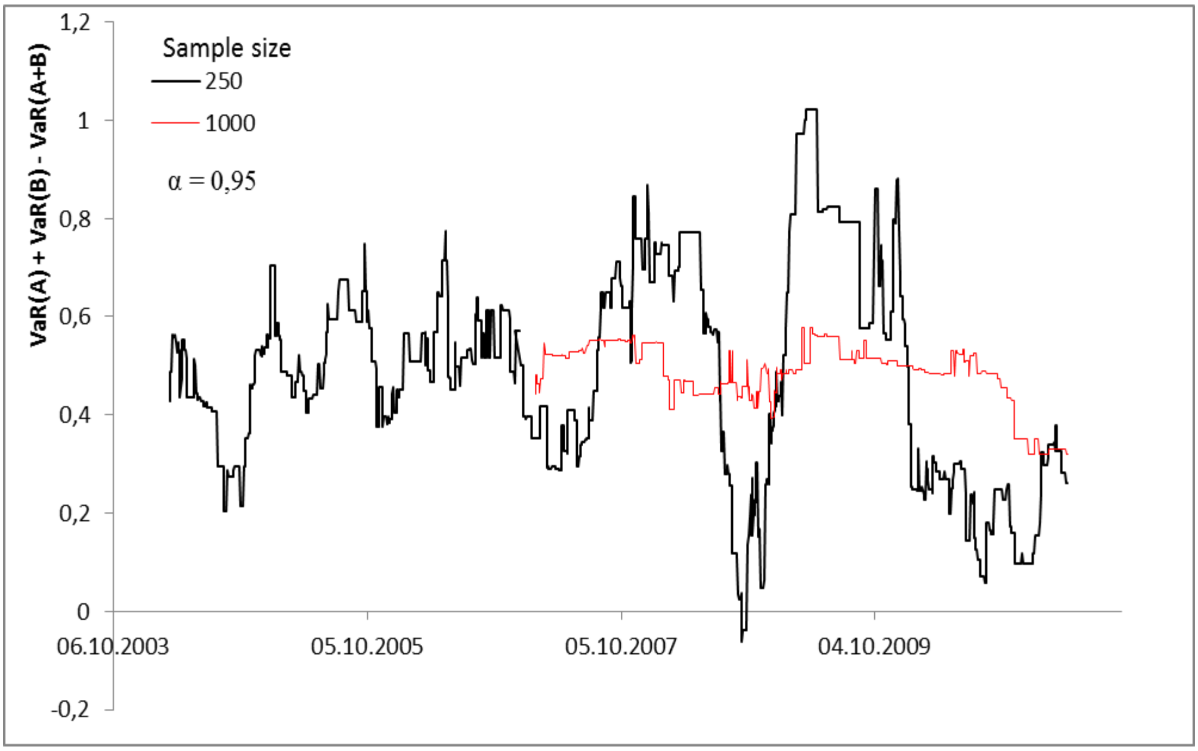

Figure 5.5: *Portfolio A: 1000$ A.0, Portfolio B: 1000$ of A.1 and 1000$ of B.3. Correlation 0,039. Violations at t=250 and α=95% accounts for 1,0%.*

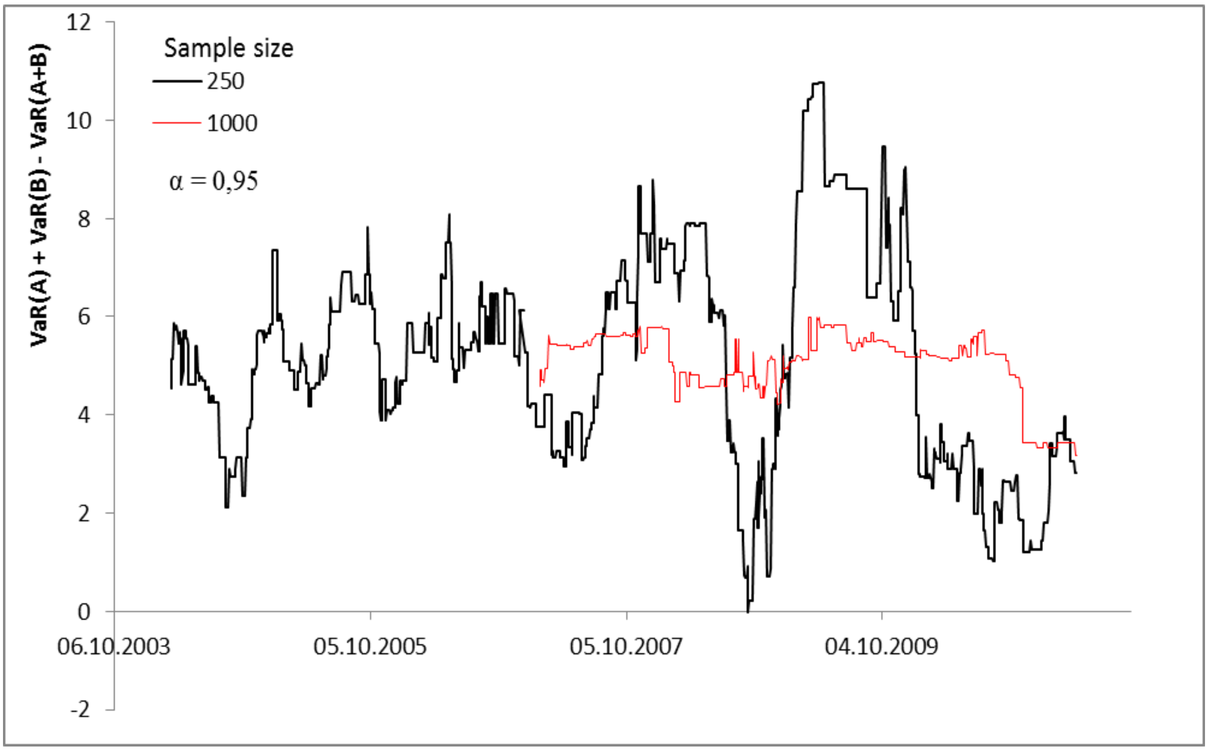

Figure 5.6: *Portfolio A: 1000$ of A.0, Portfolio B: 1000$ of A.1 and 1000$ of B.3. Correlation = -0,002. Violations at t=250 and α=95% accounts for 0,1%*

The methodology is too crude to conclude anything concrete, but it seems like a strange coincidence that it proves so difficult to create violations of subadditivity at negative correlation. When it is finally accomplished, it is only just slightly and it could be explained away with roughness in the modeling. There may be a logical reason for this, that it is in fact theoretically impossible to have both at same time.

5.6 Possible solutions

The method discussed in Chapters 5.1 and 5.2 uses equal weights for the entire data sample. While this is the simplest way of simulating historical returns, it assigns importance to the data quite unrealistically. For a sample size T=250, the return on all 250 trading days ago is given equal importance, whereas the return on the 251st trading day is not taken into account at all.

A possible solution to this problem is to assign descending weights that depend on how far in the past the data was recorded. The hybrid approach developed by Boudoukh, Richardson and Whitelaw (1999), often referred to as the BRW-model, can assign a historical weighting scheme according to a set decay factor. If λ is the decay factor, then the weight on day t can be calculated as:

$$w_t = \lambda^{t-1}(\lambda - 1)$$

The choice of decay factor can be a challenge, and has a significant impact on the VaR-estimate. It should be a number between *0* and *1*, and the most common value decay factor is no doubt $\lambda = 0{,}97$, which is the one used by the RiskMetrics standard. As t increases the weight will then approach 0, and the sum of the weights will approach 1. A large enough sample size is needed for the sum to be sufficiently close to 1. For $\lambda = 0{,}97$, a sample period of one year gives the weight sum of 0,9995 which ought to suffice. A higher decay factor will require a significantly longer history of data for this to be fulfilled, as can be noted from Table 5.8. A decay factor equal to 1 translates to no decay at all, and the standard historical is thus a special case of the BRW-model.

Sum	λ =0,97	0,98	0,99	0,995
0,90	76	114	230	460
0,95	99	149	299	598
0,99	152	228	459	919
0,995	174	264	528	1057
0,999	227	342	688	1378

Table 5.8: *Sample sizes needed in order to achieve certain weight sums at various decay factors*

Using longer historical periods also creates a problem in itself which is may be even more difficult to solve. The market regime may undergo changes during the historical period. Commodity future prices often moves in bubbles and bursts, undergo seasonal fluctuations and rapidly switching between contango and backwardation. Alexander (2008) suggests that portfolio returns could be artificially adjusted so that the volatility is approximately constant over the entire period. Volatility clustering gets removed from historical returns and imposed a constant volatility on the series, one that is equal to the conditional volatility of the series at the time VaR is estimated. Two methods are suggested in order to implement this:

1. Reconstruction of portfolio returns
2. Adjustment of individual risk factors

Alternative 1 is simpler because it only requires one volatility adjustment at the top level. The second alternative would fundamentally be more robust, but with a cost of increased complexity in the calculations.

Historical simulations, with adjustment using parametric hybrid methods such as these, will make the distribution in the tail region smoother. As was shown by Daníelsson (2010), a substantial reduction in subadditivity violations should be expected.

6 Conclusion

The non-subadditivity discussed by Artzner et al. (1999) is not only a logical issue that only arises in certain extreme cases, but is actually discovered to be quite common. This is at least true for simple positions of oil futures prices using unweighted historical simulation.

Contrary to what one would expect, the degree of subadditivity violations seems to be independent of both the choice of sample size and confidence level. It has, on the other hand, a distinct relationship with the correlation between the portfolio returns. A negative correlation causes a superadditive VaR to be virtually impossible, and even a low positive number makes it extremely unlikely. Thus, hedging a long position with a highly correlated short position is an effective way not only in reducing the actual risk, but also in eliminating superadditivity.

The degree of subadditivity in historical VaR simulation often displays clustering. Long periods of subadditivity can be offset by long periods of superadditivity. Since the historical VaR has a tendency to stabilize at a certain value after volatile periods, whether this happens in a subadditive or non-subadditive state can have a major impact on the quality of the risk measure. For two highly correlated portfolios, it is not unlikely that the combined VaR could be subadditive less than half the time. No instances were found where subadditivity was constantly violated all the time.

