



Universitetet
i Stavanger

**DET SAMFUNNSVITENSKAPELIGE FAKULTET,
HANDELSHØGSKOLEN VED UIS
MASTEROPPGAVE**

STUDIEPROGRAM:

Master i Økonomi og Administrasjon

OPPGAVEN ER SKREVET INNEN FØLGENDE
SPESIALISERINGSRETNING:
Anvendt Finans

ER OPPGAVEN KONFIDENSIELL? Nei
(NB! Bruk rødt skjema ved konfidensiell oppgave)

TITTEL:

Frykt og usikkerhet i det norske aksjemarkedet.

ENGELSK TITTEL:

Fear and Ambiguity in the Norwegian Stock Market.

FORFATTER(E)

Studentnummer:

895759

953362

Navn:

Helga Lien

Morten Hauge

VEILEDER:

Lorán Grady Chollete

OPPGAVEN ER MOTTATT I TO – 2 – INNBUNDNE EKSEMPLARER

Stavanger,/..... 2012

Underskrift administrasjon:.....

There is nothing to fear but fear itself.

- Franklin D. Roosevelt

Preface

This Master Thesis is the concluding part of our Master's degree in Economics and Business Administration at the University of Stavanger. Our specialization in the Master's program is Applied Finance.

The subject we have chosen to write is about fear and ambiguity in the Norwegian stock market. This is a subject with very little research, which is one of the main reasons why we want to write about this. In writing this thesis we have used a combination of skills learned in our studies, and new skills learned whilst writing the thesis. To use what we have learned and put it into practice have been both rewarding and a tremendous learning experience, and we hope that we will take this experience with us now that we are starting our careers.

There are several people who played a part in making the thesis becoming a reality. First we would like to give a big thank you to our advisor, Professor Loran Grady Chollete for giving us good feedback and inspiring us with his great knowledge on the subject. We would also like to thank Linn Furuvald at the Oslo stock exchange for providing us with the index option data we needed. In addition, we would like to thank Brage Refve Vik and Thomas Tardy who helped us with SAS, software which was crucial for the thesis. A big thank you also goes to our nearest families for supporting us throughout our five years of studies. We would also like to thank each other for a great semester with hard work and excellent cooperation.

Lastly, I would like to give a special recognition to my uncle, Svein Hauge, who sadly is no longer among us. Without his help in high school my Masters degree might not have been possible.

Abstract

When dealing in financial markets, knowing how much fear and ambiguity there is can be crucial. The reason for this is that when ambiguity increases to the point where it becomes fear, things can happen very fast in these markets. Over the years, considerable amounts of effort have been put in to measuring and predicting these emotions, and perhaps the most famous is the volatility index (VIX) based on S&P 500 index options. Since the VIX came there have been many replicas in other markets, but to our knowledge there have never been one for the Norwegian market. In this thesis we want to correct this, and in addition we want to look at other possibilities for measuring fear and ambiguity in the Norwegian market. We therefore defined the following problem: *Can we measure fear and ambiguity in the Norwegian stock market?*

In this thesis we make the following three contributions: First we survey the most recent literature on decision theory and risk-taking, including papers published in 2011. Second, we extend existing empirical risk research by constructing VIX and FEARS measures for the Norwegian market, which we name NVIX and NFEARS. Third we evaluate the comparative performance of our fear and ambiguity measures in Norway.

The main part of this thesis is about constructing and performing econometrical analysis on the NVIX. Especially the construction was very time consuming and it involved; retrieving and sorting all the index option trades made on the OBX from 1997-2012; learning how to make a volatility index through reading the “VIX White” made public on the CBOE website; making some sample NVIX values in excel; learning how to use SAS (Statistical Analysis Software); and finally making the NVIX in SAS. All of these points took a lot of time, but since we had no experience in using SAS the two latter points were particularly time consuming. In addition we attempt to measure investor sentiment by making the NFEARS, which consists of various negative economical search words made in Google. When the NVIX and NFEARS is made we test them using correlation and econometrical analysis.

Tables

Table 1: Outcome as a function of the number of the ball, p. 10

Table 2: Companies and weighting on the OBX index, p. 34

Table 3: Correlation matrix 2003-2012, p. 40

Table 4: Correlation matrix 2007-2012, p.40

Table 5: Descriptive statistics, p. 41

Table 6: Hypothesis, p. 42

Table 7: ADF test, p. 44

Table 8: Regression 1, 2003-2012, p. 45

Table 9: Regression 2, 2007-2012, p. 46

Table 10: Regression 3, 2003-2012, p. 47

Table 11: Regression 4, 2007-2012, p. 47

Table 12: Regression 5, NVIX relationship with OBX, 2003-2012, p. 48

Table 13: Correlation matrix 2008-2012, p. 57

Table 14: Descriptive statistics, p. 58

Table 15: Regression 6, 2008-2012, p. 59

Table 16: Regression 7, 2008-2012, p. 61

Figures

Figure 1: A hypothetical value function, p. 19

Figure 2: Utility function of a risk averse individual, p. 24

Figure 3: Utility function of a risk neutral individual, p. 25

Figure 4: Utility function of a risk loving individual, p. 25

Figure 5: Graph showing the relation between the VIX index and the S&P 500, p. 28

Figure 6: Norwegian VIX, p. 37

Figure 7: NVIX, OBX, p. 38

Figure 8: First difference NVIX, p. 43

Figure 9: VIX and NVIX, p. 49

Figure 10: Impulse response, p. 50

Figure 11: «Økonomisk krise», p. 55

Figure 12: NFEARS index 2008-2012, p. 57

Figure 13: First Difference NFEARS, p. 58

Figure 14: Trendindikatoren & UMSCENT, p. 60

Table of contents

Preface	I
Abstract	II
Tables	III
Figures	III
1.0 Introduction	1
2.0 Theory	3
2.1 Fear, ambiguity and risk.....	3
2.2 Decision Theory	4
2.3 Von Neuman and Morgenstern Theorem.....	5
2.4 Savage’s subjective expected utility model	11
2.5 Prospect theory	14
2.5.1 The value function.....	18
2.6 Local thinking	19
2.7 Saliency theory of choice under risk	21
2.8 Risk attitudes	23
2.9 Investor sentiment	25
3.0 The Volatility Index	28
3.1 Index options	29
3.2 Construction of the Volatility Index.....	29
3.3 The fear index.....	31
3.4 Critique of VIX	32
4.0 The Norwegian volatility index	34
4.1 OBX index.....	34
4.2 Collecting the data.....	35
4.3 NVIX.....	36
4.4 Econometric NVIX analyses	38
4.5 Macroeconomic factors	39
4.6 Correlation Analysis.....	40
4.8 Descriptive statistics.....	41
4.9 Hypothesis	41
4.10 Model specifications	42
4.11 Regression	45
4.11.1 Regression with lag	46
4.12 NVIX and fear.....	48

4.13 The predictive power of NVIX	50
4.14 Evaluation NVIX.....	51
5.0 FEARS.....	53
5.1 Norwegian FEARS.....	55
5.1.1 Collecting the FEARS data	55
5.1.2 Econometric NFEARS analysis	57
5.1.3 Hypothesis.....	58
5.1.4 Regression	58
5.1.5 Survey based approach.....	59
5.1.6 Evaluation NFEARS	61
6.0 Conclusion.....	63
7.0 References	65
8.0 Appendix	70
8.1 Histograms, P-P plots and scatterplots retrieved from SPSS	70
8.2 SAS Codes.....	73
8.3 STATA codes.....	80

1.0 Introduction

In August 2011 stock markets all over the world declined as the threat of the European debt crisis flared up. The main concern was (and still is) that Greece would go bankrupt and that larger European economies like Italy and Spain would also go bankrupt soon after. These problems had been well known for some time, but during the summer the investors gave it more and more focus. This attention to the potential crisis resulted in panic in the world's stock markets in August 2011, and most European stock markets declined over 20% in few weeks. Since this steep fall was due to "old news" we were amazed by this huge reaction from investors. This got us thinking about the impact fear and uncertainty has in financial markets, and that is what inspired us to write this thesis. With looking more closely at investor sentiment we want to answer the question: *can we measure fear and ambiguity in the Norwegian stock market?*

How an investor, or decision maker, acts when faced with uncertainty is not an easy thing to predict. In the mid 1900s von Neumann and Morgenstern (1944) and Savage (1954) offered two sets of axioms to explain how rational decision makers should react when facing choice under uncertainty. These theories were later critiqued by Kahneman and Tversky (1979) who offered a new theory called "Prospect theory". In this theory the focus is on explaining how the decision maker actually makes decisions when facing uncertainty and not how he should react. In this thesis we want to see if we can measure the level of fear and ambiguity in the Norwegian stock market. Since an investor in the Norwegian stock market faces a lot of uncertainty when making his decisions, it is important for us to have an understanding both of how this investor should react, and how he usually reacts. We will therefore use these theories as the basis for our thesis and explain them more thoroughly.

Nowadays there are several sophisticated indices designed to measure the level of fear and uncertainty in financial markets. However, few of these indices are made for the Norwegian market. In our opinion the most fascinating measure of fear and uncertainty is the Volatility Index (VIX). Roughly speaking this is an index who measures the volatility on the S&P 500 by looking at the bid/ask spread between S&P 500 index options. The VIX index is often referred to as the "fear index" since options are a popular tool to hedge against price drops in financial products. Since the VIX index was introduced in 1993 there have been made many

similar indices for other markets, but to our knowledge there have never been made one for the Norwegian market. Our main objective for the thesis is therefore to create a Norwegian volatility index.

When doing our research for this thesis we also found some other studies done on the subject. One article we found particularly interesting was a measure of investor sentiment by the use of Google search terms. Da, Engelberg and Gao (2009, revised 2011) uses different negative economical search terms made in Google to make investor sentiment indices called “Financial and Economic Attitudes Revealed by Search (FEARS)”. In addition to the Norwegian volatility index, we also want to make a FEARS index for Norway.

2.0 Theory

The following chapter is a representation of what we find to be the most important theoretical aspects when explaining fear and ambiguity. The chapter starts off with defining fear and ambiguity before moving on to some history involving decision theory. This is followed by an explanation of the two leading theories on how people evaluate choices under uncertainty; Expected utility theory and Prospect theory. In expected utility theory our main focus will be on explaining the axioms and theorems of von Neumann and Morgenstern, and Savage, as well as pointing out their advantages and shortcomings. We will then proceed to explain prospect theory as it was first presented by Daniel Kahneman and Amos Tversky in 1979, before we go onto explaining some more recent studies done by Gennaioli and Shleifer (2008) and Gennaioli, Shleifer and Bordalo (2010). Gennaioli and Shleifer (2008) present a memory based model of probabilistic inference called local thinking. This is a model of intuitive inference as a continuation and improvement of Khaneman and Tversky's prospect theory. Gennaioli, Shleifer and Bordalo (2010) continue this way of thinking and develop the salience theory of choice under risk. This theory claims that a decision maker's attention is drawn to salience payoffs. We round of the chapter with a presentation of the different risk attitudes before we talk about the different aspects of investor sentiment.

The reason why we take the time to present these theories is that they explain both how a decision maker should choose, and how the decision maker actually chooses when making choices under uncertainty. This can be related to the choices being made in stock markets all around the world today. Stock markets will always have some portion of uncertainty or ambiguity. What we find interesting is when this uncertainty takes over, and creates fear in the market, thereby causing the markets to rapidly decline.

2.1 Fear, ambiguity and risk

Although this thesis seeks to investigate the amount of fear and ambiguity in the Norwegian stock market, we will also present some theory involving risk. This is because fear, ambiguity and risk are closely related terms in financial theory. However, we will not attempt to measure risk, as it would prove near impossible to know all outcomes and probabilities in the stock market.

Ambiguity or uncertainty is a situation where probabilities of the outcome cannot be given or found from previous data, while risk is a situation that exists when all possible outcomes and their probabilities are known (Ackert & Deaves, 2010). People's attitude toward risk and ambiguity is an ongoing topic in research, and has been for a long time. There is convincing evidence that most people tend to avoid risk in most circumstances. However, if people are sufficiently compensated, they are willing to take the risk. Expected utility theory presents the term "risk attitude" and how people's attitude towards risk differs. The utility function is used to define risk preferences and they present the "risk averse", "risk seeker" and the "risk loving" person. Even though people tend to prefer the safe outcome to the risky, they rather know the probabilities than to be faced with uncertainty. This is called ambiguity aversion and is the tendency to prefer risk over uncertainty (Ackert & Deaves, 2010). Ackert and Deaves (2010) present a view that ambiguity aversion is more an emotional behavior, than a heuristic one. This emotional behavior could be interpreted as fear.

As mentioned, people tend to be risk averse but are willing to take the risk if compensation for taking the risk is high enough. In life one cannot avoid all risks and we can't always expect a return when taking the risk. For example, there is always a possibility that we may get sick, get in a car accident or that our house burns down. In modern society we have developed insurances that cover us financially if any of these things should happen to us. Insurance is a form of risk management used to hedge against possible risks. In this paper we concentrate on the financial markets and monetary risk. We want to investigate how fear and ambiguity of losing money affects the Norwegian financial market. In the last 30 years, option trading has become increasingly important in finance and is traded on many exchanges around the world. An option is a derivative financial instrument that gives the holder the right to buy or sell the underlying asset by a certain date for a certain price. Options can be traded as insurance for another trade, e.g. stock trade. This is called hedging and can be very useful to provide protection against adverse events (Hull, 2011). Thus, by long option trades like buying either a call or put investors can reduce their exposure to risk and ambiguity.

2.2 Decision Theory

Decision theory and decision under uncertainty has been studied since the mid- 17th century. The concepts of probability and expectation were one of the inspirations to this theory, and are most associated with Blaise Pascal and his famous decision theory "wager". The theory of

Pascal was designed to convince non-believers that it was better to believe. Pascal introduced several basic notions of decision theory: (i) the decision matrix, (ii) domination between acts, (iii) expected utility maximization, according to which the choice between un-dominated acts should be according to the mathematical expectations of the utility of the outcomes they yield, (iv) subjective probability over the states and (v) non-unique probabilities. (Gilboa, 2009, May)

Whether uncertainty can be quantified probabilistically has been a topic of dispute from the beginning of the studies on probability to this very day. Frank Knight (1921) argued that this is not the case, and he separates the situations of risk, where probabilities can be assumed given, and uncertainty where probabilities can't be given or found from past statistical data. Frank Ramsey (1931) had an opposite view and suggested defining and measuring one's subjective probability by one's willingness to bet. He believed that a reasonable decision maker would behave as if he had a subjective probability that guided his decision. Bruno de Finetti (1937) was the first to introduce a set of conditions on presumably observable choices and showed that they are equivalent to the claim that the decision maker maximizes expected value relative to a measurable probability. The conditions are called axioms because they are presented as intuitive and normative. Thus, expected utility theory is the theory of decision-making under risk based on a set of axioms for a preference ordering. John von Neumann and Morgenstern derived a similar concept of utility axioms as a by-product to their introduction of game theory in "Games and Economic Behavior" (1944). This theorem and its axioms will be presented through below.

2.3 Von Neuman and Morgenstern Theorem

In financial decision making, as well as in decision making in general, there normally exists a great deal of uncertainty about different outcomes. In 1944 John von Neumann and Oskar Morgenstern published the book "Theory of Games and Economic Behavior", where they attempted to define rational behavior for people when facing uncertainty. In doing so von Neumann and Morgenstern developed the *expected utility theory*. This normative theory seeks to explain how an individual *should* act when he is confronted with making a decision under uncertainty. The fact that the theory is normative means that it describes how people should rationally behave, as opposed to a positive theory, which characterizes how people actually behave. (Ackert & Deaves, 2010)

Itzhak Gilboa (2009) writes that, in their theory, “Neumann and Morgenstern considered a presumably-observable preference relation between pairs of “lotteries”, namely random variables with known distributions, and they showed that a set of axioms on the relation is equivalent to the claim that this relation can be represented by a utility function, such that, confronted with any two choices, the decision maker would opt for the one that has a higher expected utility.” (Gilboa, 2009, May, pp. 2-3))

Lotteries:

Hal R. Varian (1992) offers an explanation of lotteries, which we find useful to know before moving on to the axioms of rationality.

Lotteries have different outcomes with different probabilities, all summing to one. A lottery can be written in the form $p*x + (1-p)*y$, which says that the consumer receives prize x with probability p and prize y with probability $(1-p)$. The prizes can be money, bundles of goods, or further lotteries. Varian (1992) makes three assumptions about the consumer’s perception of the lotteries open to him.