Increased complexity in the structure of the portfolios makes subadditivity violations occur less often. Possibly, this is only due to the low probability of creating two highly correlated complex portfolios, but there may also be some benefit in complexity in itself. This paper has mainly investigated quite simple positions, but a similar study by Dahl (2011) concludes that when the number of positions is larger than 10, superadditivity becomes increasingly unlikely even at high correlation.

7 Bibliography

Alexander, C. (2008). *Market Risk Analysis* (Vol. IV). John Wiley & Sons Ltd.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1997). Thinking Coherently. *Risk Magazine*.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical finance, 9*, 203-228.

Asche, F., Gjølberg, O., & Völker, T. (2003). Price Relationships in the Petroleum Market: An Analysis of Crude Oil and Refined Product Prices. *Energy Economics 25*, 289-301.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (1998, May). The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk. *Risk Magazine*(11), 64-77.

Dahl, R. (2011). On the coherency of value at risk using historical simulation.

Dahl, R. E. (2008). Value at risk analysis with Monte Carlo simulation.

Daníelsson, J., Jorgensen, B. N., Samorodnidsky, G., Sarma, M., & de Vries, C. G. (2010). Fat tails, VaR and subadditivity.

Dowd, K. (1998). *Beyond Value-at-Risk.* John Wiley & Sons.

Edwards, D. W. (2010). *Energy trading and investing.* McGraw-Hill.

Garcia, R., Renault, É., & Tsafack, G. (2005). Proper conditioning for coherent VaR in portfolio management. *Management Science, 53*(3), 483-494.

Giot, P., & Laurent, S. (2003). Market risk in commodity markets: A VaR approach. *Energy economics*(25), 435-457.

Hall, K. G., & Rankin, R. A. (2011, May). *Speculation explains more about oil prices than anything else.* Hentet 2011 fra http://www.mcclatchydc.com/2011/05/13/114190/speculation-explains-more-about.html

Hamilton, J. D. (2008). Understanding crude oil prices.

Holton, G. (2008). *The case for incoherence*. Hentet February 2011 fra http://glynholton.com/2008/09/the-case-for-incoherence/

Hull, J. C. (2006). *Risk management and financial institutions.* Prentice Hall.

Ibragimov, R. (2005). New Majorization Theory In Economics And Martingale Convergence Results In Econometrics.

J.P. Morgan. (1996). RiskMetrics - technical document.

Jorion, P. (2001). *Value at risk: the new benchmark for managing financial risk.* McGraw-Hill.

Mandelbrot, B. B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*(36), 394-419.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments.* John Wiley & Sons.

McNeal, A. J., Frey, R., & Embrechts, P. (2005). Quantitative Risk Management - Concepts, Techniques and Tools. Princeton University Press.

Pritsker, M. (2001). The hidden dangers of historical simulation.

Riley, G. (2006, September). *Market for Oil*. Hentet February 2011 fra tutor4u: http://tutor2u.net/economics/revision-notes/as-markets-oil.html

Schofield, N. C. (2007). *Commodity derivatives - Markets and applications.* John Wiley & Sons.

Žiković, S. (2008). Friends and Foes: A Story of Value at Risk and Expected Tail Loss.

Žiković, S., & Aktan, B. (2010). Extreme Movements and Measuring Risk in WTI Oil Prices .

8 Appendices

Appendix A – Single-day tests

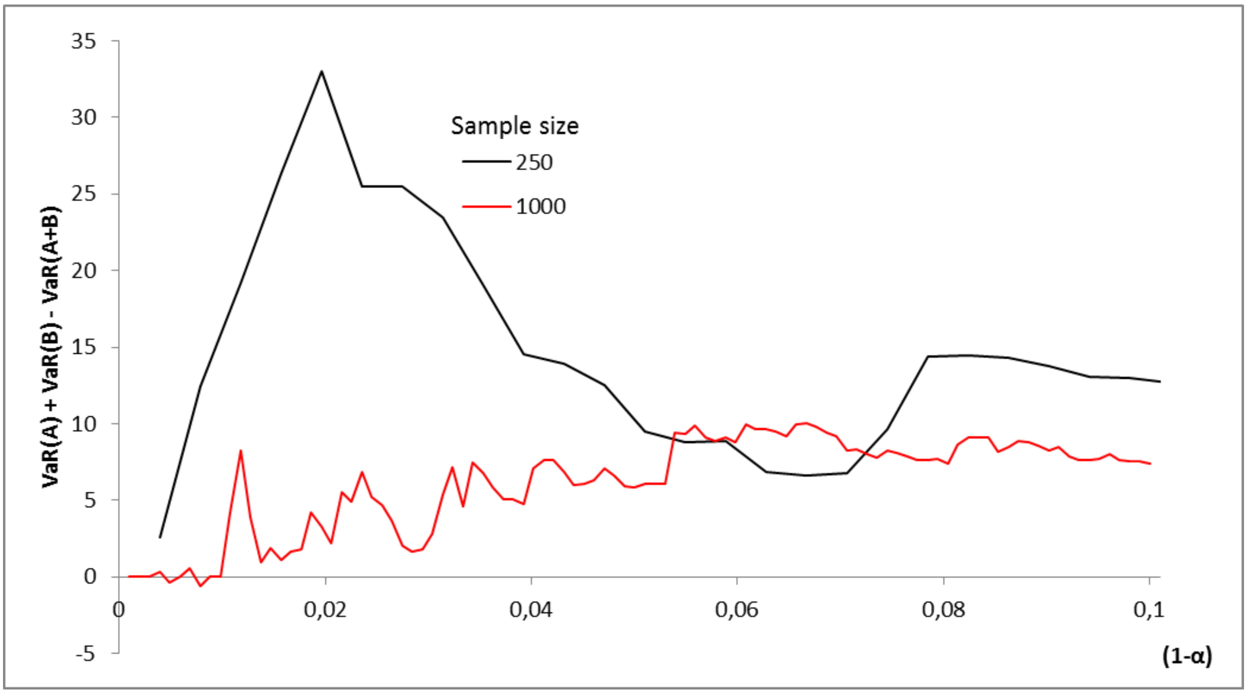

Figure A.1: *Test 2 - Portfolio A: $1000 of D.0. Portfolio B: $1000 of D.2.*

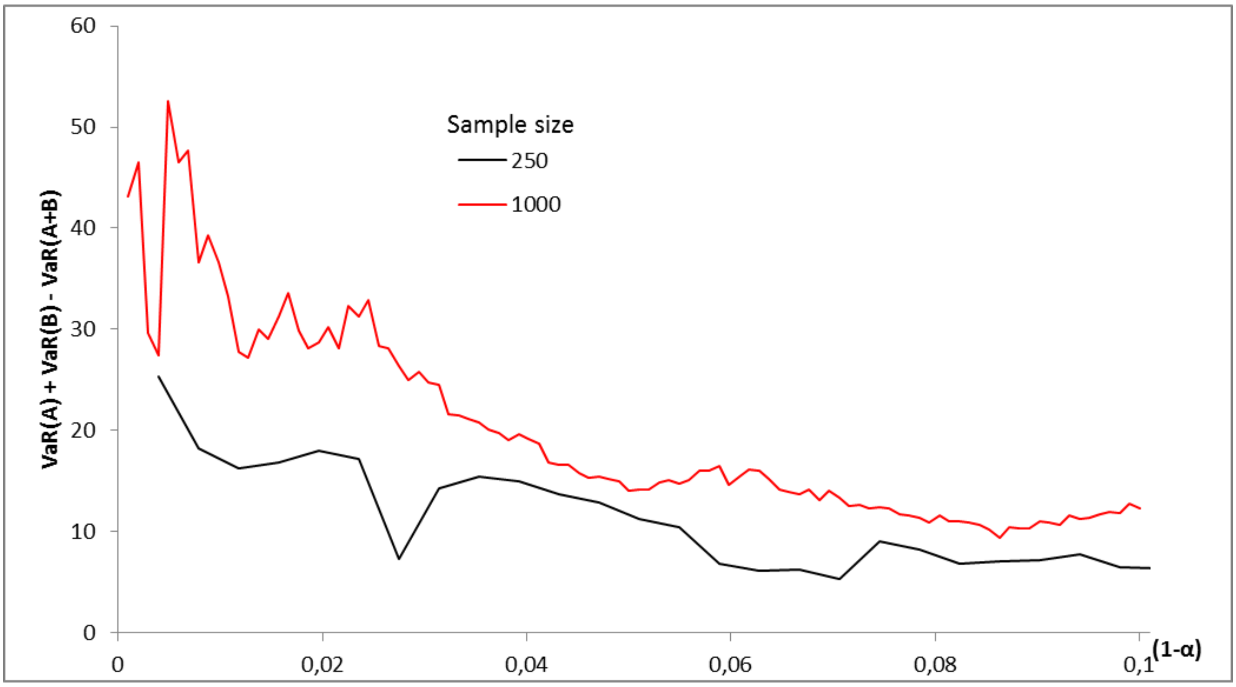

Figure A.2: *Test 3 - Portfolio A: $1000 of B.0. Portfolio B: $1000 of E.2*

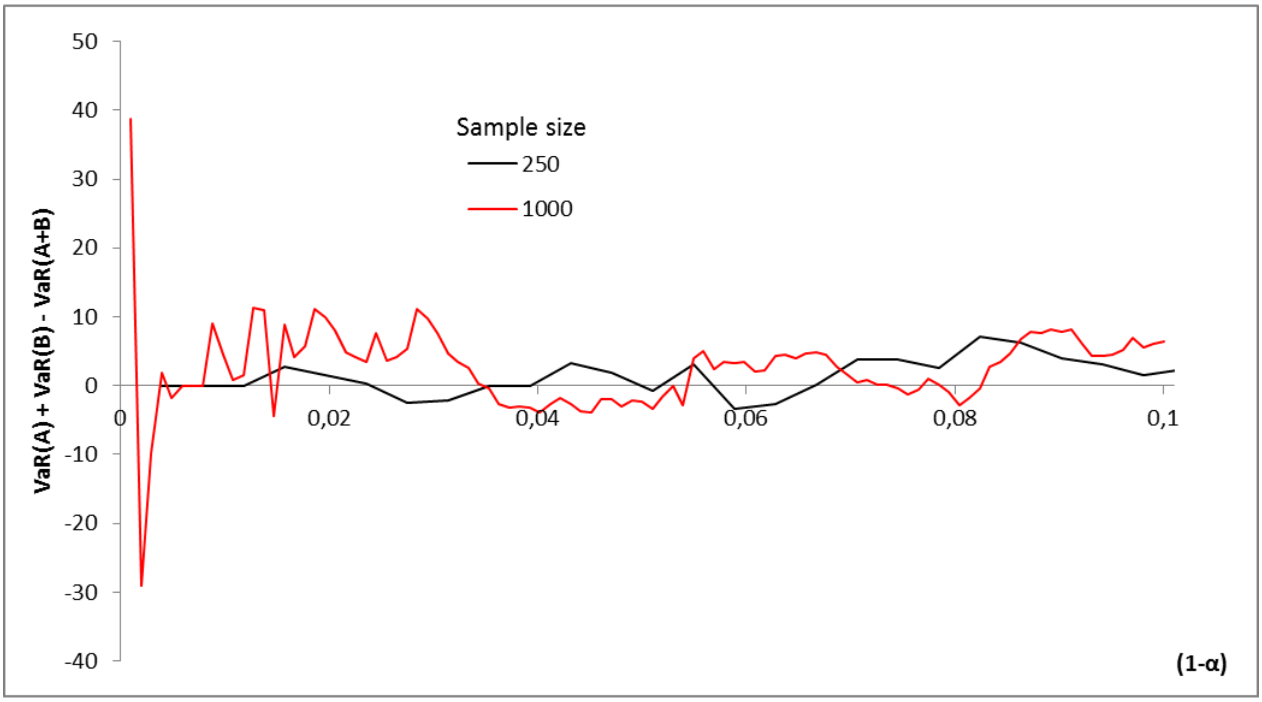

Figure A.3: *Test 4 - Portfolio A: $1000 each of D.0 + D.3 + E.2. Portfolio B: $1000 each of C.0 + D.2 + E.2.*

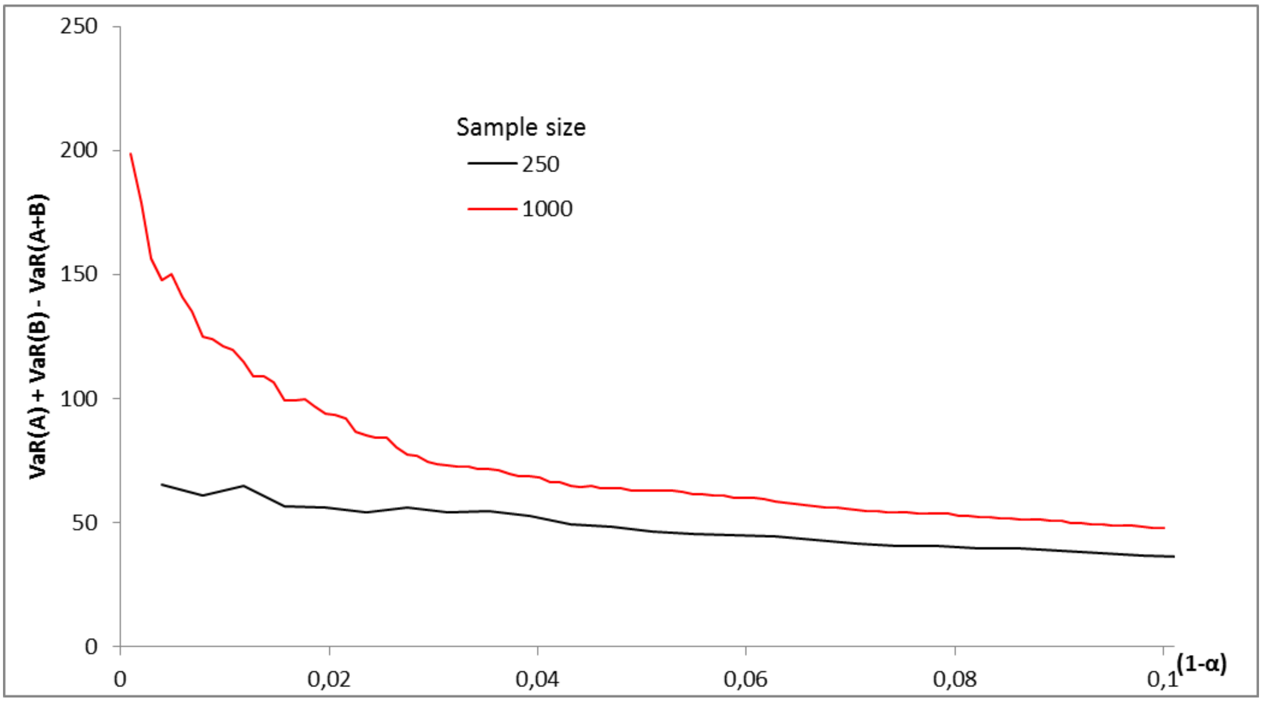

Figure A.4: *Test 5 - Portfolio A: $1000 long position of A.0. Portfolio B: $1000 short position of A.1*

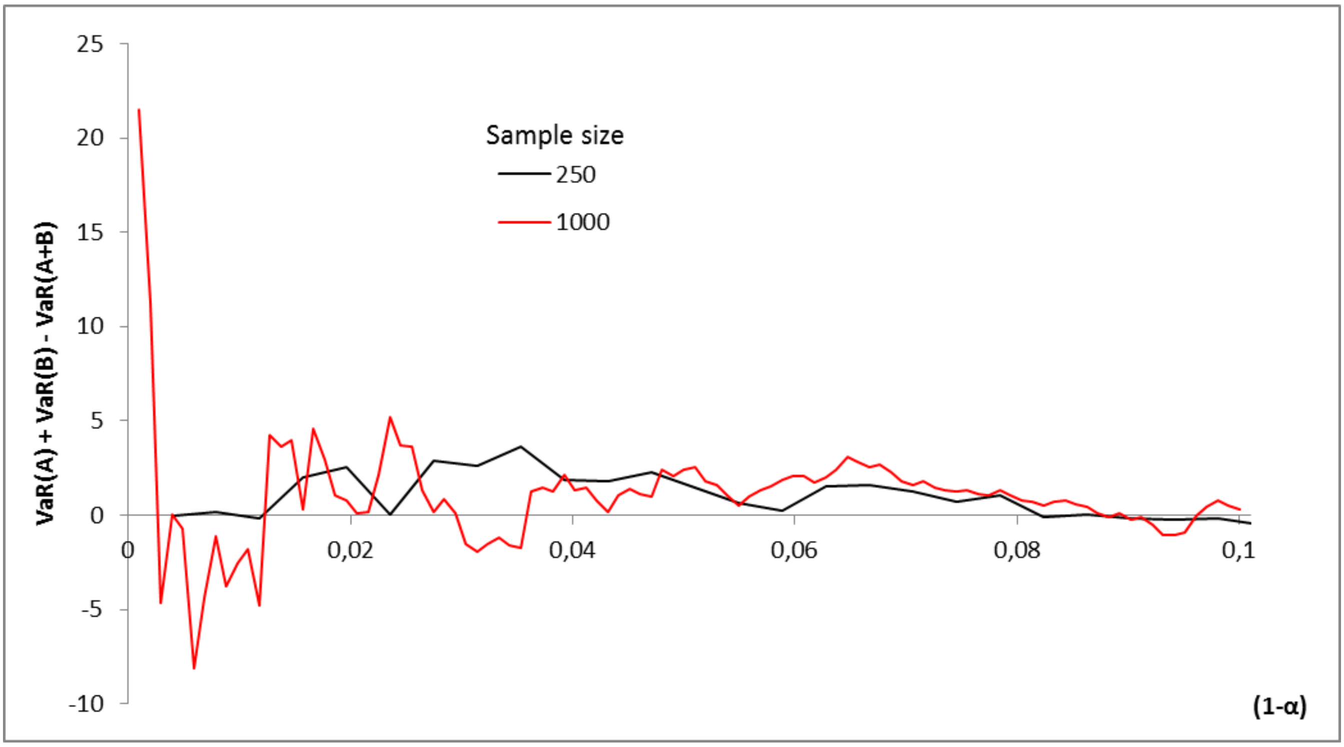

Figure A.5: *Test 6 - Portfolio A: $1000 short position of A.0. Portfolio B: $1000 short position of A.1*