- The first assumption is that if the probability of getting a prize is one, then that is the equivalent of getting a price for certain: $1*x + (1-1)*y \sim x$
- The second assumption is that the consumer doesn’t care about in which order the lotteries are described: $p*x + (1-p)*y \sim (1-p)*y + p*x$
- The third assumption says that the consumer’s perception of how attractive a lottery is depends only on the net probabilities of receiving the various prizes:
 $q*(p*x + (1-p)*y) + (1-q)*y \sim (qp)*x + (1-qp)*y$

Under these three assumptions Varian (1992) defines λ to be “the space of lotteries available to the consumer”. He assumes the consumer has preferences on this lottery space, meaning that the consumer is free to choose between two lotteries if two lotteries should be available. (Varian, 1992, pp. 172-173)

To explain rationality in people when facing uncertainty von Neumann and Morgenstern made three axioms of rationality: *Weak order*, *Continuity*, and *Independence*. (Gilboa, 2009, p.115)

- V1. **Weak order:** \geq is complete and transitive.
- V2. **Continuity:** For every $P, Q, R \in L$, if $P > Q > R$, there exist $\alpha, \beta \in (0, 1)$ such that $\alpha P + (1-\alpha)R > Q > \beta P + (1-\beta)R$.
- V3. **Independence:** For every $P, Q, R \in L$, and every $\alpha \in (0, 1)$, $P \geq Q$ iff $\alpha P + (1-\alpha)R \geq \alpha Q + (1-\alpha)R$.

The first axiom, weak order, is by many researchers divided into two axioms, namely completeness and transitivity. To offer a more extensive explanation we will do the same here, and therefore explain completeness and transitivity separately. The completeness axiom assumes that an individual has well defined preferences of choice. This means that in a choice set consisting of x and y , the individual will prefer x to y ($x > y$), prefer y to x ($x < y$), or the individual will be indifferent between x and y ($x \sim y$).

The transitivity axiom simply assumes that the individual's preference is consistent across any three options, meaning that if the individual prefers x to y and y to z , he will also prefer x to z . So if $x > y$, and $y > z$, then $x > z$.

The second axiom, continuity, says that under a sufficiently small deviation in probabilities we can maintain a separation in any preferences. Since this axiom requires a "sufficiently small deviation in probabilities" we cannot design a real-life experiment where this axiom would be directly violated. This is because its violation would require infinitely many observations. We can however "test" this axiom by the use of some thought experiments. We could, for instance imagine an extreme kind of lottery that will pay $P = \$10$ if you win, $R =$ death if you lose, or $Q = \$0$ if neither. Most individuals will then obviously prefer $\$10$ to nothing, and both to death, so that the preferences would be $P > Q > R$. What the continuity axiom now says is that for a high enough $\alpha < 1$, the individual will have the preference $\alpha P + (1-\alpha)R > Q$. Meaning that the individual would be willing to risk his life with the probability $(1-\alpha)$ in order to win $\$10$.

The third and final axiom, independence, assumes that an individual's preference will hold independently of the possibility of another irrelevant outcome. For instance, if we assume a similar situation as described in the lottery under the second axiom, were we are faced with

three outcomes: P = winning \$10, Q = nothing happens, and R = death. If we now argue that the possibility of death is always present in some way, we can assume that the choice between P and Q should not be affected by R, and therefore the most preferable outcome of P and Q should be chosen regardless of R, thus making P and Q independent of R.

Now that we have stated and explained the three axioms we can present von Neumann and Morgenstern's theorem:

*" $\succeq \subset L * L$ satisfies V1 – V3 if and only if there exists $u: X \rightarrow \mathbb{R}$ such that, for every $P, Q \in L$
 $P \succeq Q$ iff $\sum P(x)u(x) \geq \sum Q(x)u(x)$." (Gilboa, 2009, p. 118)*

This theorem has been subject for some critique during the years. The "framing effect", documented by Daniel Kahneman and Amos Tversky in 1974, show that different representations of the same problem may result in the individual making different choices. Specifically this effect shows that individuals have a tendency to make inconsistent choices, depending on whether the question is framed to focus on gains or losses. In the book "The Framing of decisions and the psychology of choice" Kahneman and Tversky (1981) describes an experiment where the participants are faced with different strategies for preventing a disease. In the experiment the same scenario is presented twice, with different phrasing. The first problem given to the participants offers two alternative solutions to rescue some or all of 600 people who were affected by a deadly disease.

- Option A: Save 200 peoples' lives guaranteed.
- Option B: 33% probability of saving all 600 people and a 66% probability of saving no one.

Both these options have the same expected number of survivors, 200. However, option B is risky. In Kahneman and Tversky's (1981) experiment 72% of the participant chose option A and 28% chose option B.

The second problem was presented to a different group of participants. This problem also has two different solutions giving an expected number of survivors of 200. This problem is however described differently.

- Option C: 400 people will die for certain.
- Option D: 33% probability of no one dying and a 66% probability of everyone dying.

In this group 78% of the participants chose option D (which is the equivalent of option B), and 22% of the participant chose option C (which is equivalent to option A). The difference in preferences in these two groups is essentially the framing effect. The two groups were presented the same problem, but with different language. In the first group there was a positive emphasis on lives saved, whereas in the second group there was a negative emphasis on lives lost. So changing the way you present the problem to an individual will most likely affect his preferences.

This framing effect can be related to von Neumann and Morgenstern's continuity axiom. If we imagine the same example as described in explanation of the axiom above, where you are presented with a lottery where $P = \$10$, $Q = 0$, and $R = \text{death}$. Continuity says that for a small enough $(1-\alpha)$ you would be willing to risk death in order to gain \$10. Now say that $(1-\alpha) = 0.000001$ and $\alpha = 0.999999$, you would possibly be less likely to accept this gamble if you were presented it in a way that gave you a sure gain of \$10 and a 0.000001 chance of dying, as opposed to if you were presented it as a sure gain of \$10 and a 0.999999 chance of living. So under the right type of framing, von Neumann and Morgenstern's continuity axiom could be violated.

It is however important to remember that the continuity axiom is more like a technical condition needed for the mathematical representation of the theorem, rather than a stated fact. As mentioned above we would need infinitely many observations to prove the violation of this axiom, and likewise we would probably need infinitely many observations in order to make the continuity axiom a proven fact.

The independence axiom has also been debated and challenged through the years. The first and probably the most famous challenge to this axiom was proposed by the French economist

Maurice Allais in his book “Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’école Américaine“, from 1953. Allais proposes the following experiment, described by Christian Gollier (2004) in “The economics of risk and time”.

An urn contains 100 balls that are numbered from 0 to 99. There are four lotteries whose monetary outcomes depend in different ways on the number of the ball that is taken out of the urn. The outcomes are described in table 1.

Table 1: Outcome as a function of the number of the ball

Lottery	0	1-10	11-99
L^α	50	50	50
L^β	0	250	50
M^α	50	50	0
M^β	0	250	0

(Gollier, Christian: The economics of risk and time, 2004, p. 14)

Decision makers are subjected to two choice tests. In the first test, they are asked to choose between L^α and L^β , whereas in the second test, they must choose between M^α and M^β . Many of the decision makers reported that they preferred L^α to L^β , but they preferred M^β to M^α . Since L^α and L^β have the same outcome when the number of the ball is larger than 10, the independence axiom tells us that these people prefer L^α that gives you value 50 with certainty, to L^β which gives value 0 with probability 1/11 and value 250 with probability 10/11. The Allais paradox is that this argument can also be used with the result in the second test, when considering the preference of M^β over M^α . (Gollier, 2004)

So, we see from this experiment that the outcome between 11 and 99 does have an impact on the choices being made, even though it shouldn't, according to the independence axiom. When outcomes between 11 and 99 are removed L^α and M^α are exactly the same and L^β and M^β are also exactly the same. So in order for the independence axiom to be fulfilled the decision makers should prefer either L^α and M^α or L^β and M^β . As shown above however, the decision makers choose L^α and M^β , thereby proving that the outcome between 11 and 99 does have an impact on their preferences, and are in fact not independent from the other outcomes. Thus resulting in these decision makers violating the independence axiom.

2.4 Savage's subjective expected utility model

Von Neumann and Morgenstern, and Bruno de Finetti's theory concerning expected utility were both powerful and sophisticated, but at the same time they also had their shortcomings. Von Neumann and Morgenstern provided a definition of utility based on a primitive notion of probability, while Bruno de Finetti did the opposite; defined subjective probability based on a primitive notion of utility (Gilboa, 2009, May). Savage (1954) addressed this problem and showed that both utility and subjective probability can be derived with the expected utility maximization rule.

In Savage's book, *The Foundations of Statistics*, he developed a theory of decision making under uncertainty, and used that theory to define choice-based subjective probabilities.

“Savage's ... theorem shows that the axioms are equivalent to the existence of both a utility function and a probability measure, such that decisions are being made as if to maximize the expectation of the utility relative to the probability measure.” (Gilboa, 2008, p. 136)

Savage's model include two concepts; outcomes and states. The set of states (S) represents a list of all the possible scenarios. Savage's “states” resolves all uncertainty and should specify the answer to any question of interest. An event (A) is any subset of (S), $A \subset B$. The set of outcomes (X) is assumed to specify all that are relevant to you when making your decision. The object of choice are acts (F), which are defined as functions from states (S) to outcomes (X). Savage also defines an event that is considered practically impossible, the “null event”. This event is not a logical impossibility, but you don't assign any weight to it when making your decision

Savage's axioms that are needed to hold in order for the theory to hold are stated below as P1-P7. *P for postulate has the same meaning as axiom: assertion without proof.*

P1: states that \geq is a weak order. This axiom is also called completeness. Meaning that all functions are assumed to be comparable; $a > b$, $b > a$ or $a \sim b$.

P2: says the preference between two acts, f and g , should only depend on the values when they differ. For example, assume that f and g only differ on an event A . If A does not happen, f and g has the same outcome. So when comparing the acts f and g we can focus on event A

and ignore other events like A^b . This axiom is often referred to as the “Sure- Thing Principle”.

P3: is convincingly called the monotonicity axiom when the outcomes are monetary payoffs like in our case. A simple explanation is if you take an act (F), which guarantees an outcome (X) on an event A, and you change the outcome (X) to outcome (Y), the preference between the acts should follow the preference between the two outcomes (X) and (Y).

Formally, for every $f \in F$, non null events $A \subset S$ and $x, y, \in X$, $x \geq y$ if $f^x_A \geq f^y_A$.

P4: This axiom is often referred to as the “Independence of probabilities”. We wish to measure the ranking of outcomes and events. The choice of an event must be independent of the values given their ordering. Another way to put it is that the choice of event a person prefers should not be effected by the size of the price or payoff. This axiom also emphasize that we must be able to say which event A, B is more likely.

Formally, for every $A, B \subset S$ and every $x, y, z, w \in X$ with $x > y$ and $z > w$,

$$y^x_A \geq y^x_B \text{ if } w^z_A \geq w^z_B$$

P5: This axiom is important when it comes to showing a unique probability measure, which will be the “subjective probability of the decision maker”. P5 states that there are $f, g \in F$ such that $f > g$. If this does not hold, we get $f \sim g$ for every $f, g \in F$. That again would say that $\geq = F \times F$ and the probability measure would not be unique and the axiom would be violated. In order for the utility function to be unique there should be $f > g$ in every $f, g \in F$. This relation is represented by expected utility maximization.(savage book p.146)

P6. The two remaining axioms are more technical, which are needed for the mathematical proof but maybe not necessary for the conceptual grounds. This axiom consists of two types of constraints. As it is stated in the book (side 147) there is a “flavor of continuity with an Archimedean twist. Let’s say we have acts $f, g \dots f > g$, and we wish to state some notion of continuity. Meaning we want to say that f' is “close” to f . This axiom will make a weaker requirement, allowing the partition to depend on the values of f' .

Formally, P6 requires that for every $f, g, h \in F$ with $f > g$ there exists a partition of S , $\{A_1, \dots, A_n\}$ such that, for every $i \leq n$, $\{ \}$
 $f^h_{A_i} > g$ and $f > g^h_{A_i}$

We have to be able to partition the state space into finitely many events and outcomes, each of which is not too significant. But the states need to be infinitely many and the probability measures need not to be dividable. “No atoms”.

P7. States the following: “Consider acts $f, g \in F$ and events $A \subset S$. If it is the case that for every $S \in A$, $f \geq_A g(s)$, then $f \geq_A g$, and if for every $S \in A$, $g(s) \geq_A f$, then $g \geq_A f$. P7 requires” that if f is weakly preferred to any particular outcome that g may obtain, and then f should be weakly preferred to g ”. This makes sense because we already have the axioms P1-P6 to rely on. Now that we have explained Savages axioms, we will present his theorem.

“ \geq satisfies P1-P7 if and only if there exist a nonatomic finitely additive probability measure μ on $S (=S, 2^S)$ and a non- constant bounded function $u: X \rightarrow \mathbb{R}$ such that, for every $f, g \in F$

$$f \geq g \text{ if } \int_S u(f(s))d\mu(s) \geq \int_S u(g(s))d\mu(s)$$

Furthermore, in this case μ is unique, and u is unique up to a positive linear transformation.”
 (Gilboa, 2008, p. 156)

Savage believed that rational people preferences should satisfy the above axioms, and he showed that the axioms imply that the preferences are congruent with a ranking by subjective expected utility. These axioms describe how he thinks rational people ought to behave, not how they behave in real life. Almost from the moment Savage developed this theory there have been critiques about the descriptive validity. In particular axiom P2, “The sure thing principle”. The most severe criticism to this is due to Ellsberg (1961), who demonstrated that individuals display choice patterns that are inconsistent with the existence of beliefs representable by probability measure by using simple mind experiments.

Axiom P3 and P4 implies that the preference relation is state independent. This also means that the ranking of consequences and bets are independent of the underlying events. This implies risk attitudes that are event-independent, but does not rule out the effect the state could have on the decision-makers well-being, or that the utility of the consequences is state dependent. Since the utility and probability presented by Savage are unique as a pair it is possible to define new probability measures and state- dependent utility functions, and from there obtain new subjective expected utility representation without violating any of the axioms. This shows that the uniqueness of the probability is conditional on the understanding that the utility function is state independent.

Another aspect of the model that is unsatisfactory concerns the interpretation of null events. Ideally a null event should be ascribed zero probability only if the decision maker believes it is impossible. In savages model the “definition of a null event is if the decision maker displays indifference among all acts that agree on the payoff on the complement of the said event.” (Karni, 2005). This definition does not separate events that the decision maker thinks is impossible and ones whose possible outcome he perceives as equally desirable. Possible or even likely events in the decision maker’s view could therefore be defined as null events and assigned non probability. Edi Karni presents an example in his paper that describes such a situation. “... a passenger who is indifferent to the size of his estate in the event that he dies is about to board a flight. For such a passenger, a plane crash is a null event and is assigned zero probability, even though he may believe that the plane could crash” (Karni, 2005, p. 10). This problem shows the implicit and unverifiable assumption that in every event some outcomes are strictly more desirable is needed, not to result in misrepresentations of beliefs.

2.5 Prospect theory

Now that we have given a presentation of what we believe to be the most important theories in expected utility theory, we will move on to present a recognized theory which originated as a response to expected utility theory.

Prospect theory was introduced in 1979 by Daniel Kahneman and Amos Tversky in a paper about decision making under uncertainty called ”Prospect theory: An analysis of decision under risk.”. In this paper prospect theory is presented as a critique of expected utility theory as a descriptive model. Descriptive theory looks at which choices people actually make and

then bases models on these observations. In contrast to normative theory which says that reasonable people should act in a certain way. So what Kahneman and Tversky said with prospect theory was that the expected utility theory wasn't good at describing actual behavior. (Ackert & Deaves, 2010). To prove this, Kahneman and Tversky designed several questions about monetary decisions involving different outcomes and probabilities. These questions were then presented to students in Israel, Sweden and the US. The pattern of results was essentially identical across the countries, and showed that people don't tend to act in accordance with expected utility theory (Kahneman & Tversky, pp. 264-265).

Prospect theory distinguishes between two phases in the choice making process; editing and evaluation. In the editing phase, the decision maker makes an analysis of the different prospects, which will often result in a simplification of these prospects. The process of editing can be divided into four different operations, these are; *coding*, *combination*, *segregation* and *cancellation*. These operations are described in the next paragraph. In the evaluation process, the decision maker considers the edited prospects and chooses the prospect with the highest value.

In expected utility theory it is assumed that the decision maker perceives outcomes as final states, whereas in prospect theory the outcomes are perceived as gains or losses. The process in which the outcome will be viewed as a gain or loss is called *coding*. Every decision maker has his own reference point which corresponds with his current asset position. So any potential gain or loss will be defined relative to this point. However, the position of the reference point, and the consequent coding of outcomes as gains or losses, can potentially be affected by the expectations of the decision maker, and by the formulation in which the prospect is offered.

Sometimes prospects can be simplified if they have identical outcomes. This process is called *combining*, and is done by combining the probabilities associated with these outcomes. If we for example have a prospect that have two outcomes in which we can win 200, both with a probability of 25% (200, 0.25; 200, 0.25), the decision maker will perceive this as a 50% chance of winning 200 (200, 0.5), and evaluate the prospect in this form.

The *segregation* operation happens in a prospect that contains a riskless component which the decision maker can segregate from the risky component. For example if the decision maker is

presented with a prospect yielding 300 with 80% probability and 200 with 20% probability (300, 0.80; 200, 0.20), this prospect will be decomposed into a sure gain of 200 and a possibility of receiving 100 more (100, 0.80).

The preceding operations are applied to single prospects. *Cancellation* however is applied to a set of two or more prospects. This operation involves the discarding of identical outcomes with identical probabilities in different prospects, or put in another way, the discarding of “outcome-probability pairs”. For instance, the choice between two prospects yielding (200, 0.20; 100, 0.50; -50, 0.30) and (200, 0.20; 150, 0.50; -100, 0.30) can, by cancellation of the outcome-probability pair (200, 0.20), be reduced to a choice between (100, 0.50; -50, 0.30) and (150, 0.50; -100, 0.30).

The editing phase is followed by the evaluation phase, where we assume the decision maker evaluates all the edited prospects, and then chooses the prospect with the highest value. Kahneman and Tversky make two equations showing the evaluation process. In these equations the overall value of an edited prospect is denoted V , and this value is expressed in terms of two scales, π and v . The first scale, π , associates each probability p with a decision weight $\pi(p)$, this decision weight reflects the impact p has on the total value of the prospect. It should be mentioned however that π is not a probability measure. The second scale, v , assigns a number $v(x)$ to each outcome x , this number reflects that outcome’s subjective value. As mentioned above all outcomes are defined relative to a reference point, v is therefore a measure of the value of deviations from this reference point, thus we can say that v measures whether an outcome is perceived as a gain or a loss.

The equations describes simple prospects with at most two non-zero outcomes. These prospects take this form $(x, p; y, q)$. In this prospect the decision maker will receive outcome x with probability p and outcome y with probability q . If a prospect is strictly positive, its outcomes will all be positive, i.e., if $x, y > 0$ and $p + q = 1$. Likewise the prospect will be all negative if its outcomes are all negative. Should the prospect be neither strictly positive nor strictly negative, we say that it is regular. The first equation is made for evaluating regular prospects. The equation describes how π and v are combined in order to determine the total value of a regular prospect. So if $(x, p; y, q)$ is a regular prospect, then its evaluation can be described by the following equation:

$$(1) \quad V(x, p; y, q) = \pi(p)v(x) + \pi(q)v(y)$$

where $v(0) = 0$, $\pi(0) = 0$, and $\pi(1) = 1$. As mentioned above V is the total value of a prospect, while v is defined on each single outcome. If we have a sure prospect of either a certain gain or a certain loss, these two scales will coincide. Equation form: $V(x, 1.0) = V(x) = v(x)$.

The equation for evaluating strictly positive and strictly negative prospects is somewhat different from the first. Strictly positive and strictly negative prospects will be separated into two components in the editing phase; the riskless component and the risky component. Where the riskless component consist of gains or losses that are given, no matter what, and the risky component consist of some additional gain or loss that is at stake. The evaluation of these prospects is described in the following equation:

$$(2) \quad V(x, p; y, q) = v(y) + \pi(p)[v(x) - v(y)].$$

So what this equation says in words is that, the value of a strictly positive or strictly negative prospect equals the value of the riskless component plus the difference in value between the outcomes, multiplied with the decision weight of the risky outcome. The essential feature of this equation in relation to the first is the application of a decision weight to the difference in value of the outcomes $v(x) - v(y)$, which represents the risky part of the prospect, but not to the riskless part $v(y)$.

These equations of prospect theory presented by Kahneman and Tversky have the same general bilinear form that underlies expected utility theory. The main difference is that Kahneman and Tversky assumes that values are not attached to final states, but rather to changes from a reference point, and that the decision weights does not correlate with the stated probabilities. These differences from expected utility theory will lead the decision maker into violating the expected utility theory axioms, and thereby making irrational choices. This kind of irrational behavior would normally be corrected once the decision maker realizes his mistakes. However, in many situations the decision maker will not have the opportunity to discover that his preferences are irrational and a violation of the decision rules he logically wishes to obey. In circumstances such as these one can expect the anomalies presented by prospect theory to occur (Kahneman & Tversky, 1979)

2.5.1 The value function

One essential part of prospect theory, that separates it from expected utility theory, is the fact that it not considers changes in value as final states, but rather as gains or losses. This should however not be seen as the value of a particular change is independent from the initial position. Kahneman and Tversky said that “value should be treated as a function in two arguments: the asset position that serves as reference point and the magnitude of the change (positive or negative) from the reference point”.

The latter can be shown by imagining that a gain of 100 is more valuable to an individual if the reference point is 200, rather than if the reference point was 1000. This idea is plausible, and Kahneman and Tversky therefore hypothesized that for changes in wealth, the value function will normally be concave above the reference point. Thus saying that the marginal value of gains and losses generally decrease as they get larger. This hypothesis is concerning the shape of the value function when gains and losses are presented in a riskless context. When dealing with risky choices however, Kahneman and Tversky had a hypothesis that the value function is concave for gains and convex for losses. To illustrate this idea they presented the following question to a group of students:

A choice between: A. (6000, 0.25) or B. (4000, 0.25; 2000, 0.25)

And a choice between: C. (-6000, 0.25) or D. (-4000, 0.25; -2000, 0.25)

Most of the students chose B and C. If we imply these preferences into the second evaluation equation from above, we get the following results:

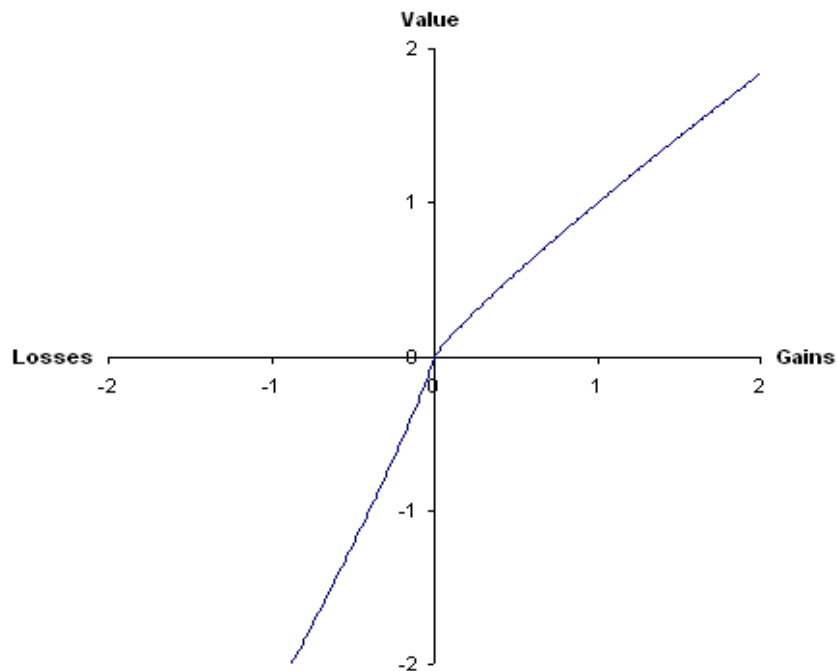
$$\pi(0.25)v(6000) < \pi(0.25)[v(4000) + v(2000)] \text{ and}$$

$$\pi(0.25)v(-6000) > \pi(0.25)[v(-4000) + v(-2000)].$$

Since the decision weights are the same on each side of the equations we can remove them to simplify, and we get, $v(6000) < v(4000) + v(2000)$ and $v(-6000) > v(-4000) + v(-2000)$. So now we can see from this question that peoples preferences changes towards more risk seeking if talking about losses rather than gains. This fits the hypothesis that the value function is convex for losses and concave for gains. This is also consistent with risk aversion

from expected utility theory which says that losses loom larger than gains.

Figure 1: A hypothetical value function



<http://prospect-theory.behaviouralfinance.net/> February 20, 2012

It should be mentioned however, as Kahneman and Tversky (1979) points out, that when discussing the utility function for money, we must leave room for that some special circumstances can have an effect on people's preferences. If an individual for instance should need \$100.000 in order to purchase a house, we might witness a very steep rise in preference near this critical value. The same can also be shown in the realm of losses. So because of these special circumstances, an individual's value function will not always reflect "perfect" attitudes to money. Such circumstances can potentially produce convex regions for gains and concave regions for losses in the value function (Kahneman & Tversky, 1979).

2.6 Local thinking

Nicola Gennaioli and Andrei Shleifer presents a model of intuitive inference called "local thinking" (2008). This is a memory based model of probabilistic inference as a continuation and improvement of Khaneman and Tversky's Prospect Theory from 1979.

The model is based on a heuristic belief that people evaluate hypothesis quickly based on

what first comes to mind. The quick and intuitive inference is called “local thinking” and is based on the idea that only some decision relevant data come to mind initially.

According to this model individuals evaluate the likelihood of a hypothesis based on some partial evidence. When evaluating, the decision maker fills in from the memory what is missing and completes the “picture”. This process of filling in from the memory the missing details is here called “framing”.

Gennaioli and Shleifer developed two assumptions for judgment under uncertainty (Gennaioli & Shleifer, 2008, p. 2):

1. One assumes that frames come to mind in order of their ability to predict the hypothesis being evaluated relative to other hypothesis. There also has to be a clear distinction between diagnosticity and the relevant frequency of frames.
2. Second, we assume that agents have limited memory and not all potentially relevant information comes to mind. This assumption is essential because if this is not the case the decision making would be entirely Bayesian and recalling missing data wouldn't matter.

Central results depend on the difference between the diagnosticity and likelihood of the frame. In most cases, the most diagnostic frames are also the most likely ones. The local thinker mostly makes modest judgment errors in this case. When on the other hand there is a mismatch between likelihood and diagnosticity of frames the local thinker probability assessment becomes very inaccurate. This mismatch can lead to underestimation and substantial biases. This model can also account for conjunction and disjunction fallacies. To explain this more closely I want to use an experiment Kahneman and Tversky famously used. They described a young woman, Linda as an activist in college, and asked their panel about the relative likelihood of her various activities today. People that were asked found it more probable that Linda was a bank teller and a feminist, rather than just a bank teller. This is an interesting result because there are surely bank tellers that are not feminists.

The “local thinker” model shows that the conjunction fallacy can be explained if the details of “Linda” are filled in differently by the local thinker depending on what data he/she are given

in the first place. If in example a former bank teller is represented with a diagnostic but very unlikely frame of a political moderate, the local thinker can assume that there are fewer people like this than there are formerly activists who are now feminist bank tellers (Gennaioli & Shleifer, 2008, p. 4).

The model also accounts for anomalies related to demand for insurance. It is illustrated that the local thinker is willing to pay more for insurance against specific risk than the rational thinker. But when presented the insurance that covers any risk they are willing to pay the same. This is because for the local thinker, only one risk comes to mind.

The point of this model is that, when making quick decisions, people do not consider everything they know and think. Only some information is recalled from the passive memory and this information is not always even the most useful.

2.7 Salience theory of choice under risk

Gennaioli, Shleifer and Bordalo present “a theory of choice among lotteries in which the decision maker’s attention is drawn to salient payoffs. In this case the payoffs that draw the decision maker’s attention are “salient”. This is a new psychologically founded model of choice under risk, which exhibits the systematic instability of risk preferences and accounts for the puzzles” (Bordalo, Gennaioli & Shleifer, 2010, p. 1). True probabilities are replaced by decision weights distorted in favor of salient payoffs. This model specifies decision weights as a function of payoffs and thereby provides a new and unified account for frequent risk seeking behavior, invariance failures like the Allais paradox, and preference reversals. There are also some new predictions which distinguish this model from prospect theory.

Important violations of expected utility have shown that attitudes toward risks are unstable. At the basic level, people exhibit both risk loving and risk averse behavior depending on the situation. People also participate in unfair games, pick highly risky occupations over safe ones and invest without diversification, while simultaneously buying insurance. This systematic instability support several paradoxes of choice under risk.

The results of this model rely on three assumptions; the first two are called *ordering* and *diminishing sensitivity*, and formalize the salience of payoffs. Salient payoffs are very

different in percentage than other payoffs in the same state or world. This captures the idea that we focus on differences rather than values, and that we perceive changes on a logarithmical scale (Weber's law). The third assumption states that "the extent to which decision weights are distorted depends on the salience of the associated payoffs, and not on the underlying probabilities" (Bordalo et al., 2010, p. 2). Under these assumptions the model describes how the decision maker develops a context dependent representation of each lottery.

In many ways this approach is similar to the one pursued by Gennaioli and Shleifer with their "Local thinking". Both studies have the idea that decision makers do not fully take into account all the information available to them, but rather over emphasize the information their mind focus on. Local thinkers neglect potentially important but unrepresentative data, while in this theory the decision makers analogously overweight states that draw their attention and neglect states that do not.

The strongest deviation from expected utility theory in this model occurs in the presence of extreme payoffs, and especially if they occur with a low probability. This property leads to an explanation of the Allais paradox which shows an inconsistency of actual observed choices with the predictions of expected utility theory. In sync with the definition of salience the model predicts that the subjects in the Allais experiments are risk loving when the common consequence is small and attention is drawn to the highest lottery payoffs, and risk averse when the common consequence is large and attention is drawn to the lowest payoffs.

Like Prospect Theory, this model assumes that decision makers focus on payoffs rather than the absolute wealth when choosing amongst risky alternatives. Prospect Theory also assumes that the probability weight used to make choices is different from objective probabilities, while salience theory believes that these weights depend on the actual payoffs and their salience.

Both of these models presented above explore how the limitations of the mind cause people to focus their attention on some but not all aspects of the world, which is called local thinking. Salience theory argue that salience shape this focus. In the case of choice under risk we can say that the contrast between payoffs shapes the salience, and people overrate the decision weights associated with the salient payoff. Decision makers overweight salient payoffs, and when the salient payoff is the upside of a risky choice they are risk seeking, and on the other

hand if the salient payoff is the downside the decision maker behave in a risk averse way.

Other aspects of salience have been used by economists to examine the consequences of people overrating salient data. For example Barber and Odean (2008) find that stock traders respond to “attention grabbing” news. In their paper “*All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors*” they find that individual investors respond to important news about firms by buying stocks. This news also affects the institutional investors but their respond is both buying and selling. Barber and Odean propose that investors manage their problem of choosing stocks by limit the selection to the ones that recently has grabbed their attention. Thus, making the investors more likely to buy salient stocks.

2.8 Risk attitudes

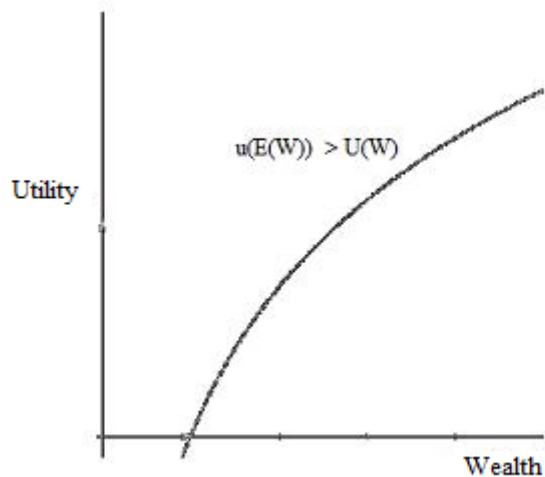
So far, this chapter has focused on fear and ambiguity, and little on risk. The main reason for this is that when considering the stock markets, ambiguity and fear is the main concern, and this is what we wish to study. As mentioned in the beginning of this chapter, the main difference between uncertainty and risk is that when dealing with risk, the outcomes and probabilities are known, while when dealing with ambiguity the outcomes and/or probabilities are unknown (Ackert & Deaves 2010). All possible outcomes and probabilities are rarely known in stock markets, therefore is uncertainty more interesting to look at in this thesis. However, risk attitudes are somewhat interesting to this thesis because they tell us something about how “brave” people generally are when investing. We will therefore include this part on risk attitudes.

From behavioral finance we can divide people into three different risk attitudes; *risk averse*, *risk neutral* and *risk loving*. People are divided into one of these categories in regards to if they are willing to accept a fair game or not, given that the fair game is of some size. Snyder and Nicholson (2008) offer this definition to a fair game “A “fair game” is a random game with a specific set of prizes and associated probabilities that has an expected value of zero.” (Snyder & Nicholson, 2008, p. 203). An example of a fair game could be a coin toss. In a coin toss there are two possible outcomes and a 50% chance of each outcome, so if you have a number of rounds in this game, your expected value would be zero.