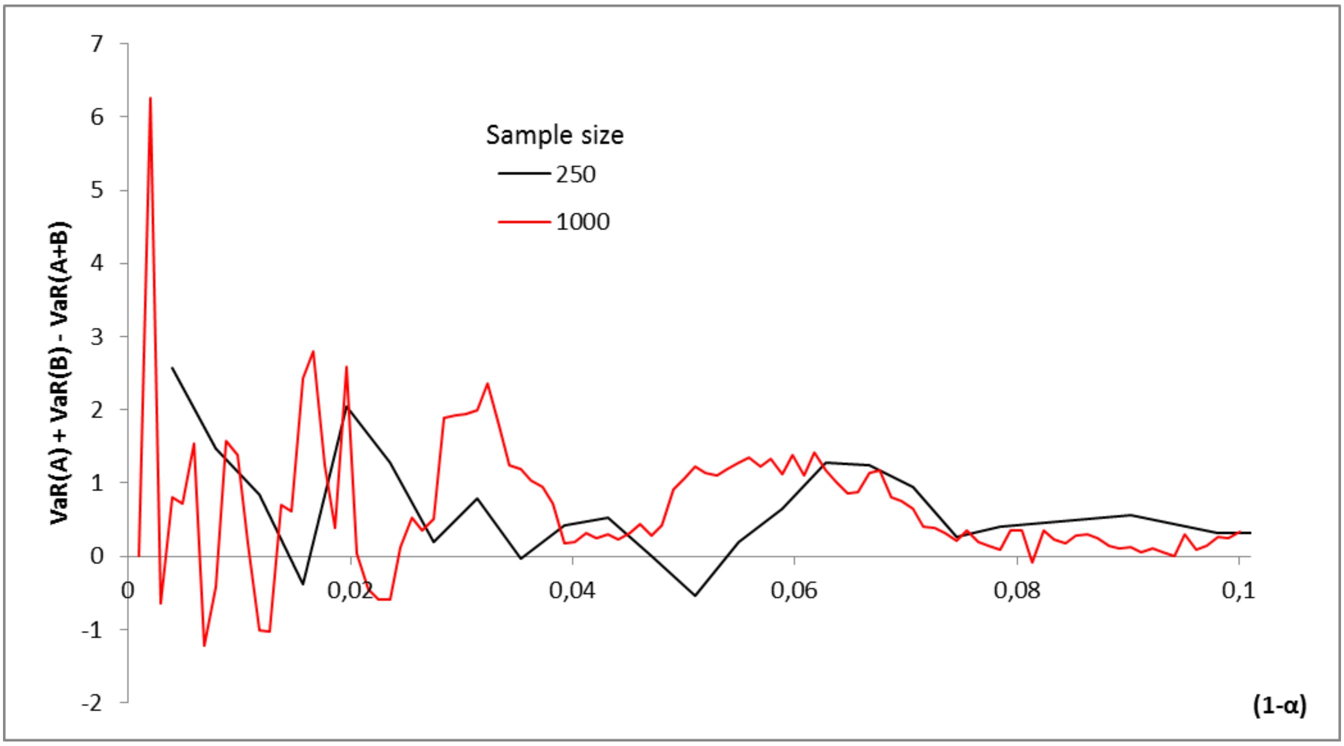

Figure A.6: Test 7: *Portfolios as described in Table 5.5*

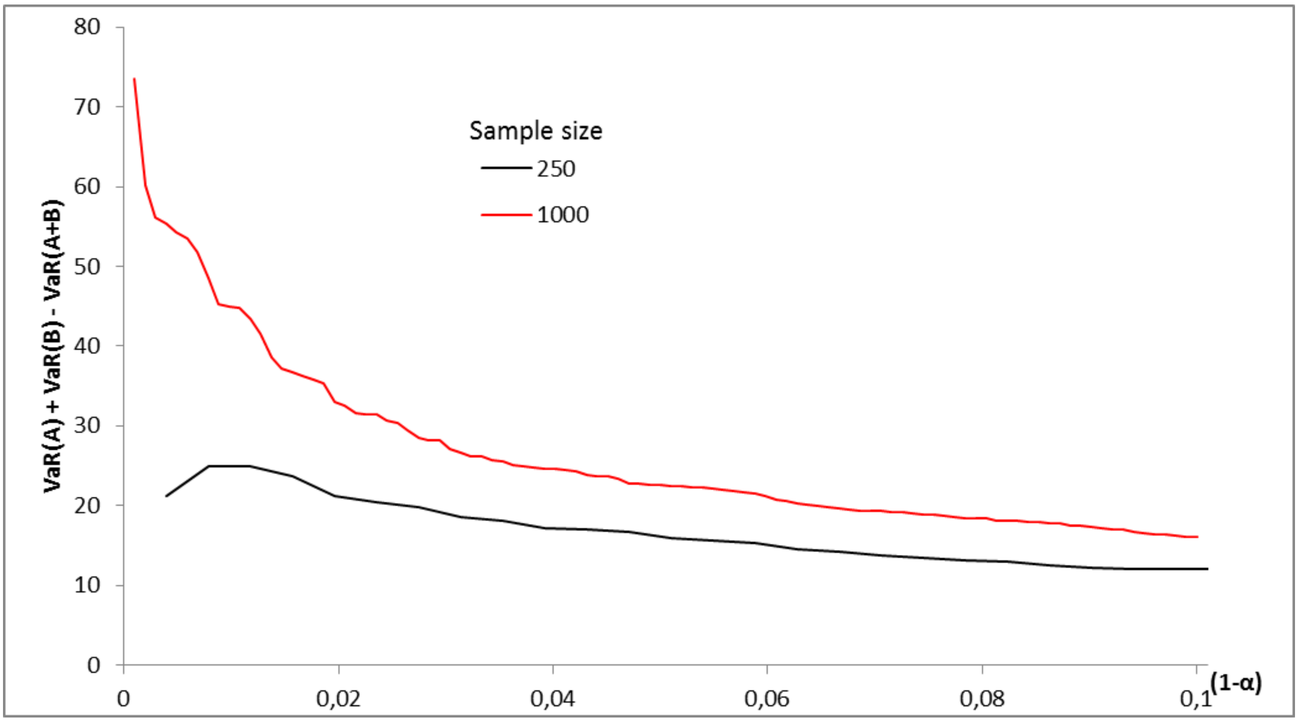

Figure A.7: Test 8 - *Portfolios as described in Table 5.5*

Appendix B - Backtests

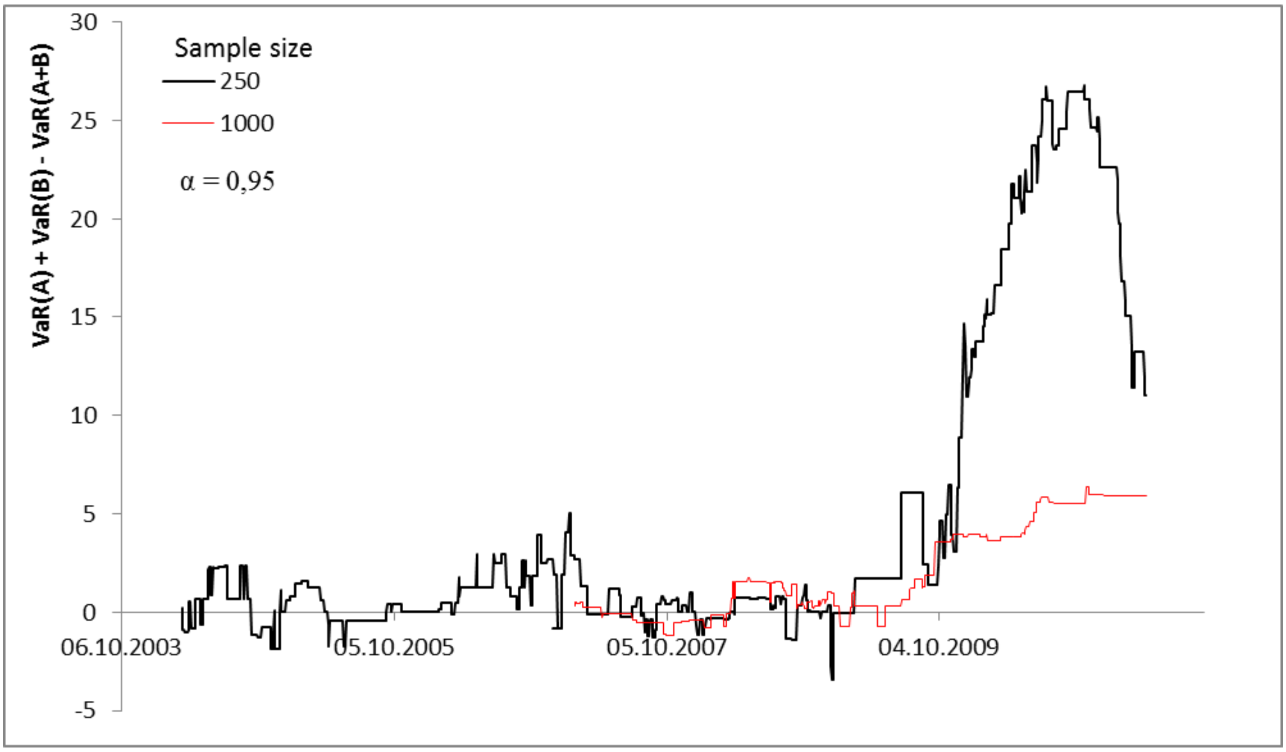

Figure B.1: *Test 2 - Portfolio A: \$1000 of D.0. Portfolio B: \$1000 of D.2.*

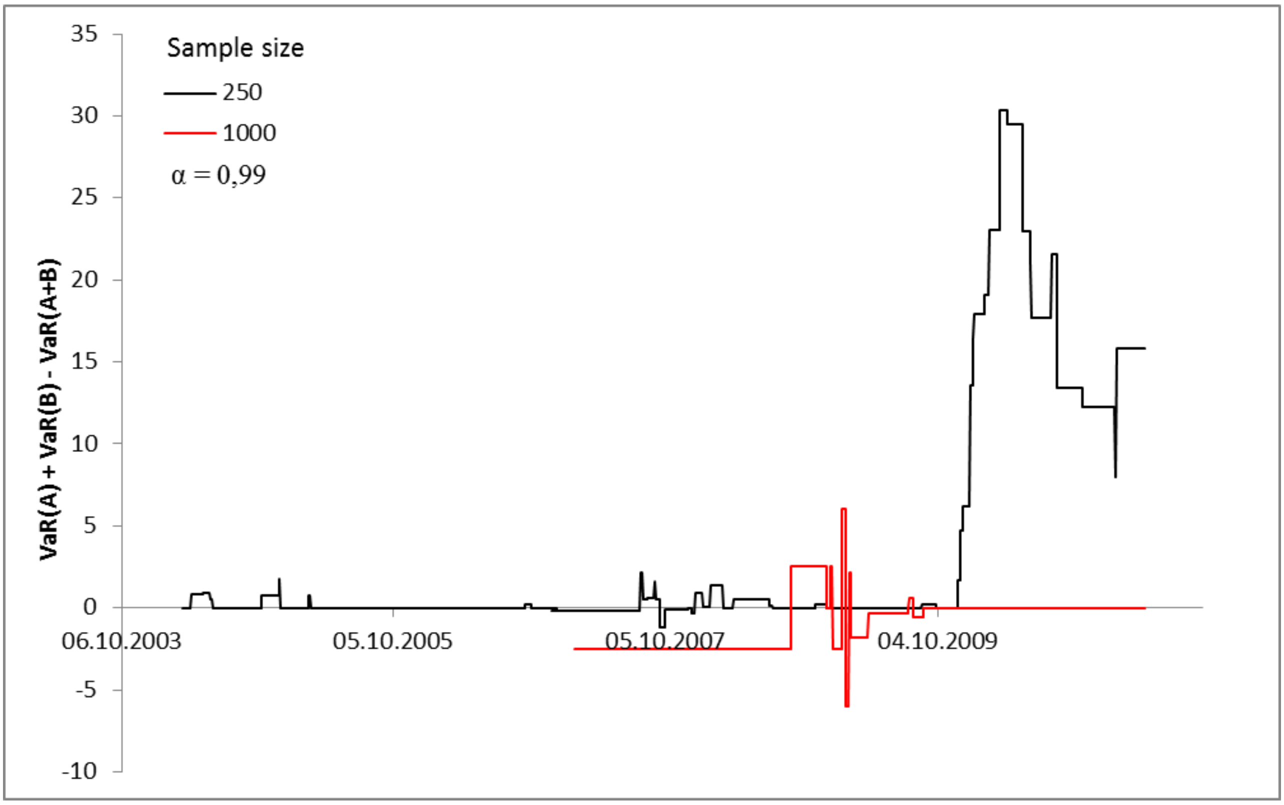

Figure B.2: Test 2 - *Portfolio A: \$1000 of D.0. Portfolio B: \$1000 of D.2.*