A risk averse person would not be willing to take a gamble on a fair game. For this person there would have to be some additional prize for him to be willing to take the bet, assuming that the cost of the bet is of some size (even a risk averse person would possibly participate in a fair game just for fun as long as the stakes were low enough). This is due to the diminishing marginal utility of wealth. Diminishing marginal utility of wealth means that if you already have a fair portion of wealth gaining more of it would mean increasingly less to you, thereby causing a potential loss of \$1000 to loom larger than a potential gain of \$1000. A risk averse person will therefore have a concave utility function, which means that he will prefer the utility of the expected value of a prospect to the expected utility of the prospect. In equation form:

$$u(E(W)) > U(W)$$

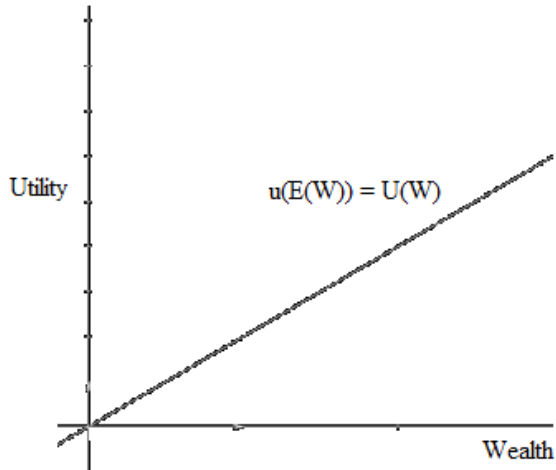
Figure 2: Utility function of a risk averse individual



A risk neutral person will rank his options after the expected value, and choose accordingly. The risk neutral person will therefore have a linear utility function, which means that this person will be indifferent between the utility of the expected value of a prospect and the expected utility of the prospect. In equation form:

$$u(E(W)) = U(W)$$

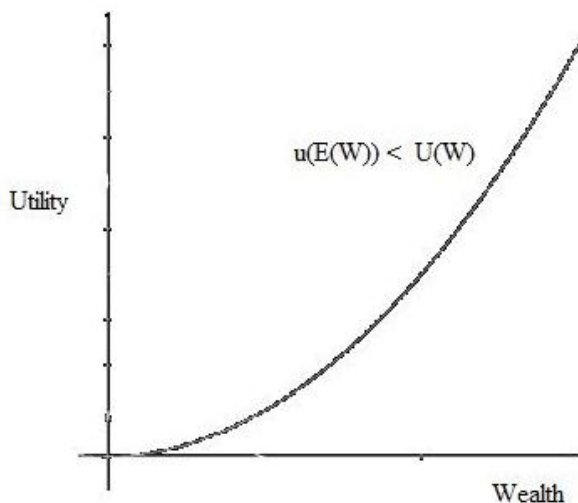
Figure 3: Utility function of a risk neutral individual



Although rare, some people are risk lovers, and will always prefer a fair game to not participating. People with these kinds of preferences have a convex utility function, meaning that the expected utility of a prospect will be preferred to the utility of the expected value of a prospect (Ackert & Deaves, 2010). In equation form:

$$u(E(W)) < U(W)$$

Figure 4: Utility function of a risk loving individual.



2.9 Investor sentiment

In theoretical finance you learn that stock prices, and the prices of other securities, are set by unemotional investors who set the price to the present value of expected future cash flows. In the real world however, you find that stock prices often are more affected by trends and

investor sentiment, rather than a valuation of fundamentals. As Baker and Wurgler (2007) puts it: “Now, the question is no longer, as it was a few decades ago, whether investor sentiment affects stock prices, but rather how to measure investor sentiment and quantify its effects”.

As long as there have been financial markets there have been events so shocking that the traditional financial (valuation) models have had no way of explaining them. The great crash of 1929, the black Monday crash of October 1987, and the more recent Subprime mortgage crisis (financial crisis) of 2008. These crashes are not a new phenomenon, but something that have existed for a long time. In fact, to find what is considered to be the first “bubble” we need to go back to 1637 and the Dutch “Tulip mania”, where the tulip became so popular that at the most extreme a single tulip bulb had a value of more than twenty times the annual income of a skilled craftsman (<http://penelope.uchicago.edu>). While this last example may seem extreme we don't need to go further back in history than to the years 1995 through 2000 when the so called “Dot.com” bubble were building up in stock markets, particularly in the US.

All these events have one thing in common; they are the result of human emotions such as greed, fear, and ambiguity. We can safely claim that during the last portion of the build up, and during the crashes in these events, valuation by “fundamentals” was not present. Baker and Wurgler (2007) find that different stocks are differently affected by investor sentiment. They find that larger stocks that are paying dividend on a regular basis are less affected, whilst small companies who are, amongst other things, more volatile, non-profitable and have an extreme growth potential are more affected by investor sentiment. Baker and Wurgler (2007) measure investor sentiment by constructing an index consisting of five proxies; trading volume as measured by NYSE turnover, the number and first-day returns on IPOs, the dividend premium, the closed-end fund discount, and the equity share in new issues.

Their findings show that when the investor sentiment is low, the average future returns on the smaller companies are higher than the larger ones. On the other hand, when the investor sentiment is high, the average future returns on the smaller companies are lower than the larger companies. These findings are inconsistent with classical asset pricing models, in which smaller, more risky, companies should have a higher expected return in order to make it worthwhile for the investor to take the risk.

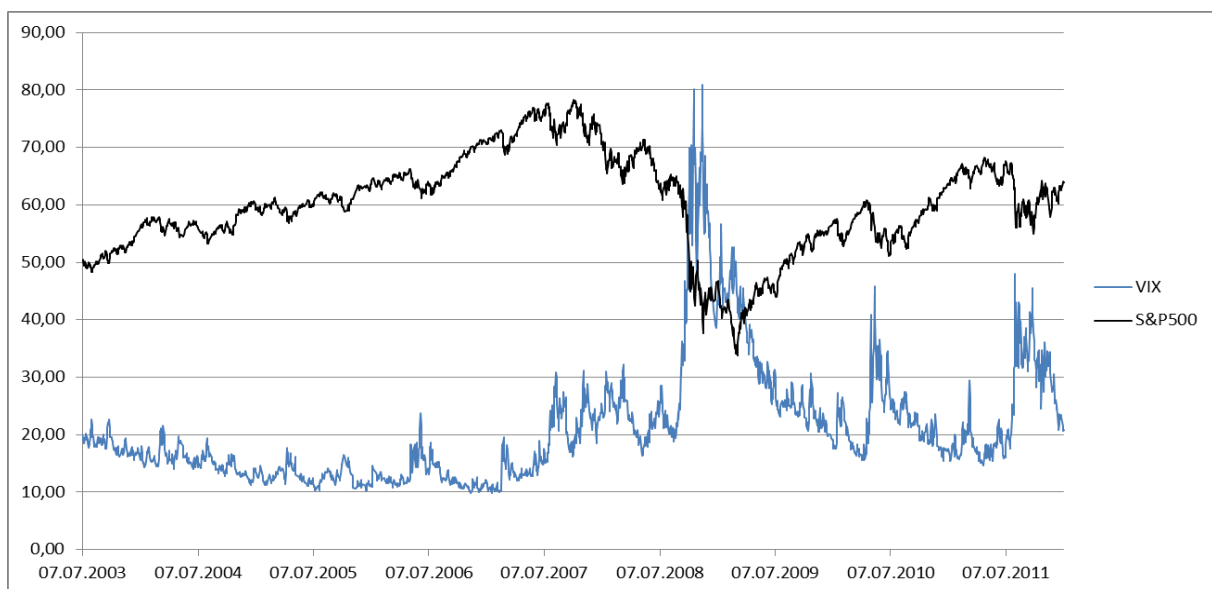
So if investor sentiment plays such a huge role in financial markets, how can we measure it? For the next part of this thesis we will attempt to measure what we find to be the most interesting parts of investor sentiment, namely fear and uncertainty.

3.0 The Volatility Index

The Volatility index (VIX) was developed in 1993 by Professor Robert E. Whaley and was introduced in the market by the Chicago Board Options Exchange (CBOE). The VIX index was originally designed to measure how the market expected the 30-day volatility to change, implied by at-the-money S&P 100 index option prices. In addition to this one of its purposes was to provide investors with an index upon which options and futures contracts on volatility could be bought and sold. The VIX index soon became the preferred benchmark for measuring stock market volatility in the U.S. and it is often referred to as the “fear index”.

In 2003 the VIX index was updated, by the CBOE together with Goldman Sachs, in order to reflect a new way of measuring expected volatility. This updated VIX index is based on the S&P 500 index, which is considered to be the core index for U.S. equities. Here the expected volatility is measured by taking an average of the weighted prices of the S&P 500 puts and calls over a wide range of strike prices. The idea is that the price of each option reflects what the market is expecting in terms of future volatility. The graph below this paragraph show how the VIX index and the S&P 500 moves according to each other, the S&P 500 data that went into this graph have been divided by a 20 to make it more comparable to the VIX in a graph. Following this paragraph we offer a more detailed description of how the VIX index is constructed. (The CBOE Volatility Index – VIX, 2009)

Figure 5: Graph showing the relation between the VIX index and the S&P 500



3.1 Index options

Before presenting the Norwegian volatility index we will explain some of the most important components and the construction process of both the VIX and Norwegian VIX. Like the VIX index, our Norwegian volatility index is based on index options.

An option is defined as the right to buy or sell an asset, where an asset can be a stock, currency, commodities, etc., an option to sell an asset is referred to as a “put option”, while an option to buy an asset is referred to as a “call option”. An index option is an option contract on a stock or some other index. In our thesis the index options used are on a stock index, specifically the OBX index, these index options are both puts and calls with various strike prices. A stock index is an index monitoring the value of a portfolio of stocks.

3.2 Construction of the Volatility Index

The VIX index consists of near- and next-term out-of-the-money put and call options, which are usually in the first and second S&P 500 contract months. An option with a bid price of zero will not be included in the calculation. Options that are near-term must be at least one week away from expiration, due to the pricing anomalies which might occur when an option is close to expiration. In order to have a high precision in pricing, the VIX calculation measures the time to expiration for each day in minutes. The risk free interest rate used in the calculation is the yield of the U.S. T-bill maturing closest to the expiration date of the relevant S&P 500 options.

When the S&P 500 options are sorted the VIX calculation can begin. The first step is to determine the forward S&P 500 level for each contract month. This is done by sorting the options by their strike prices and then finding the strike price where the absolute difference between the call and put is at the lowest. When this strike price is found it can be inserted into this forward index level formula:

$$F = \text{Strike Price} + e^{RT} * (\text{Call Price} - \text{Put Price})$$

This calculation is done for both the near- and next-term options. Now that the forward prices are found, the strike price which will be used can be determined by finding the strike price

immediately under the forward index level. When this strike price is found, the options to be included in the calculation can be found by selecting the out-of-the-money *put* options with strike prices less than the found strike price, and moving downwards to successively lower strike prices. Then the out-of-the-money *call* options with strike prices higher than the found strike price is selected, and moving upwards to successively higher strike prices. The choosing of options with lower and higher strike prices, in relation to the found strike price, stops when hitting two consecutive zero bids.

Now that the options to be included in the calculation have been found and selected, their prices are determined by using mid-quote prices, which is an average between the bid and ask price. When these prices are found they can be used to calculate the index values using this generalized formula:

$$\sigma^2 = 2/T \sum \Delta K_i / K_i^2 * e^{RT} * Q(K_i) - 1/T [F/K_0 - 1]^2$$

where:

$$\sigma = \text{VIX}/100 \rightarrow \text{VIX} = \sigma * 100$$

T = Time to expiration

F = Forward index level derived from index option prices

K_0 = First strike below the forward index level, F

K_i = Strike price of i^{th} out-of-the-money option; a call if $K_i > K_0$ and a put if $K_i < K_0$; both put and call if $K_i = K_0$

ΔK_i = Interval between strike prices – half the difference between the strike on either side of K_i :

$$\Delta K_i = (K_{i+1} - K_{i-1})/2$$

(Note: ΔK for the lowest strike is simply the difference between the lowest strike and the next higher strike. Likewise, ΔK for the highest strike is the difference between the highest strike and the next lower strike.)

R = Risk-free interest rate to expiration

$Q(K_i)$ = The midpoint of the bid-ask spread for each option with strike K_i

After the VIX calculation formula is applied for both the near- and next term option, the 30-day weighted average of σ_1^2 and σ_2^2 (near- and next term option, respectively) is calculated. Then, all that remains in order to get the VIX value for this day is to take the square root of

this value and multiply it with 100. (The CBOE Volatility Index – VIX, 2009, pp. 3-9)

$$\text{VIX} = 100 * \sqrt{\{T_1\sigma_1^2[(N_{T_2} - N_{30})/(N_{T_2} - N_{T_1})] + T_2\sigma_2^2[(N_{30} - N_{T_1})/(N_{T_2} - N_{T_1})]\} * N_{365}/N_{30}}$$

where:

N_{T_1} = number of minutes to settlement of the near term options

N_{T_2} = number of minutes to settlement of the next term options

N_{30} = number of minutes in 30 days

N_{365} = number of minutes in a 365-day year

3.3 The fear index

The VIX index is designed to measure volatility and not fear in particular. Technically volatility means unexpected moves either up or down. So why is the VIX index often referred to as the fear index? The main reason for this is that the S&P 500 index option market has become dominated by hedgers. According to John C. Hull (2011) the objective for hedgers is to use futures markets to reduce their risk against the underlying as much as possible. On the S&P 500 index option market hedgers achieve this by buying index puts which is the right to sell the index option at a certain predetermined price. The demand for these index puts is higher when the hedgers are concerned about a potential drop in the stock market. The higher the demand for index puts, the higher the price. This higher price is reflected by a higher VIX index, so one could say that the VIX index is an indicator of the demand for portfolio insurance.

This asymmetry in trading put index option as opposed to call index options should lead to that the change in the VIX will rise at a higher absolute rate during a stock market fall as opposed to when it rises. The man who developed the VIX, Professor Robert E. Whaley (2008), tests this by regressing the daily rate of change of the VIX, the rate of change of the S&P 500 portfolio, and a dummy variable measuring the change of the S&P 500 portfolio if the market is going down and 0 otherwise. The results are an R-squared of 55.7% and with an exception of the intercept all the regression coefficients are significant at a 1% level. His results show both an inverse relation in movements in the VIX and the S&P 500, and an asymmetry of movements due to portfolio insurance. His coefficients say that if the S&P 500 rises by 100 basis points, the VIX will fall 2.99%. Should the S&P 500 fall by 100 basis points Whaley's results say that the VIX-index will rise by 4.493%.

Because of this asymmetry of movements due to the demand for portfolio insurance, we could say that although the VIX index is designed as a measure of future volatility, it can also be used to measure the amount of fear in the stock market.

So at what VIX level would we say that there is fear in the market? One way to get some sense of this is to characterize which levels are normal and abnormal. Whaley does this in his paper, he considers a sample period from January 1986 through October 2008 and calculates the median for each year and for the period as a whole. He finds that over the whole period the median is 18.88. Whaley (2008) also calculates at which value the VIX is 5% likely to occur. His finding for the whole period is that this value is 34.22. Since this value or higher only occurs 5% of the time over the whole period, it can safely be considered as rare or abnormal. So when the VIX remains above this level for a number of days one could say that the market is nervous or even afraid. Looking back in history, Whaley finds four periods in which the VIX index had more than 20 consecutive days above the 34.22 value: October 16 through December 22, 1987 (47 days), January 8 through February 8, 1988 (22 days), August 28 through October 31, 2002 (46 days), and September 26 through April 17, 2008/2009 (140 days). This shows that the anxiety in the markets is high when we experience large downward market movements. We especially see this during the financial crisis in 2008 when the VIX was above 34.22 for 140 days straight.

3.4 Critique of VIX

Whether the volatility implied in option prices is our best estimate of future volatility is not a subject in which everyone agrees upon. Philippe Jorion (1995) in his article “Predicting Volatility in the Foreign Exchange Market” points out some shortcomings in the S&P 100 options (which the VIX was based on at the time). Jorion (1995), among other things, points out that because of the transaction costs in going long or short on all the stocks in the index simultaneously, the arbitrage between the option and the underlying asset would be hard to implement.

An earlier study by Lamoureux and Lastrapes (1993) measures prices and matching the forecast horizon on ten individual stock options at the New York Stock Exchange. Their findings in this study are that historical time-series has more predictive power when compared to implied volatility. However, in their study, Lamoureux and Lastrapes (1993) also find that

when forecasting in a 90- to 180-calender-day horizon the price process of the underlying stock contains useful information which does not exist in the historical data.

However, in a more recent study done by Ferris, Kim and Park (2008) on daily S&P 100 and S&P 500 index option prices from January 2000 through September 2006, the conclusion is that implied volatility is a good and unbiased estimator of future volatility. Their findings also suggest that implied volatility measures obtained from constant volatility models, such as the Black and Scholes, contain valuable information about future volatility.

Other studies on the subject have found that volatility is a term that can easily be confused or misinterpreted. Goldstein and Taleb (2007) show this in an experiment, where they ask three groups, consisting of different finance professionals and Ivy League graduate students, a question designed to see if they can distinguish between mean absolute deviation and standard deviation. The groups are asked to calculate daily and yearly sigma from the question. The results are that as few as 3 of 87 of the respondents managed to calculate the correct daily sigma and none managed to calculate the correct yearly sigma. The reason for these results is that the respondents were given data in the question that were in mean absolute deviation, but they treated the data as if it was standard deviations. The mean absolute deviation is about 0.8 times the standard deviation so a confusion of the two terms can lead to large differences when doing calculations. Goldstein and Taleb (2007) also mention reports about financial journalists who make the same mistake when explaining the VIX index to the general public.

Some of the more recent criticisms of the VIX are that it is too simplistic to estimate future volatility for all stocks. The reason for claiming this is that different stocks have different volatility. Technology stocks typically have higher volatility than utility stocks, and small stocks tend to have higher volatility than larger ones. Though this is true it is important to remember that the VIX does not measure the volatility for each stock on the S&P 500, but rather on the index as a whole.

4.0 The Norwegian volatility index

Now that we have given a presentation of the U.S. volatility index based on S&P 500 index, we will present our volatility index based on the Norwegian OBX stock index. We will start by giving a presentation of how the process of making the Norwegian volatility index was from beginning to end, before presenting the Norwegian volatility index and performing some analysis.

4.1 OBX index

The index option data we have for this thesis is for the OBX index. The OBX consists of the 25 companies in the main index (OSEBX) with the highest turnovers. The decision for which companies that should be included or excluded from this index is done once every six months. Therefore some of the companies in this index will have changed over the period in which we are considering. Even though the OBX index consist of only 25 companies, most of these companies are quite large and they in fact make up around 90% of the main index on Oslo stock exchange, at the present date. As per 17.04.2012 the OBX index consist of the following companies (company ticker symbol in the first column and their weights on the OBX in the second column):

Table 2: Companies and weighting on the OBX index

PGS	2,519
RCL	1,6468
ORK	5,2934
SDRL	6,4487
AKSO	2,1849
SFR	0,779
PRS	1,4654
TGS	2,407
FRO	0,1636
STB	1,3872
STL	25,2249
TEL	12,2456
CEQ	0,6039
YAR	7,6448
FOE	1,1078
NHY	4,0757
DNO	0,8946
MHG	1,1455

SCH	2,6145
GOL	1,12
GJF	2,0438
SUBC	6,0139
DNB	10,0827
ALGETA	0,6258
REC	0,2615

<http://www.newsweb.no/newsweb/search.do?messageId=302703> April 7, 2012

4.2 Collecting the data

What we want to investigate in this thesis is if there are some way in which we can measure the degree of fear and ambiguity that exists in the Norwegian stock market. In our minds the best way to accomplish this is to make a volatility index for the Oslo stock exchange similar to the volatility index which currently exists in the US for the S&P 500. When we set out doing this task we wasn't quite sure whether we would get the data we needed or of it even could be done with the time and resources at hand. Even so we wanted to give it our best shot, and began the process of retrieving the data.

We first started investigating whether there already existed an Norwegian volatility index or something similar, to do this we used a program called Datastream which is found on the university library computers. When using Datastream we encountered a few difficulties, the first being that since this program was brand new there was no one in the library or university staff who knew how to operate it properly. After this problem was solved by reading a very long manual, we discovered that there was an error in the program which did that only one computer could retrieve data from the program at a time. This error made the whole process a bit more tedious, but eventually we found a solution to the problem. When this was done we found that there is no Norwegian volatility index. We also found that Datastream did not have the data we needed to make one, and therefore decided to contact the Oslo stock exchange directly.

Linn Furuvald at Oslo stock exchange was very helpful and provided us with the data for all the index option trades done on the OBX index over the past fifteen years. This is quite a lot of data and consists of about 300.000 rows in an excel spreadsheet. This raw data we received needed quite a lot of sorting before we could start to work on them. Most of this sorting could be done in excel, but some elements had to be taken out manually, which takes quite some

time when you have 300.000 rows of data. Eventually, when the sorting was done we could set about trying to make our own Norwegian volatility index.

We started by visiting the Chicago Board Options Exchange website (www.cboe.com) since this is the company that has produced the volatility index. On their website we found a paper called “VIX White” which essentially is a recipe for how the volatility index is made. After fine reading this recipe several times we set about trying to make our own volatility index. After some time we first accomplished making an index value for one day, and after this we made index values for a couple more days to confirm that we were doing the right calculations. Our main problem at this point was that we were working in excel which have some limitations, because of these limitations, calculating the index value for just one single day took roughly one hour. Since the calculations in excel had such limitations we were advised to use a program called SAS (Statistical Analysis System) in calculating the rest of the index. This software is not available on the university, but we were fortunate enough to get access to it on Helga Lien’s workplace.

None of us had ever used or even seen SAS being used before, so this naturally represented a significant challenge for us. We spent quite a long time figuring out how we were going to use this software. However, after a lot of manual reading and trying and failing, we were lucky enough to get in touch with Brage Refve Vik and Thomas Tardy who had experience in using the software, and they were kind enough to help us get started. Although our Norwegian volatility index may not seem like a lot of work on paper, this is not the case. From the time when we received the first batch of index options data, to the point where we finally managed to make our Norwegian volatility index went more than two months of hard work. Firstly, a good portion of this time was first spent on figuring out where to begin when making the index, and making the first “prototype” in excel. Secondly, a lot of time was spent on learning how to use SAS which, as mentioned before, was completely new to us. It is, however, worth all the hard work now that we have our very own Norwegian volatility index, from here on referred to as the NVIX, which we will present to you.

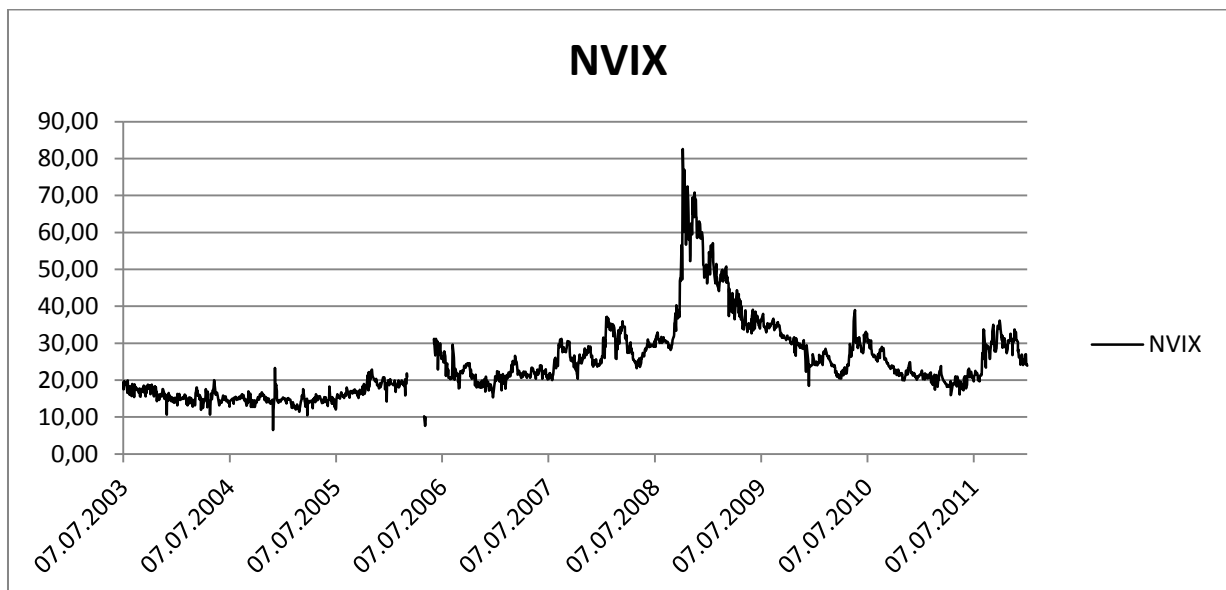
4.3 NVIX

The index option data received from Oslo stock exchange have now been sorted and calculated in line with the CBOE Volatility Index method, and the outcome is a Norwegian

volatility index based on OBX index options.

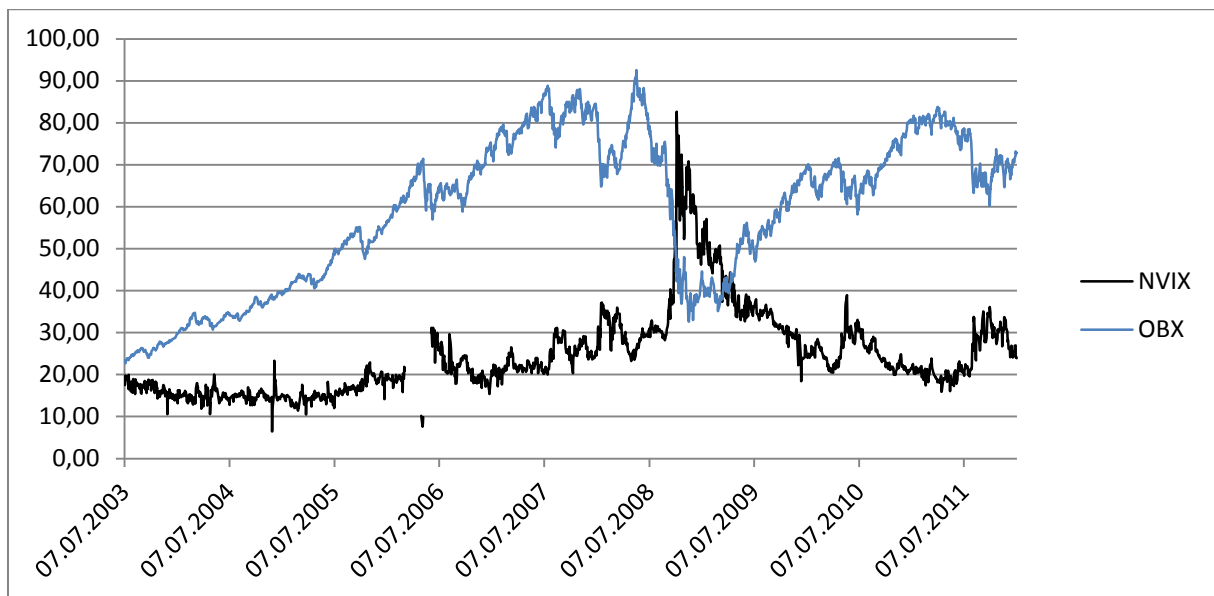
The only part in the construction where our NVIX have a different construction than the VIX is when it comes to the risk free rate. Originally, the risk free rate should be found by finding the yield of the U.S. T-bill maturing closest to the expiration date of the relevant S&P 500 options, as is done for the VIX. On this matter we had to make a simplification, finding this Norwegian risk free rate for each day would be very time consuming, and the data available was not detailed enough for this purpose. We therefore decided to have one fixed risk free rate throughout the whole index. The risk free rate we decided to use was the one given as an example in the “VIX White” (<http://www.cboe.com/micro/vix/vixwhite.pdf>), which was 0.38%. We decided on this rate since the U.S. rate and the Norwegian rate are relatively similar, and we figured since the rate is also relatively small due to the close expiration there would probably be only small differences across the period from 2003 to 2012. Before analyzing the data more thoroughly we present the NVIX, values in a graph.

Figure 6: Norwegian VIX



The gap just before July 2006 comes from some missing NVIX values in that period. The reason is that in order to calculate the NVIX you need both next and near option trading. In this period there were only near trading and we were unable to calculate values. The big spike in 2008 is as expected due to the credit crunch. When comparing NVIX and OBX in figure 7 we observe that it looks like there is a negative correlation between the two indices. The OBX data is divided by five for better comparison to the NVIX in the graph.

Figure 7: NVIX, OBX



4.4 Econometric NVIX analyses

We want to test if the Norwegian VIX has any predictive power to OBX and investor sentiment, and to explain any possible connection between these variables. This is done by performing a regression and a correlation analysis in Excel and evaluating the outcomes. We use time series data from 07.07.2003-06.01.2012. Originally we received index option data from Oslo stock exchange from year 1997, but because of an index change done in year 2000 we found it better to use the data from then as it was clearer to us what the signs meant. The index option data had different signs telling us whether it was a call or put, which month it expired and at which month it started trading.

After turning the option data into a volatility index in SAS it turned out to be very poor data from year 2000-2003 due to very little trading, therefore we decided to leave this time period out of our analysis. Data from 2003-2006 also had some parts with low trading but we chose to keep them for the sake of explanatory power of time. In the Analysis you will see that we look at both the whole period 2003-2012 and the period with the most stable data 2007-2012 separately.

4.5 Macroeconomic factors

In order not to over evaluate the connection between OBX and our Norwegian VIX we want to include other variables in the model that correlate with the explained variable or have a significant connection to the explanatory variable. These other variables are listed below. All the data found are in daily values.

Interest

The interest is the Norwegian Inter Bank Offered Rate 3 month nominal. This is the money market rate that Norwegian banks are willing to lend each other money for. This rate is referred to as the NIBOR rate. Gjerde and Sættem (1999) find in their study that the change in this interest will have an impact on the Norwegian stock market. The data is found at www.norges-bank.no (NIBOR).

Oil price

It is natural to include this variable since Oslo stock exchange is very oil-based. Many of the biggest companies on the exchange are oil companies or oil service companies and rely on the oil price. Falling demand for oil will lower the price and thereby the earnings for these companies. Daily oil prices are found at Thomson Reuters (Brent_USD).

Exchange rates

The oil industry has a large impact on the Norwegian economy. Oil is normally traded in US dollar so we want to include the exchange rate between Norwegian kroner and US dollar. The rates were found at www.mataf.net/no (NOK_USD).

S&P 500

The Norwegian stock market is a small market that is largely affected by what happens in the rest of the world. The US is the biggest economy in the world and has a huge effect on the rest of the world, and especially the western world. The historical data was found at finance.yahoo.com. (S&P500)

US VIX

The US VIX is based on S&P500 index options and is a natural variable to include in our model. The method used to make this index is the same that we used to make the Norwegian VIX and it is therefore expected to have some connection or explanatory value to the OBX. Daily historical data for this index was also found at finance.yahoo.com. (VIX)

4.6 Correlation Analysis

Correlation between two variables measures the degree of linear context between those variables. The scale goes from -1 which means it is negatively correlated. If one variable goes up by one percent, the other goes down by one percent. A correlation value of 0, means that there is no correlation at all. If one variable moves by one percent the other variable is moving independently from this. Correlation of 1 means the two variables is perfectly correlated. Meaning, both variables move in the same direction at all time. Below we find a correlation matrix for the whole period 2003-2012 and one for the period with better and more consistent data 2007-2012.

Table 3: Correlation matrix 2003-2012

	<i>NVIX</i>	<i>OBX</i>	<i>VIX</i>	<i>NIBOR</i>	<i>NOK_USD</i>	<i>Brent_USD</i>	<i>S&P 500</i>
<i>NVIX</i>	1						
<i>OBX</i>	0,0611	1					
<i>VIX</i>	0,8907	-0,0983	1				
<i>NIBOR</i>	0,4519	0,4495	0,3399	1			
<i>NOK_USD</i>	0,0838	0,8043	0,0012	0,4779	1		
<i>Brent_USD</i>	0,1865	0,8139	0,1280	0,4225	0,8881	1	
<i>S&P 500</i>	-0,4123	0,7274	-0,5291	0,4152	0,5547	0,4407	1

Table 4: Correlation matrix 2007-2012

	<i>NVIX</i>	<i>OBX</i>	<i>VIX</i>	<i>NIBOR</i>	<i>NOK_USD</i>	<i>Brent_USD</i>	<i>S&P 500</i>
<i>NVIX</i>	1						
<i>OBX</i>	-0,8298	1					
<i>VIX</i>	0,8982	-0,7703	1				
<i>NIBOR</i>	0,2471	0,2271	0,1669	1			
<i>NOK_USD</i>	-0,5249	0,7174	-0,4764	0,3415	1		
<i>Brent_USD</i>	-0,4412	0,6104	-0,3069	0,2010	0,8363	1	
<i>S&P 500</i>	-0,6842	0,9293	-0,6739	0,4295	0,6383	0,4575	1

We can see that the correlation between OBX and NVIX change from almost no correlation 0,0611 in table 1 to negatively correlated -0,8298 in table 2. If the volatility index is working as we expect it to it should be negatively correlated to OBX. Another interesting observation is that the VIX and NVIX is highly correlated in both tables and has the same pattern over

time. We can also observe that S&P 500, Brent_USD and NOK_USD are all positively correlated to the OBX in both periods. This indicates that stock markets tend to move in the same direction and that the oil price and exchange rate react similar.