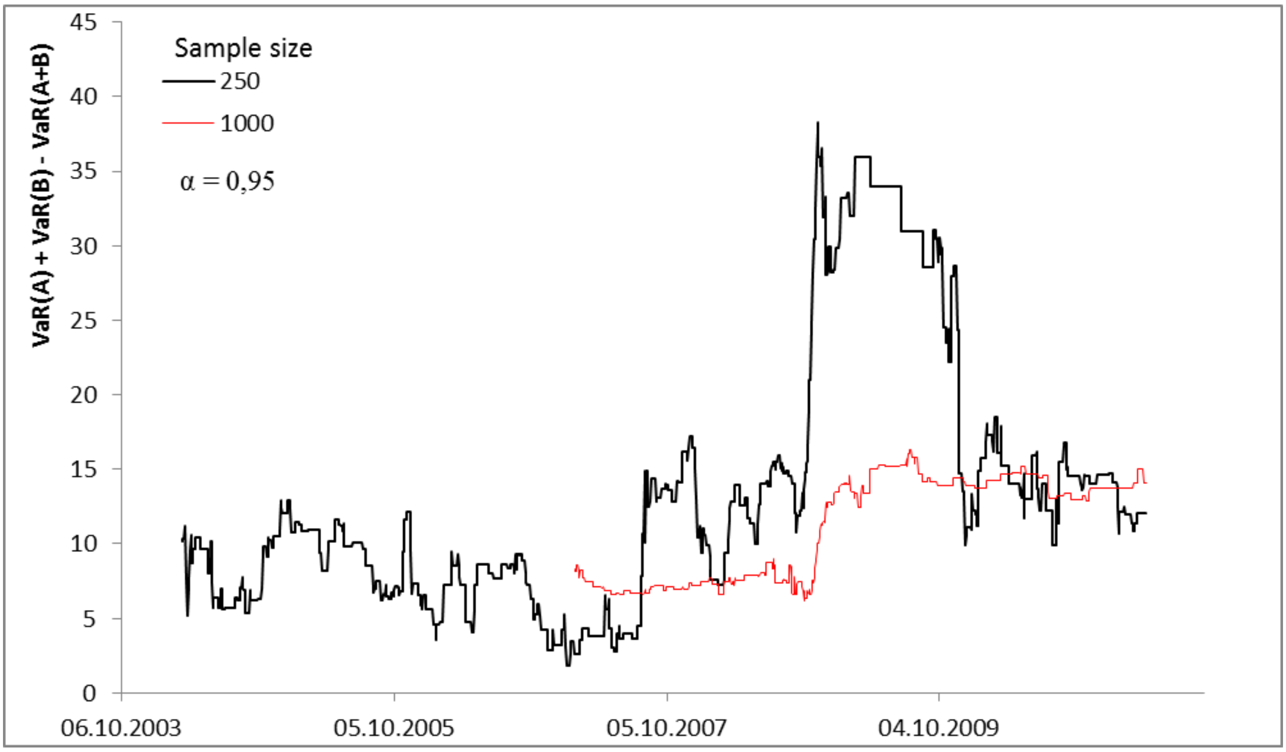

Figure B.3: Test 3 - *Portfolio A: $1000 of B.0. Portfolio B: $1000 of E.1*

.

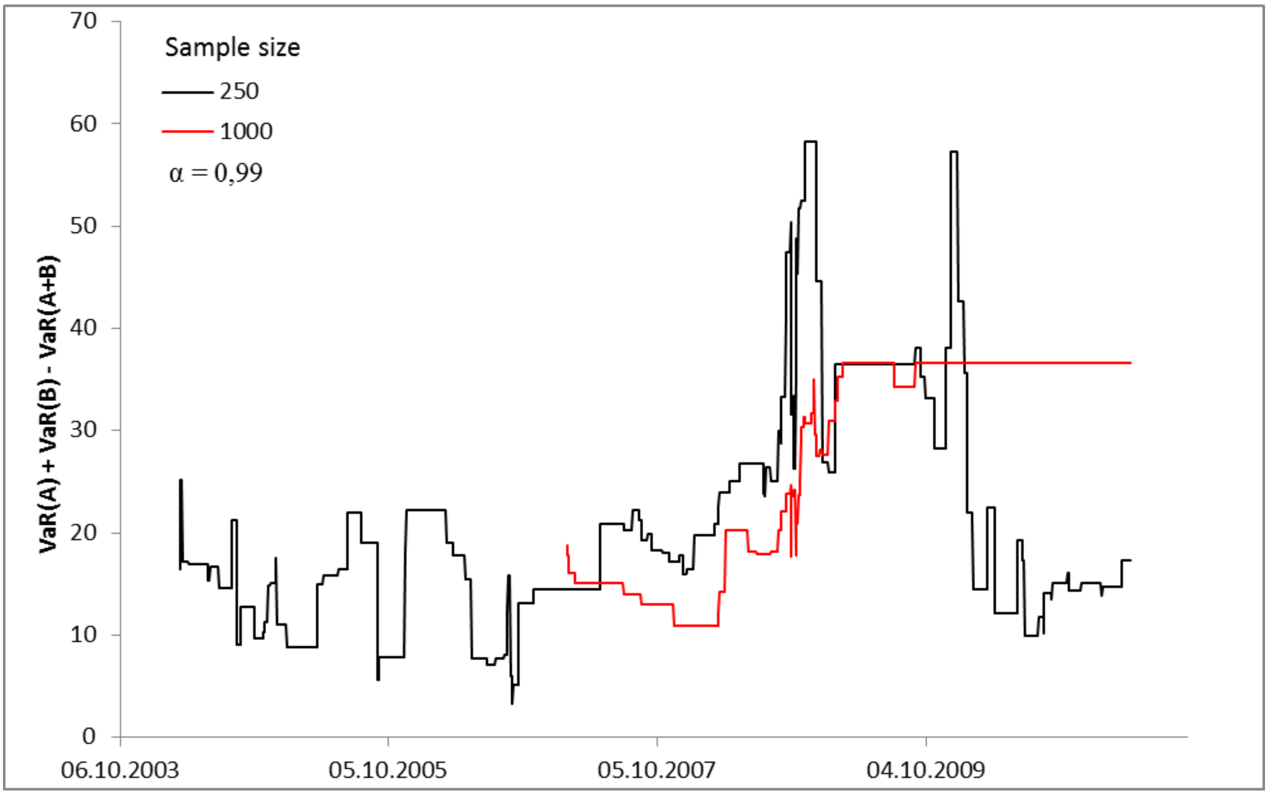

Figure B.4: Test 3 - *Portfolio A: $1000 of B.0. Portfolio B: $1000 of E.1*

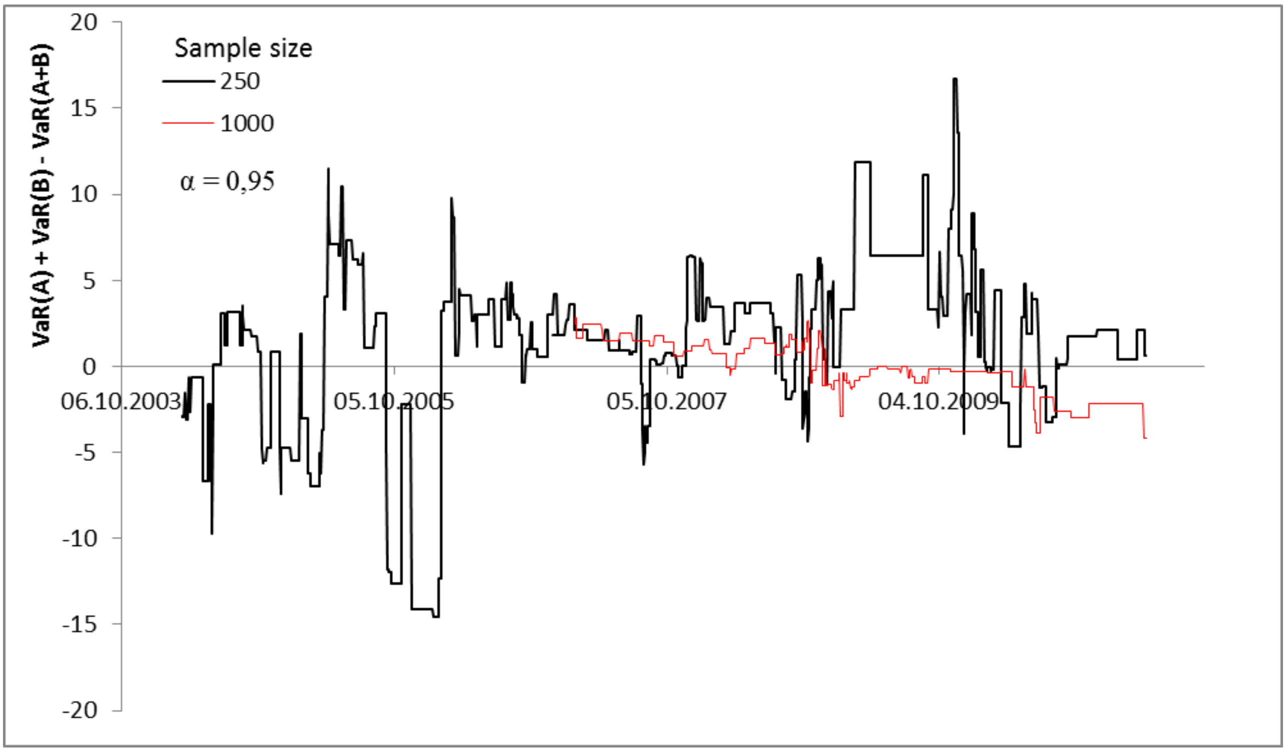

Figure B.5: *Test 4 - Portfolio A: $1000 each of D.0 + D.3 + E.2. Portfolio B: $1000 each of C.0 + D.2 + E.2.*