These high correlation values can result in multicollinearity. Multicollinearity is when the independent variables in a regression are highly correlated with each other and can lead to imprecise estimations.

4.8 Descriptive statistics

This matrix shows the descriptive statistics for use in the regression. We observe that the NVIX and the VIX are skewed to the right and have a high and sharp central peak by the kurtosis. OBX are skewed to the left and have a lower and smoother peak.

Table 5: Descriptive statistics

	Number	Minimum	Maximum	Mean	Std. deviation	Skewness	Kurtosis
NVIX	1997	6,480	82,613	24,927	10,261	1,808	4,545
OBX	1997	113,532	462,697	296,983	90,483	-0,357	-1,095
VIX	1997	9,890	80,860	21,507	10,557	2,084	5,561
NIBOR	1997	1,690	7,910	3,325	1,459	1,144	0,059
NOK_USD	1997	0,130	0,202	0,164	0,015	0,295	-0,640
Brent_USD	1997	25,510	143,950	71,032	27,181	0,386	-0,613
S&P 500	1997	676,530	1 565,150	1 200,734	172,854	-0,211	-0,161

4.9 Hypothesis

The main issue we want to test is: The NVIX can explain OBX. When we test hypothesis there is always one null hypothesis (H₀) and one alternative hypothesis (H₁). The null will then be that the NVIX will not affect or predict OBX and the alternative will be that it does.

H₀: $\beta_{NVIX} = 0 \rightarrow$ NVIX does not predict or have any effect on OBX volatility.

H₁: $\beta_{NVIX} \neq 0 \rightarrow$ NVIX does predict or affect OBX volatility

Similar procedure also accounts for the other variables.

Table 6: Hypothesis

	H ₀	H ₁
S&P500	$\beta_{S\&P500} = 0$	$\beta_{S\&P500} \neq 0$
Brent_USD	$\beta_{Brent_USD} = 0$	$\beta_{Brent_USD} \neq 0$
VIX	$\beta_{VIX} = 0$	$\beta_{VIX} \neq 0$
NVIX	$\beta_{NVIX} = 0$	$\beta_{NVIX} \neq 0$
NIBOR	$\beta_{NIBOR} = 0$	$\beta_{NIBOR} \neq 0$
NOK_USD	$\beta_{NOK_USD} = 0$	$\beta_{NOK_USD} \neq 0$

When testing hypotheses there are two things you can do wrong. This is if you discard H₀ when H₀ is true, or if you keep H₀ when H₀ is false. In order to avoid this problem we use significance levels that exclude very uncertain values. We will use both 1% and 5% significance level when testing the hypothesis.

4.10 Model specifications

The main goal with a regression is to explain the context between two or more variables. Our model is a multiple regression because we choose to include more than one explanatory variable.

$$\text{E.g. } Y_t = \alpha + \beta_1 X_{1t} + \dots + \beta_k X_{kt} + u_t \quad t=1, 2, \dots, T \text{ (time)}$$

Where Y_t is the dependent variable, and is explained by the constant α and the relation β_k to the explanatory variables X_k. The error term u_t captures all the other unobserved factors that affect the dependent variable. In order to estimate the coefficients α and β_k we use the OLS method. This model has T-k degrees of freedom, where T is the number of observations and k is the number of parameters in the model.

Model assumptions for Ordinary Least Square (OLS)

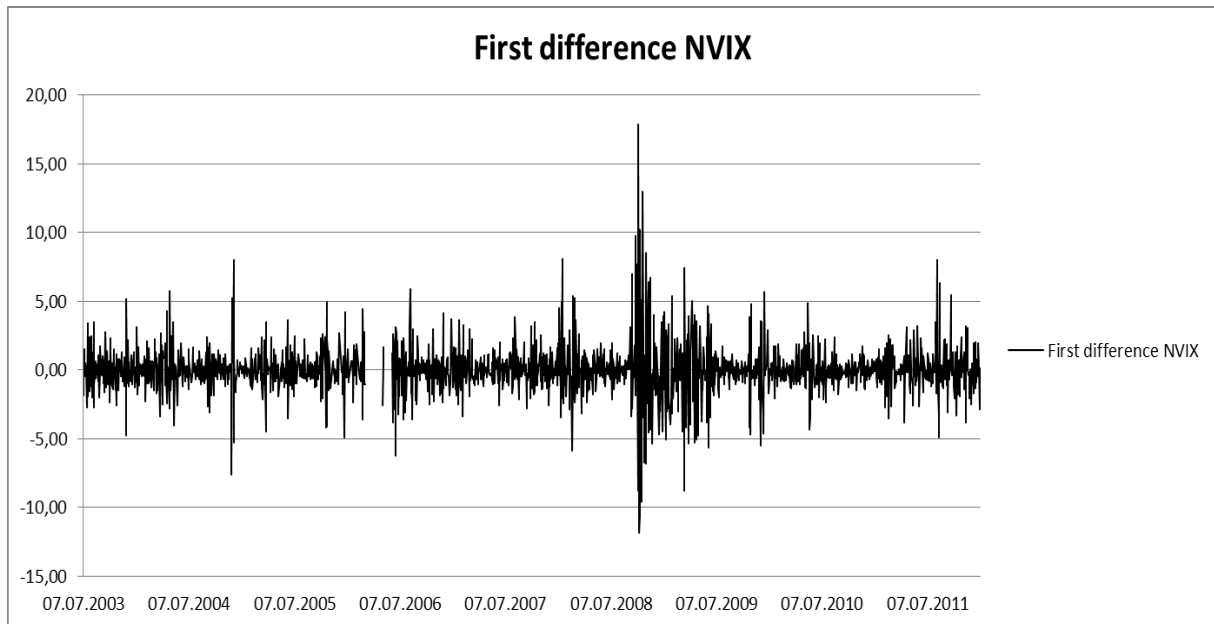
There are five assumptions to be met in order for the error term u_t:

1. The expected value of the error term is zero: $E(u_t) = 0$

2. Homoscedasticity: $\text{Var}(u_t) = \sigma^2 < \infty$

The variance of the error term is constant and finite for all values of X_k . If this is not the case the residuals will be heteroskedastic. A visual inspection of the NVIX first difference points to heteroskedasticity in the data, as can be seen in figure 8.

Figure 8: First difference NVIX



A solution to heteroskedasticity is to transform the data into logarithmic form. This will reduce the impact of extreme observations.

3. Avoiding autocorrelation: $\text{Cov}(u_i, u_j) \neq 0$

To avoid autocorrelation we need the error terms to be statistically independent of each other. Autocorrelation is a common problem when dealing with time series data. One can measure this by performing a Durbin-Watson test. One can reduce the likelihood of autocorrelation by differentiating the variables to percentage change.

4. Stochastic explanatory variables: $\text{Cov}(u_i, X_t) \neq 0$

With stochastic explanatory variables there is no connection between the error terms and the explanatory variables. Time series data can be tested for this by the Augmented Dickey-Fuller

test (ADF) (Stock & Watson, 2007). A time series with non-stationary data can be stationary if we differentiate the variables to percentage. But it is still important to test the data. As we can see in table 7 the t – values are all lower than the critical values and we don't have to worry about non – stationary data

Table 7: ADF test

Variable	p-value	t-value	Critical value 1% level	Critical value 5 % level
OBX	0,01	-12,8682	-3,47	-2,88
NVIX	0,01	-13,8412	-3,47	-2,88
VIX	0,01	-13,4086	-3,47	-2,88
NIBOR	0,01	-10,3424	-3,47	-2,88
NOK_USD	0,01	-4,1574	-3,47	-2,88
Brent_USD	0,01	-12,3704	-3,47	-2,88
S&P 500	0,01	-11,9234	-3,47	-2,88

5. Normally distributed error term: $u_t \sim N(0, \sigma^2)$

The error term has to be independent and normally distributed. When you put the error terms in to a histogram they should have the shape of a clock if the condition are met.

Summarized we use the first change logarithm of all the data when estimating the regression to reduce the likelihood of autocorrelation, multicollinearity, non- stationary and heteroskedasticity.

Measurements

When analyzing the outcome of a regression we look at the coefficient estimates and whether they are negative or positive. For all the coefficients there are stated a t-test which we can use to state if the variable is significantly different from zero. The p-value is the likelihood to achieve a result as least as extreme as the observed if H_0 is true. E.g. with a significance level of 5% we cannot reject H_0 if the p-value is larger than 0,05. For the regression as a whole we use the f-test. This measures all the variables at once unlike the t-test that test only one variable at the time.

R^2 tells us the regressions explanatory power by telling us how much of variation in the dependent variable is explained by the independent variables. The explanatory power is measured between 0 and 1 where 1 is the highest. Adjusted R^2 takes into account the loss of

degrees of freedom when adding one more explanatory variable.

4.11 Regression

The model is turned into a linear regression by using log form on all the variables, like so:

$$\ln(y_t) = \ln(\alpha) + \beta_1 \ln(X_t) + \ln(u_t)$$

Log-log model coefficients can be interpreted as elasticity measures. The coefficient value can then tell you how much the dependent value will change if the explanatory variable changes by 1%. As mentioned we also used the first difference values to avoid problems with heteroskedasticity and autocorrelation in the error term. This makes it possible to also see what a change in the dependent variable will do to the independent variable.

In our first model the OBX is the dependent variable and the other variables independent. We want to see which of the independent variables that is statistically significant. Δ stands for $\ln(\text{first difference})$.

$$\Delta \text{OBX}_t = \beta_0 + \beta_1 \Delta \text{NVIX}_t + \beta_2 \Delta \text{VIX}_t + \beta_3 \Delta \text{NIBOR}_t + \beta_4 \Delta \text{NOK_USD}_t + \beta_5 \Delta \text{Brent_USD}_t + \beta_6 \Delta \text{S\&P500}_t + u_t$$

Table 8: Regression 1, 2003-2012

<i>Significance-F</i>	<i>R-squared</i>	<i>Adjusted R-squared</i>
9,9814E-262	0,460027329	0,458397633

	<i>Coefficient</i>	<i>t-Stat</i>	<i>P-value</i>
Intercept	0,00029034	0,912574335	0,361577109
NVIX	-0,045541724	-10,64901211	8,60565E-26
VIX	-0,002268907	-0,484924905	0,627783043
Brent_USD	0,294753027	21,25575147	1,63106E-90
S&P 500	0,576241682	24,52258579	3,0343E-116
NOK_USD	0,081266787	3,946532497	8,20472E-05
NIBOR	0,025153701	1,144810268	0,252425568

Table 9: Regression 2, 2007-2012

<i>Significance-F</i>	R-squared	Adjusted R-squared
8,8687E-193	0,517048111	0,514720632

	<i>Coefficient</i>	<i>t-Stat</i>	<i>P-value</i>
Intercept	-0,000123712	-0,288031034	0,773370923
NVIX	-0,071214346	-8,821088281	3,75414E-18
VIX	-0,001788464	-0,311832713	0,755219874
Brent_USD	0,356959703	18,47683886	1,49175E-67
S&P 500	0,529571741	19,51383253	3,33953E-74
NOK_USD	0,117380339	4,150322588	3,54577E-05
NIBOR	0,021288669	0,801308639	0,42310589

R-squared and adjusted R-squared improve slightly from regression one to two, which mean that the second regression have a better explanatory power than the first. The critical value for the F-statistics at a 5% significance level is 2,10 and the critical value for a significance level at 1% is 2,80. Looking at both regressions above we see that both are well below the critical value and explain the variation of the dependent variable Y. When looking at each variable separately we see that NIBOR and VIX are below the critical values and we cannot reject the H_0 that these variables don't predict or have any effect on OBX volatility. The two sided critical t-value is 1,96 at a 5% significance level and 2,58 at a 1% significance level. NVIX and all the other variables are well above the critical t-value and we can reject H_0 . The coefficient values tell us how much OBX will change if one of the independent variables changes by 1%. We observe that NVIX has a negative relationship with OBX and they will move in opposite directions. Another observation is that OBX move closest with the price of oil and S&P500 as they have the largest coefficient values.

4.11.1 Regression with lag

Time series data consists of observation by the same variable over time. In time series there occur context between the independent variables in earlier time periods and the dependent variable. A lag is a time-delayed effect and could be that it takes some time for the stock market to react on a change in VIX value or the other way around. In an informed market we expect the variables to react immediately, and that it should not be necessary to consider lags when our data is daily and not by minute or hour. However we want to test this. After testing

the variables one day at the time we found that VIX had significance at its best when leading three days forward. This meaning in practice, that the other variables are lagged three days after VIX.

$$\Delta OBX_t = \beta_0 + \beta_1 \Delta NVIX_t + \beta_3 \Delta NIBOR_t + \beta_4 \Delta NOK_USD_t + \beta_5 \Delta Brent_USD_t + \beta_6 \Delta S\&P500_t + \beta_2 \Delta VIX_{t+3} + u_t$$

Table 10: Regression 3, 2003-2012

<i>Significance-F</i>	<i>R-squared</i>	<i>Adjusted R-squared</i>
2,7506E-280	0,483315437	0,48175367

	<i>Coefficient</i>	<i>t-Stat</i>	<i>P-value</i>
Intercept	0,000317185	1,017941229	0,308829937
NVIX	-0,040757145	-9,654283107	1,39361E-21
Brent_USD	0,283818039	20,78606392	5,49957E-87
S&P 500	0,603028576	26,00557361	1,4899E-128
NOK_USD	0,070538669	3,494271499	0,000485781
NIBOR	0,022851149	1,060472491	0,289058712
VIX	-0,043406474	-9,295609884	3,72865E-20

Table 11: Regression 4, 2007-2012

<i>Significance-F</i>	<i>R-squared</i>	<i>Adjusted R-squared</i>
5,0685E-201	0,532448927	0,530190226

	<i>Coefficient</i>	<i>t-Stat</i>	<i>P-value</i>
Intercept	-0,000113583	-0,268343537	0,788479459
NVIX	-0,062359383	-7,621176499	4,96816E-14
Brent_USD	0,341688985	17,80000076	2,62521E-63
S&P 500	0,561240966	20,60785907	2,13465E-81
NOK_USD	0,100505628	3,593108984	0,000339537
NIBOR	0,023379706	0,894243834	0,371364697
VIX	-0,036129381	-6,108769566	1,34107E-09

Regression three and four show us that when leading three days ahead VIX is also significant, and we can reject H_0 . At the same time we observe that adjusted R-squared is higher. Lagging and leading was tested at all the variables including the non-significant NIBOR with no change in the significance.

4.12 NVIX and fear

In his article “Understanding VIX” Whaley (2008) points out the fact that the VIX frequently spikes upwards and after each spike the VIX returns to more normal levels. The fact that VIX spikes during market turmoil makes it suitable as the “investor fear gauge”. Whaley (2008) argue and document that increased demand to buy index puts affect the level of VIX and that we should expect to find that the change in VIX rises at a higher absolute rate when the stock market falls than when it rises.

We want to test this proposition with a simple regression on our Norwegian VIX. Included in the test are the first difference NVIX, first difference OBX and a dummy variable that takes the value 1 if the market is going down and 0 otherwise, ΔOBX_t^- . Model two looks like this:

$$\Delta NVIX_t = \beta_0 + \beta_1 \Delta OBX_t + \beta_2 \Delta OBX_t^- + u_t$$

Table 12: Regression 5, NVIX relationship with OBX, 2003-2012

<i>Significance-F</i>	R-squared	Adjusted R-squared
0	0,086952651	0,086035937

	<i>Coefficients</i>	<i>t-Stat</i>	<i>P-value</i>
Intercept	-0,001956832	-0,887016882	0,375176889
OBX	-0,916338506	-6,028321976	1,969E-09
OBX ⁻ (dummy)	-0,415691906	-1,81274446	0,070021715

If the proposition is true, the intercept term should not be very different from zero and the slope coefficients should be significantly less than zero. The adjusted R-square in this regression is only at 8,6%. This indicates that we have left out important variables that explain the NVIX. The F-test is good and the regression in itself is significant. OBX is significant but OBX⁻ (dummy) is not significant from H_0 at 1% or 5% significance level but pass at 10%.

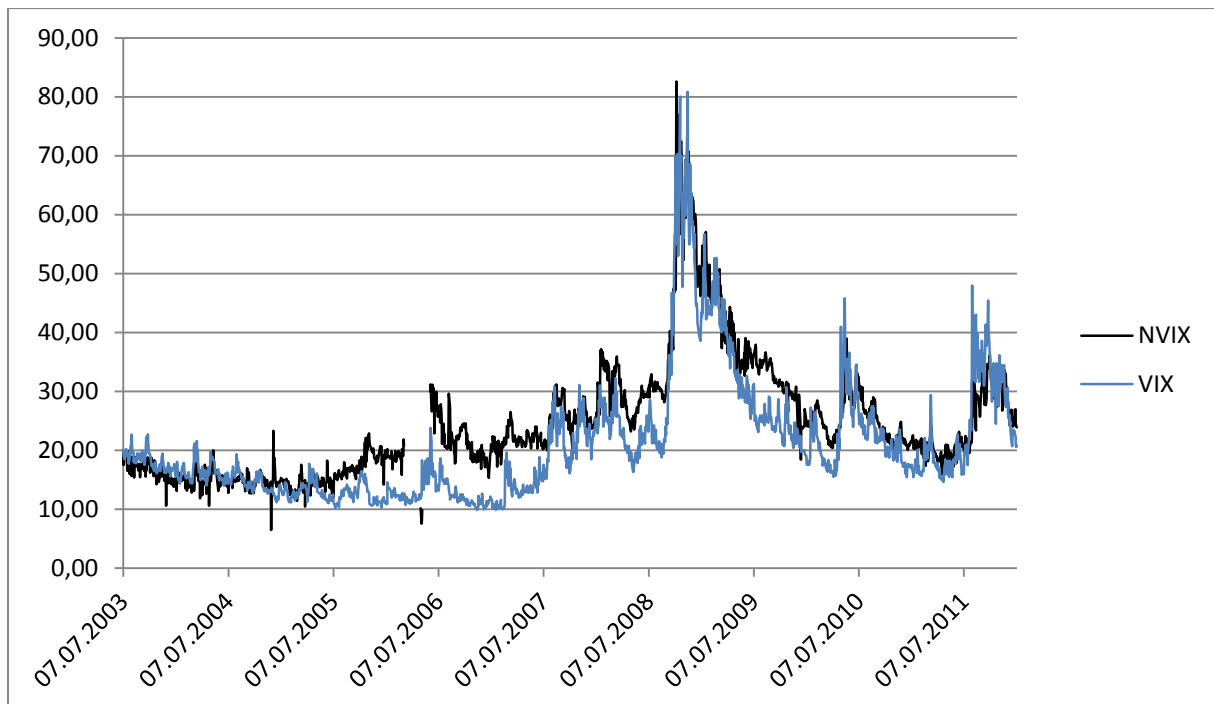
Thus if we allow lifting the significance level to 10% we find the prediction to be true. The fact that the intercept value is close to zero means that if the OBX does not change over the day, the rate of change in NVIX should be almost negligible. The estimated slope coefficients are both negative and reflect the inverse relation between movements in NVIX and OBX, and the asymmetry of the movements. We can interpret the coefficients as follows;

If the OBX rise by 1% the NVIX will fall by -0,916 %. $\rightarrow \Delta NVIX_t = -0,916(1) = -0,916 \%$

If the OBX falls by 1% the NVIX will rise by 1,332 %. $\rightarrow \Delta NVIX_t = -0,916(1) - 0,416(1) = 1,332 \%$

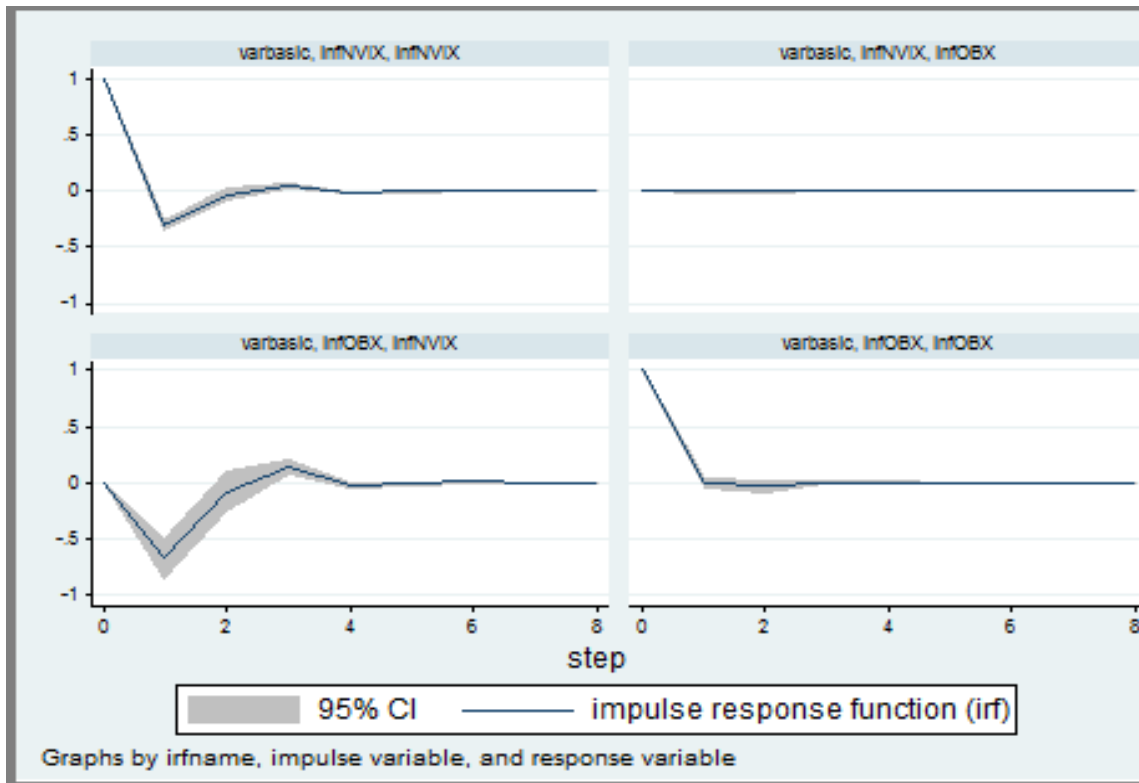
This shows that the NVIX is more a barometer of investor fear of the downside than it is a barometer of investor's greed in a market rally. Comparing our result to Whaley's result we notice a scale difference. He found that if S&P 500 rise by 1%, the VIX fall by 2,99 % and if S&P 500 fall by 1% VIX rise by 4,49%. This is an indication that the VIX spike more than the NVIX and react stronger on the underlying stock market. Looking at figure 9 you can visually see the difference between the NVIX and the VIX, and their reaction pattern.

Figure 9: VIX and NVIX



In addition to the reaction pattern shown above we want to investigate the impulse response NVIX has on OBX and the other way around. This is done in STATA and is based on a simple VAR regression. The impulse response graphs show the effects of shocks on the adjustment path of the variables. The data used is the NVIX and OBX 2003-2012 in log first difference.

Figure 10: Impulse response



The graphs in figure 10 tell us how a shock in the first variable affects the second variable. We are most interested in the second (northeast) graph InfNVIX, InfOBX, and the third (southwest) graph InfOBX, InfNVIX. The first of these shows how OBX respond to a shock in NVIX at a 95 % confidence interval. As we can see there is little or no movement in the OBX. The southwest graph on the other hand shows that a shock to OBX will affect NVIX significantly and the step axes tell us how long it takes before its back to “normal”.

4.13 The predictive power of NVIX

The VIX is, as we know, originally meant as a tool for predicting future volatility on the underlying stock market. If you have a NVIX value at 20, divide 20% by the square root of 12(months in a year) and you have the predicted volatility of 5,8% at OBX the next 30 days. By comparing NVIX prediction with OBX 30 day volatility we found a positive correlation at

30%. In the same time period (2007-2012) we found the VIX to predict the volatility on S&P 500 33% of the time.

4.14 Evaluation NVIX

The main reason for constructing the NVIX was to see whether we could measure fear and ambiguity in the Norwegian stock market. In our opinion we have accomplished this. When testing for asymmetry using Whaley's (2008) method we found a clear asymmetry indicating that the NVIX react stronger to downward movements on the OBX, as opposed to upward movements. It should however be mentioned that the adjusted R-squared for this particular regression was poor (only 0.086) and the OBX dummy variable was only significant at a 10% level.

The VIX is known as the "fear index" and is used by investors and economists as a benchmark of the market condition in the US. Our NVIX is constructed in almost an identical way, and the correlation analysis show that the VIX and NVIX have a correlation of 0.89 which is very close to perfect correlation. This analysis also show that the NVIX is strong negatively correlated to the S&P 500 in the period from 2007 to 2012. This is a strong indication that the Norwegian and US market is connected, (which we also can see from the correlation of 0.93 between the two markets) and that the NVIX is closely related to the VIX. Our regression analysis also says that the NVIX have significant explanatory power on the OBX at a 1% level, which further strengthens our belief that the NVIX can be used as a measure of fear and ambiguity in the Norwegian market.

Even though the correlation and econometric analysis gives us good results, we should also consider some of the parts that could be improved. One obvious flaw with the study is the relatively short time period we look at. The main reason for this is poor data, with several days of missing and/or low trading, which is probably one of the main reasons for why there didn't exist a Norwegian volatility index until now. However, from year 2007 the data is improving and the analysis produced from then till now have very high significance.

Another thing we need to be aware of is that the NVIX and the VIX for that matter is not a "clean" measure of fear and ambiguity. There will always be some other factors that can affect the course, for instance not all trades in index options are done for hedging purposes,

some are just purely speculative trades. In addition, the Norwegian options market is very small, so that the options trades done on the OBX are done by a small amount of people, where DNB is the largest actor. In spite of this we still believe the NVIX to be a good indicator of fear and ambiguity in the Norwegian market, as our analysis has shown us.

So why do we need a Norwegian volatility index? The world is clearly connected through globalization. In fact, the S&P 500 is the variable with the highest explanatory power to the OBX, even higher than the oil price. The VIX also have good explanatory power, and a strong negative correlation to the OBX, so why not just look at this when considering fear and ambiguity in the Norwegian market? It is our opinion that there should be an NVIX. The reason is that even though we are a small nation and very dependent of what happens in the US and the world economy in general, the world is not dependent of what happens in our economy. So if the Norwegian economy for some reason should experience a local downturn in the future, we could not depend on the VIX to say something of the fear and ambiguity in the Norwegian market. Because of this we hope that there one day will be an official NVIX with minute-by-minute trading.

5.0 FEARS

The VIX-index is not the only possibility to measure investor sentiment. Da, Engelberg and Gao (2009, revised 2011) introduces a new way to do this called “Financial and Economic Attitudes Revealed by Search (FEARS)”. The FEARS indices are made up by aggregating a number of negative economic words (such as recession and unemployed) made by US households, on Google search engine. This information can be found on a site called Google Insights for Search (www.google.com/insights/search), this site makes public the Search Volume Index (SVI) which is an index for each single word, telling you how many searches have been done on that word relative to the total number of searches done on Google over time.

Traditionally there are two approaches of measuring investor sentiment. The first one is market-based measures such as trading volume and option implied volatilities (VIX index). The second approach is survey-based indices such as the Purchasing Managers Index (PMI) or the University of Michigan Consumer Sentiment Index. The FEARS indices may have some advantages when compared to these two approaches. Market-based measures have the disadvantage of being the equilibrium outcome several other economic forces that are not related to investor sentiment. When compared to survey-based measures, the FEARS indices have the advantage of being available in real time and the fact that they reveal attitudes rather than inquire about them. It is plausible to think that people are more “honest” when typing in search words in Google as compared to when they are filling out a survey.

The construction of the FEARS indices begins with choosing the right words to best capture investor sentiment. In an earlier study done by Tetlock (2007) where the effects of pessimism in the financial media were seen in relation to the stock market movements, Tetlock found that negative words associated with negative outlook was the best way to capture investor sentiment in contrast to positive words with positive outlook. Because of this the FEARS indices consists of only negative words. The first words were gathered from the Harvard IV-4 list of “negative” words, only the negative words which were also economic was picked, which resulted in 40 unique words. When searching these words, Google Insights will provide a list of ten “top searches” related to each word. From these related words only the economic ones were kept. The final word list consists of 27 words after infrequently searched words (words with two years of missing data) have been eliminated. These words are then sorted

into micro- and macro words, and the result is 19 words concerning micro-related FEARS and 8 words concerning macro-related FEARS:

Since Google Insights data goes back to 2004 the sampling period is January 2004 through March 2011. The U.S. SVI for this period was then downloaded for each word. In order to get daily data you need to download maximum one quarter at a time, this means that the daily SVIs in a quarter are scaled by the time series SVI in that quarter. This normalization meant that Da et al. (2011) were not able to compute the daily SVI on the first trading day in each quarter.

Before making the FEARS indices, the data needed to be adjusted for seasonality, heteroskedasticity and extreme values. Seasonality exists because of change in SVI rises in the beginning of the week and then falls throughout the week. Because of considerable difference in variance among search terms the data also have heteroskedasticity. Extreme values typically occur around holidays.

When doing their regression Da et al. (2011) find that the Micro FEARS index is not significant, the Macro FEARS index is however significant, and show some interesting results. Increases in macro-related FEARS correspond with low market-level returns on the same day, but they also predict high returns over the next few days. This effect could imply a temporary mispricing due to negative investor sentiment. A similar reversal pattern is found when examining the predictive power of the Macro FEARS index related changes in the VIX index. What they find is that an increase in Macro FEARS today predicts a negative change for the VIX a few days later. This reversal pattern supports the idea that the initial increase in the VIX contains a component due to negative sentiment. Finally, when seeing the Macro FEARS index in relation to daily aggregate mutual fund flows they find that increases in Macro FEARS index triggers investors to sell equity funds, but not bond funds, thereby pushing the price down on the equity funds. Da et al. (2011) suggest that this is direct evidence for “noise” trading in the market. Noise trading refers to an investor who buys and sells securities without thinking of fundamental data, only relying on trends and typically over-reacts to good and bad news.

5.1 Norwegian FEARS

Constructing our own NFEARS index based on Norwegian Google data is an alternative method to measure investor sentiment. One interesting difference is that this is a tool that measure inputs and not outputs like the VIX. While the VIX index is a product of index options, this index is straight forward data on search volume for negative economic words. As well as measuring inputs the search results reflects concerns and interests people don't know or think about that they are being measured on.

5.1.1 Collecting the FEARS data

In order to construct the NFEARS (Norwegian Financial and Economic Attitudes Revealed by Search) index we first had to come up with relevant sentiment revealing search terms. According to Tetlock (2007) negative economic words seem to be most informative. When searching for these words on the Google insight search page we soon discover that much of the data are inadequate. A typical example of this can be seen in figure 10 below, which show the search volume for “økonomisk krise” (economic crisis) from 2004 until today. We observe that the data do not start before 2008 and have large gaps in 2009 and 2011.

Figure 11: «Økonomisk krise»



www.google.com/insights/search/

We also found that it was easier to find useable data when searching for micro-economic words than macro-economic words. Another problem we encountered was that because of the gaps in the data, it was impossible to get daily data for each quarter. Some of the quarters were in daily data, while some was in weekly, and some were only available as monthly data. Because of this we made the decision to make all the data into weekly numbers, so that we could compare them, and thereby we discarded those words with only monthly data.