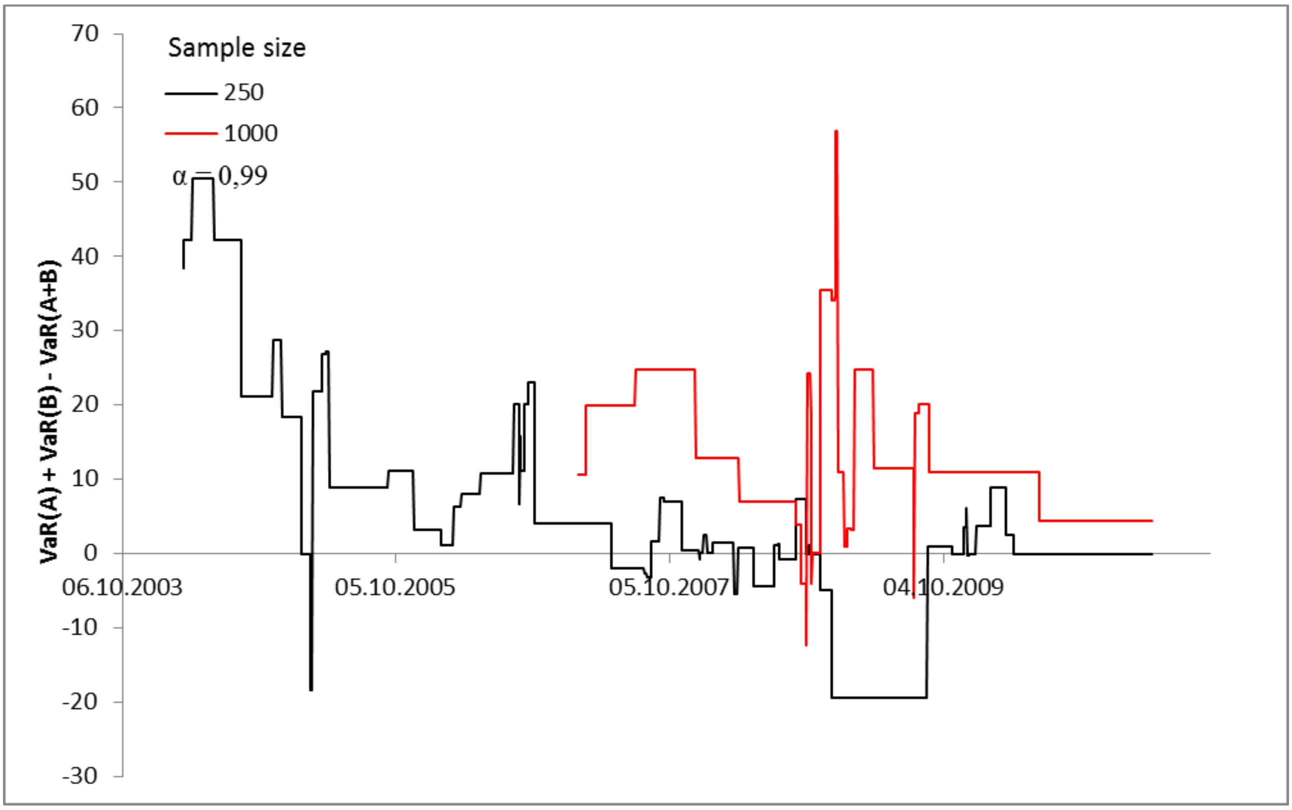

Figure B.6: *Test 4 - Portfolio A: $1000 each of D.0 + D.3 + E.2. Portfolio B: $1000 each of C.0 + D.2 + E.2.*

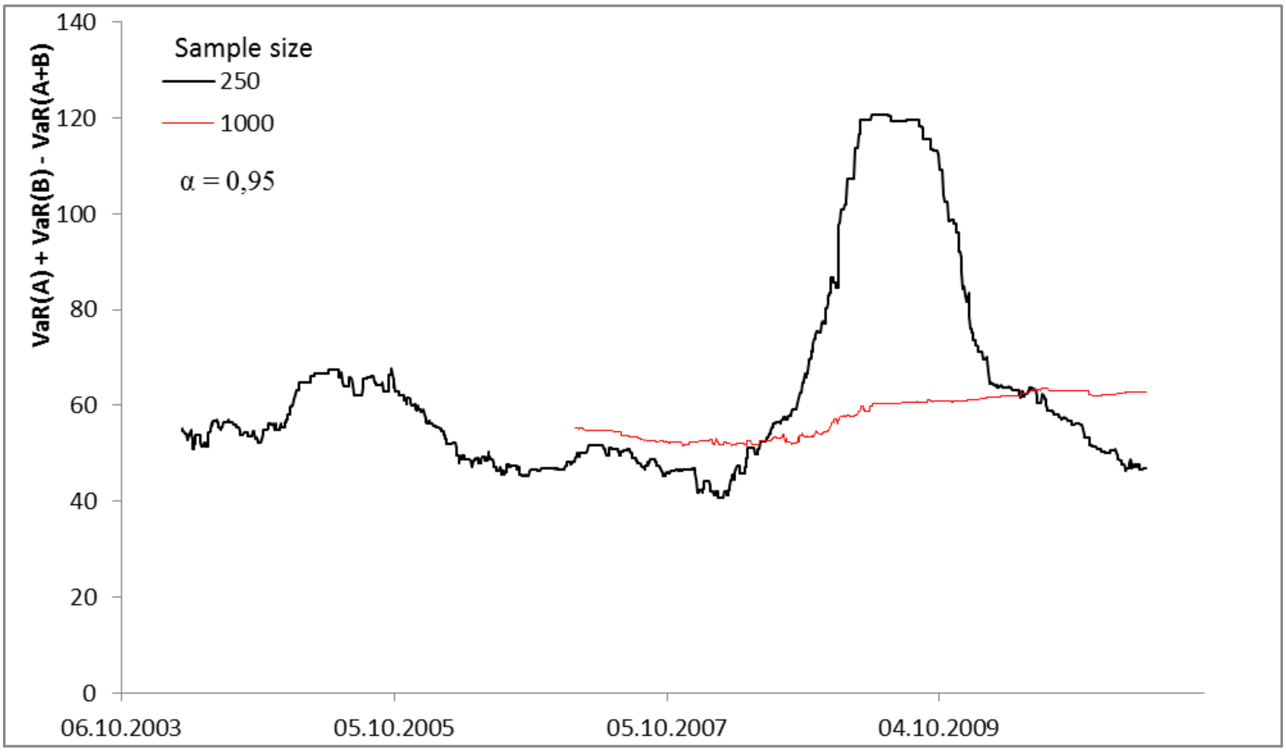

Figure B.7: *Test 5 - Portfolio A: $1000 long position of A.0. Portfolio B: $1000 short position of A.1.*

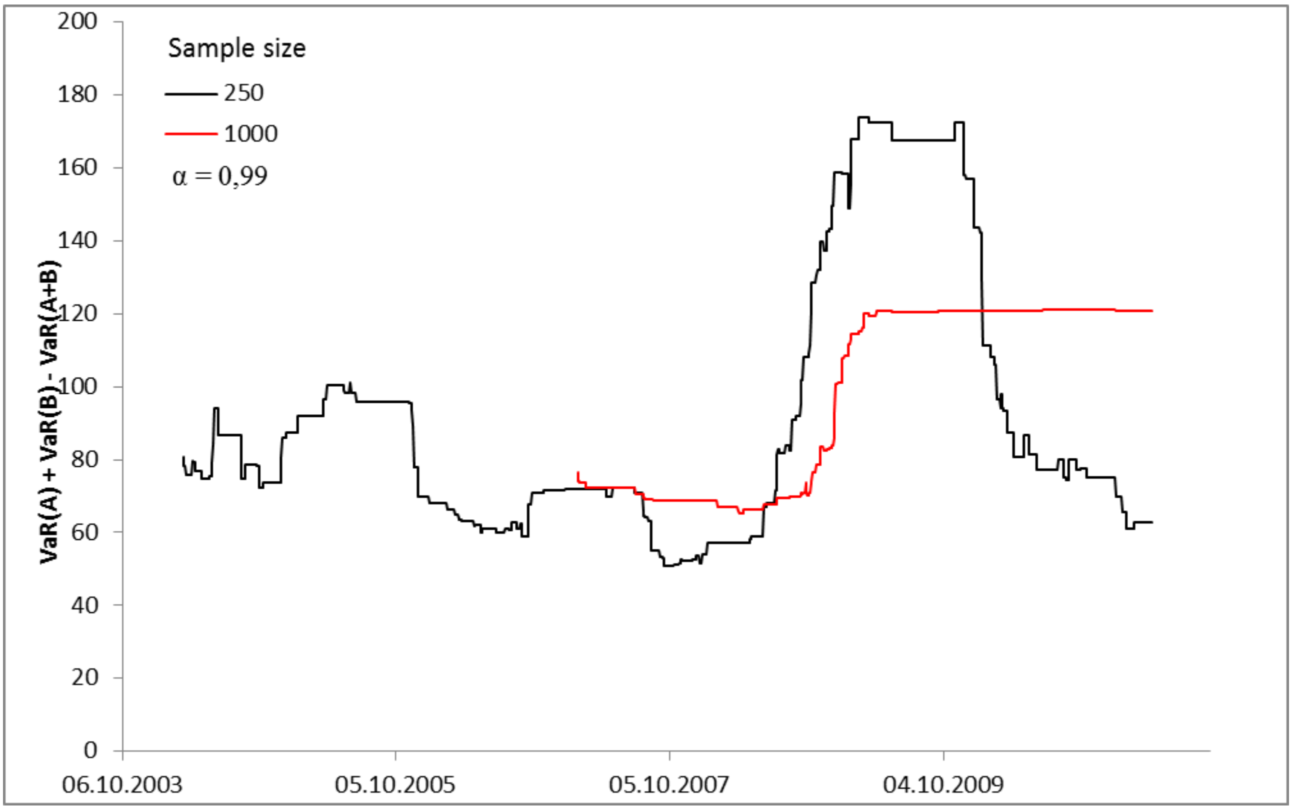

Figure B.8: *Test 5 - Portfolio A: $1000 long position of A.0. Portfolio B: $1000 short position of A.1.*

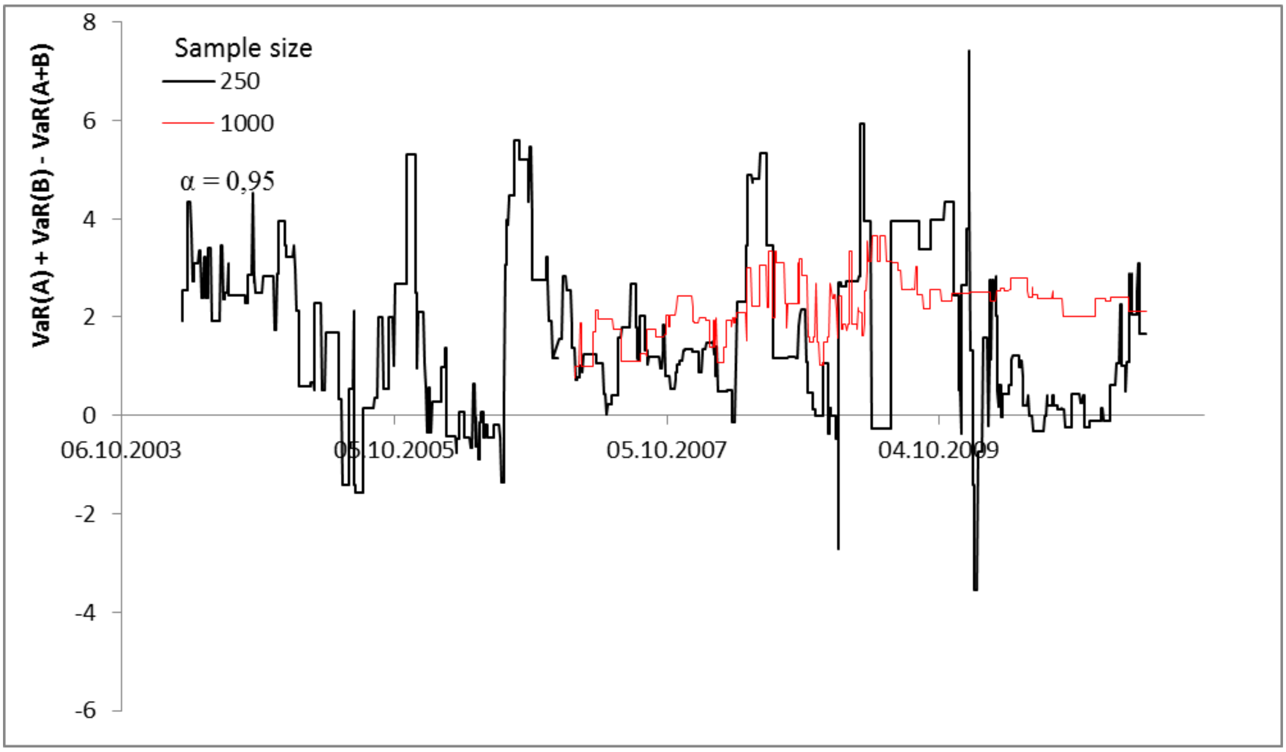

Figure B.9: *Test 6 - Portfolio A: \$1000 short position of A.0. Portfolio B: \$1000 short position of A.1.*

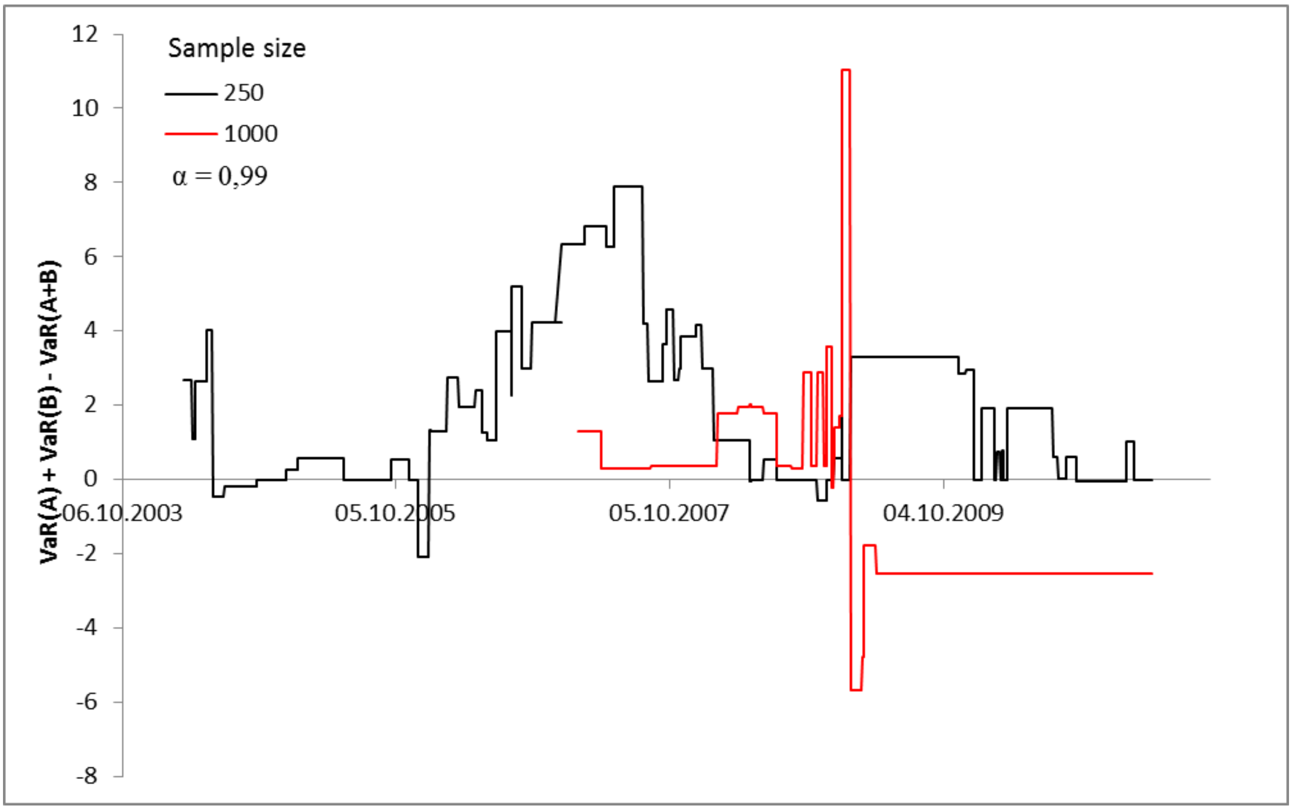

Figure B.10: *Test 6 - Portfolio A: \$1000 short position of A.0. Portfolio B: \$1000 short position of A.1.*

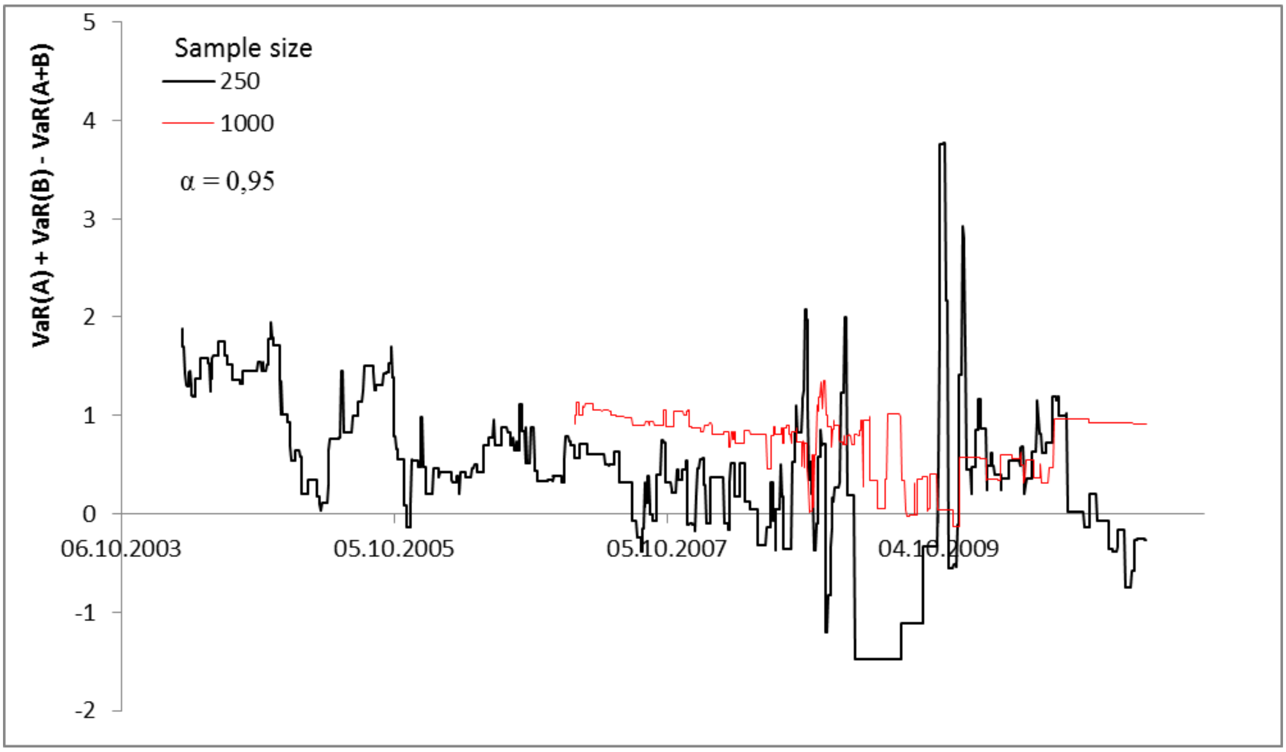

Figure B.11: *Test 7 - Portfolios as described in Table 5.5*

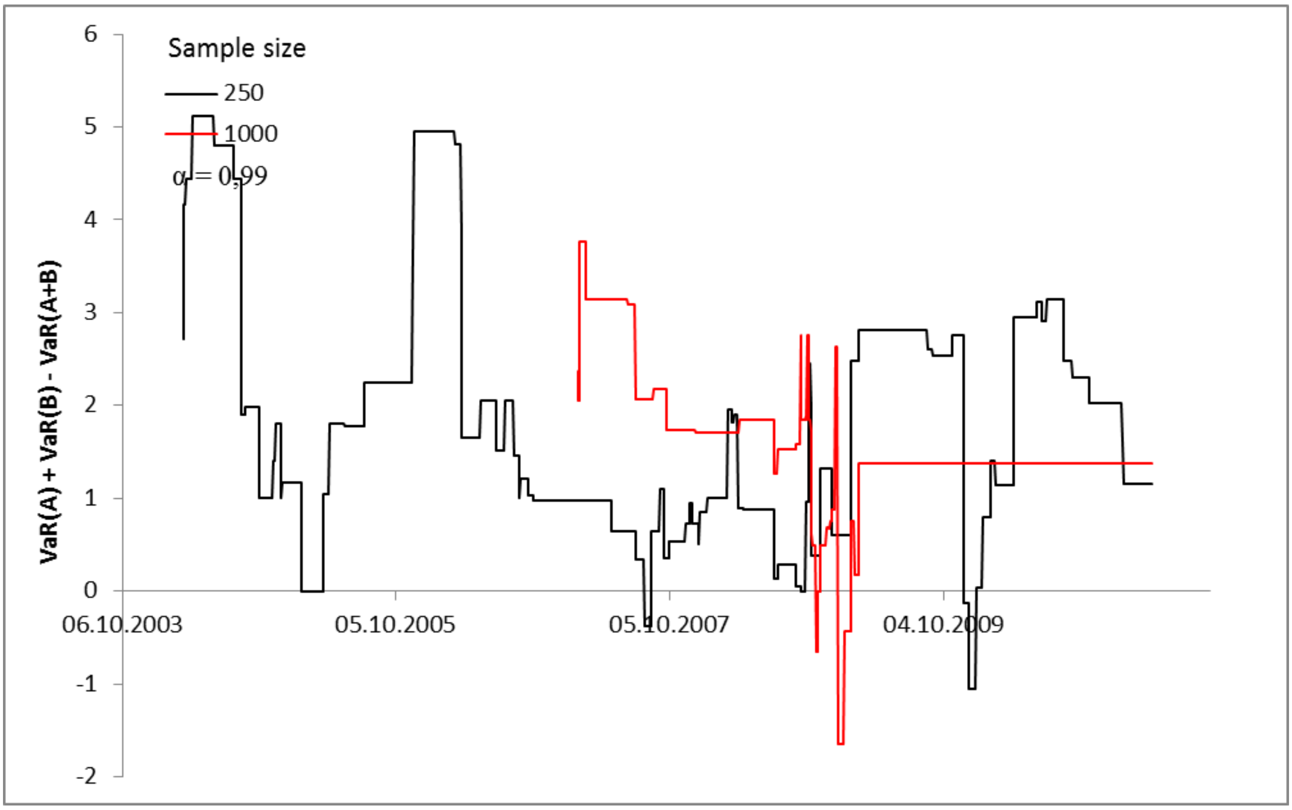

Figure B.12: *Test 7 - Portfolios as described in Table 5.5*

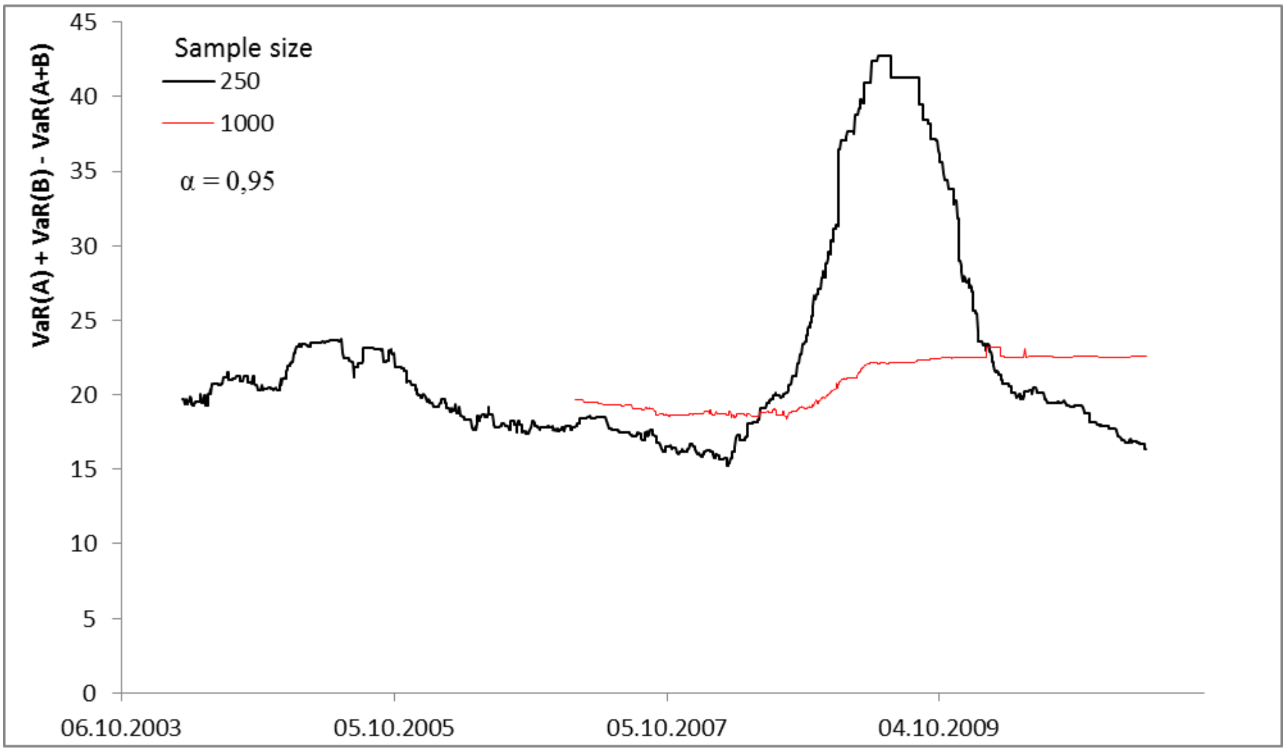

Figure B.13: *Test 8 - Portfolios as described in Table 5.5*

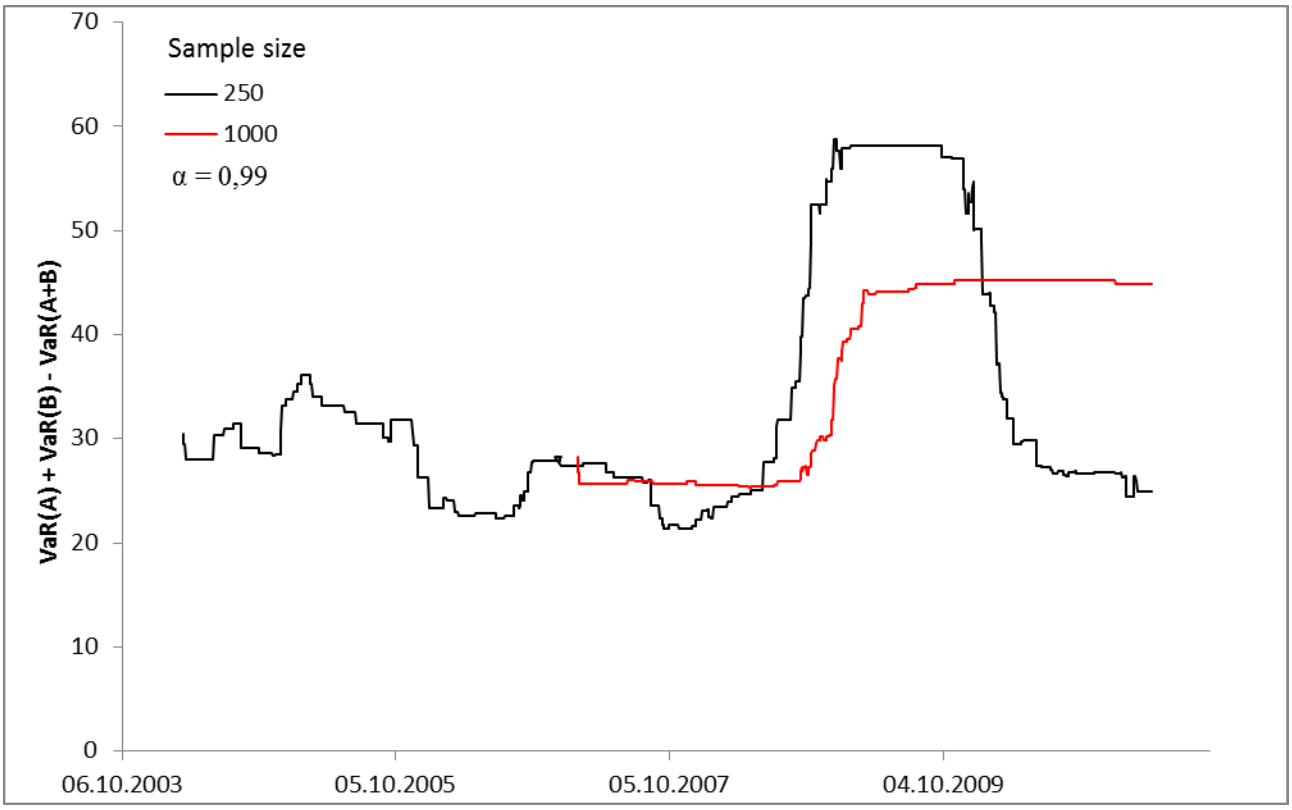

Figure B.14: *Test 8 - Portfolios as described in Table 5.5*