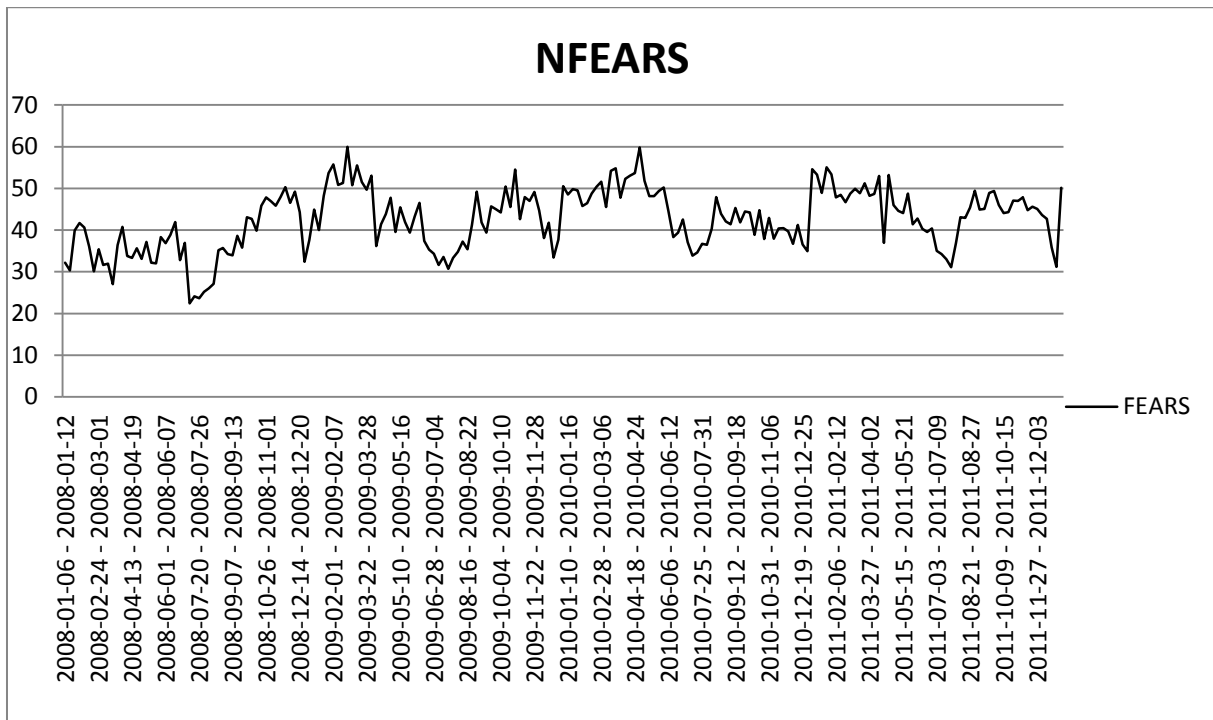
Unfortunately, this limits our range of words and data rate.

Da et al. (2011) divide their search words into macro and micro and test these indexes separately. Because of the predominance of micro words we choose to make only one index with both micro and macro words. Unfortunately Da, Engelberg and Gao only found the macro FEARS index significant in their study. Nevertheless we want to test our version of the FEARS index. The selected time period is 2008-2012 because most of the words had little or no data before 2008. The words we ended up with are listed below:

- Jobb ledig (Jobs available)
- Dagpenger (Unemployment benefits)
- Gjeld (Debt)
- Inflasjon (Inflation)
- Inkasso (Collection)
- Jobbsøknad (Job application)
- Konkurs (Bankruptcy)
- Ledige stillinger (Vacancies)
- Kostnader (Costs)
- Nav (Unemployment office)
- Oppsigelse (Termination)
- Permittering (Lay-off)
- Stilling ledig (Situation Vacant)
- Utgifter (Expences)

Figure 12 shows the FEARS index based on these words.

Figure 12: NFEARS index 2008-2012



5.1.2 Econometric NFEARS analysis

We want to test whether the Norwegian FEARS index (NFEARS) is a good measure of investor sentiment. This is done in a similar way as when testing the NVIX by a correlation analysis and a regression analysis in Excel. We want to use the same macro variables we used in the analysis of the NVIX to explain changes in OBX.

Table 13: Correlation matrix 2008-2012

	<i>NFEARS</i>	<i>OBX</i>	<i>NVIX</i>	<i>VIX</i>	<i>NOK_USD</i>	<i>Brent_USD</i>	<i>S&P500</i>	<i>NIBOR</i>
NFEARS	1							
OBX	-0,2084428	1						
NVIX	0,0700185	-0,8504394	1					
VIX	0,1893651	-0,7944002	0,9260967	1				
NOK_USD	-0,3768871	0,8256884	-0,6171343	-0,6217481	1			
Brent_USD	-0,3120136	0,8677581	-0,6144921	-0,5355240	0,8762081	1		
S&P 500	-0,2840625	0,9432675	-0,7088237	-0,6708712	0,8763309	0,89489804	1	
NIBOR	-0,4234681	0,0543563	0,3882795	0,3189632	0,3306486	0,2966837	0,2682702	1

The correlation matrix shows that there is some negative correlation between OBX and NFEARS but close to no correlation between NFEARS and NVIX. The strongest correlation is with NIBOR at -0,42, indicating a moderate negative linear relationship between the two variables.

Table 14: Descriptive statistics

<i>NFEARS</i>	
Mean	42,49
Std. deviation	7,35
Kurtosis	-0,35
Skewness	-0,25
Minimum	22,43
Maximum	60,00
Number of observations	209,00

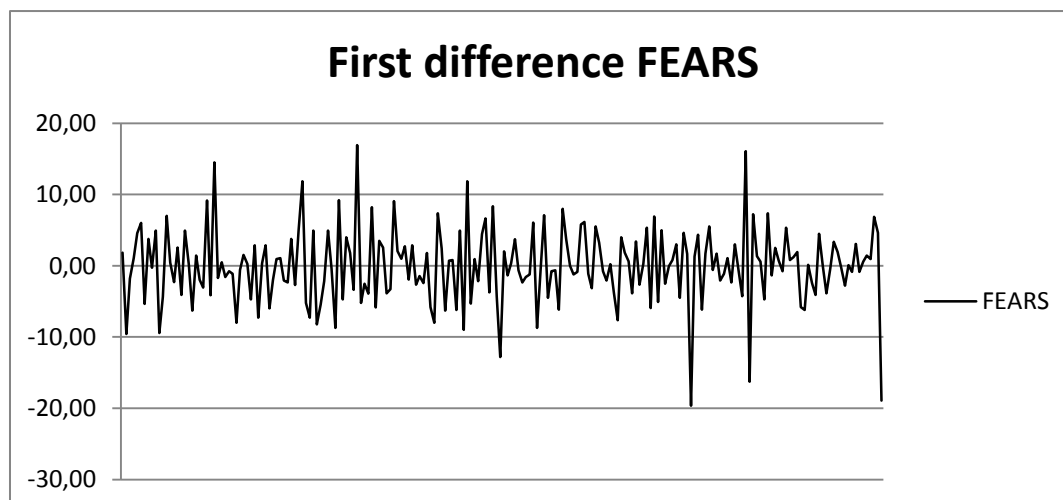
5.1.3 Hypothesis

The main purpose of this method is to test if we are able to measure investor sentiment by the constructed NFEARS index. In reality it is just adding one more explanatory variable to our first regression when testing NVIX to see if this new NFEARS index can explain movements in OBX. Our null hypothesis is therefore; $\beta_{NFEARS} = 0$ NFEARS does not have any effect on OBX, and the alternative hypothesis is; $\beta_{NFEARS} \neq 0$ NFEARS does affect OBX.

5.1.4 Regression

Before testing the hypothesis we need to prepare the input data. All the data has to be weekly and the NFEARS index needs to be transformed into log first difference like the other variables. Transforming the data into logarithmic form is a solution to avoid heteroskedasticity which is easy to spot visually in our NFEARS index below.

Figure 13: First Difference NFEARS



Both independent and dependent variables are transformed into logarithmic form and the log-log regression therefore reduces the likelihood of autocorrelation as the results are percentage changes. The Augmented Dickey-Fuller test (ADF) shows a t-value at -7,19083 and a p-value at 0,01. The t-value is lower than the critical value and we don't have to worry about non-stationary data. The regression line is presented below with the regression results in table 15:

$$\Delta OBX_t = \beta_0 + \beta_1 \Delta NFEARS_t + \beta_2 \Delta NVIX_t + \beta_3 \Delta VIX_t + \beta_4 \Delta NOK_USD_t + \beta_5 \Delta Brent_USD_t + \beta_6 \Delta S\&P500_t + \beta_7 NIBOR_t + u_t$$

Table 15: Regression 6, 2008-2012

<i>Significance-F</i>	<i>R-squared</i>	<i>Adjusted R-squared</i>
2,7656E-64	0,790118265	0,782772405

	<i>Coefficient</i>	<i>t-Stat</i>	<i>P-value</i>
Intercept	-0,000537342	-0,396295052	0,692309859
NFEARS	0,008646811	0,802263791	0,42335263
NVIX	-0,141812855	-5,908946021	1,46245E-08
VIX	0,005922549	0,373178607	0,709410743
NOK_USD	-0,04139126	-0,432111998	0,666125527
Brent_USD	0,244324523	6,544491798	4,90722E-10
S&P 500	0,84211065	11,0987627	1,30352E-22
NIBOR	-0,053149109	-1,076453835	0,2830211

The regression above tells us that our NFEARS index is not significantly different from zero and has no explanatory power of the movement in OBX. We tried to lag OBX to see if there was a delayed connection between the two but NFEARS was still not significantly different from zero. With this result we didn't find it useful to continue the analysis any further.

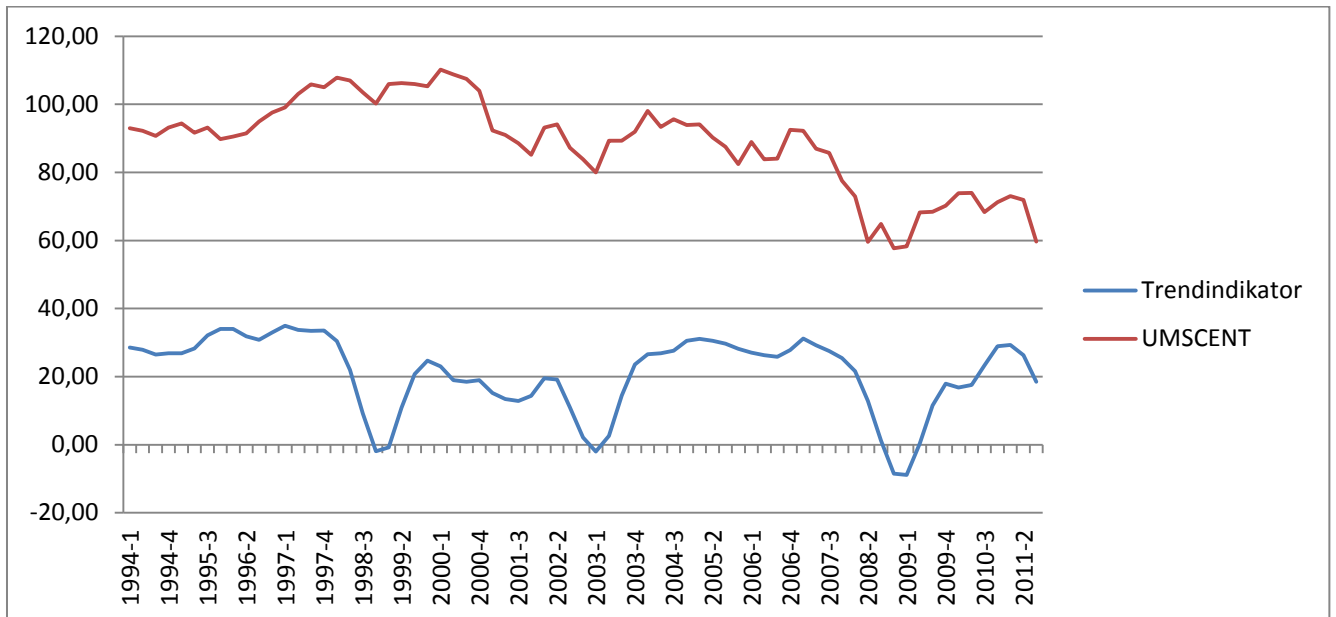
5.1.5 Survey based approach

Da et al. (2011) talk about the two traditional approaches of measuring investor sentiment, market-based and survey-based. When testing the FEARS index they use the survey-based University of Michigan Consumer Sentiment Index (UMSCENT) as an explanatory value for the market. We found a Norwegian substitute for this index, the "Trendindikatoren", produced by TNS gallup. This is the Norwegian equivalent to the UMSCENT, an index that tell us about household's outlook on their own, and the country's economy. We were lucky enough to be given this data from Finansnæringens Fellesorganisasjon, but unfortunately it was in quarterly numbers, and would therefore not be much help when testing the Norwegian

FEARS.

Even though we don't have frequently data we want to have a look at the possibility of measuring investment sentiment with survey based data. In the following graph we compare the Norwegian "Trendindikatoren" and the UMSCENT Index.

Figure 14: Trendindikatoren & UMSCENT



Just looking at the graph you can see some correlation between the two indexes even though they don't share the same value system. Measuring the correlation we found it to be positive at 36 %. To test the significance of these indexes on OBX we transformed all the data from 2008 to quarterly and ran a regression.

$$\Delta OBX_t = \beta_0 + \beta_1 \Delta NVIX_t + \beta_2 \Delta VIX_t + \beta_3 \Delta NOK_USD_t + \beta_4 \Delta Brent_USD_t + \beta_5 \Delta S\&P\ 500_t + \beta_6 \Delta NIBOR_t + \beta_7 \text{Trendindikator}_t + \beta_8 \Delta UMSCENT_t + \beta_9 \Delta NFEARS_t + u_t$$

Table 16: Regression 7, 2008-2012

<i>Significance-F</i>	<i>R-square</i>	<i>Adjusted R-square</i>
5,12197E-06	0,994211446	0,985528614

	<i>Coefficients</i>	<i>Standard error</i>	<i>t-Stat</i>	<i>P-value</i>
Intercept	-4,59421096	3,189681594	-1,440335289	0,199843043
NVIX	-0,225397764	0,13019549	-1,731225588	0,1341275
VIX	-0,080810309	0,094700034	-0,853329246	0,426225005
NOK_USD	-0,448599857	0,316956892	-1,41533397	0,206718127
Brent_USD	-0,409670355	0,2181327	-1,8780786	0,109446172
S&P 500	2,22155166	0,647710121	3,42985479	0,013973671
NIBOR	-0,119479711	0,066205866	-1,804669554	0,121166288
Trendindikator	-0,035676582	0,017230045	-2,070602938	0,083817485
UMSCENT	-0,649113697	0,273793394	-2,37081577	0,055458898
NFEARS	-0,068196826	0,084371802	-0,808289307	0,449775671

Both the “Trendindikator” and UMSCENT seem to be significantly different from zero at a 10 % significance level. The adjusted R-square is very close to one and the Significance-F tells us that the regression as a whole is significant. At the same time we can see a dramatic change in the other explanatory variables to the worse. This is probably due to the fact that the data is transformed to quarterly which makes the regression based on very few observations compared to when looking at daily or weekly data.

5.1.6 Evaluation NFEARS

This approach to measure investor sentiment failed when using Norwegian data. There can be many reasons for that. The first thing that comes to mind is the lack of data on macro words available. One though concerning this is about the number of private investors in the stock market. One would think it is more likely for private investors to search the internet for information than it is for professionals that might get their information elsewhere. According to Nordnet and Gallup.com 13 % of the Norwegian population own shares in the stock market while this number is 54 % in the US. This may be a reason why it was difficult to find Google search data on macro words. In addition we have a very different financial climate in Norway since the financial crisis didn’t affect the Norwegian people in the same scale.

The macro index made by Da et al. (2011) is dominated by recession words like “the depression”, “great depression”, “the great depression” and “recession”. This is four out of eight words that make the index. The rest of the words are “national debt”, “inflation”,

“inflation rate” and “cost of living”. Although they find the macro FEARS index significant we question the completeness of the index and its ability to measure investment sentiment because of the monotonous composition of words.

Although we failed to measure investment sentiment using Norwegian search data from Google we think it’s an interesting approach with potential. With increasing use of internet, cheaper and more available stock trading for private investors this method could yield significant data in the future.

The survey based approach is meant as an example of how this can be done. Because of the few observations we have to interpret the results more as an indication than valid results.

6.0 Conclusion

Fear and ambiguity are human emotions, and are therefore present in any financial market where people are investing. There are several ways in which researchers try to measure the level of these emotions in the market. The most famous measure of fear and ambiguity is perhaps the VIX index. This index is designed to measure future volatility on the S&P 500, but is more often referred to as a measure of fear. The VIX index has been replicated on many other indexes and countries, but not in Norway. Since fear and ambiguity is also present in the Norwegian stock market, our main goal with this thesis was to make a Norwegian volatility index, to measure fear and ambiguity on the OBX index.

In this thesis we make the following three contributions: First we survey the most recent literature on decision theory and risk-taking, including papers published in 2011. Second, we extend existing empirical risk research by constructing VIX and FEARS measures for the Norwegian market, which we name NVIX and NFEARS. Third we evaluate the comparative performance of our fear and ambiguity measures in Norway.

The main part of this thesis is about constructing and performing econometrical analysis on the NVIX. Especially the construction was very time consuming and it involved; retrieving and sorting all the index option trades made on the OBX from 1997-2012; learning how to make a volatility index through reading the “VIX White” made public on the CBOE website; making some sample NVIX values in excel; learning how to use SAS (Statistical Analysis Software); and finally making the NVIX in SAS. All of these points took a lot of time, but since we had no experience in using SAS the two latter points were particularly time consuming. In addition we attempt to measure investor sentiment by making the NFEARS, which consists of various negative economical search words made in Google. When the NVIX and NFEARS is made we test them using correlation and econometrical analysis.

Firstly, we do a correlation analysis on the NVIX for the period from 2003 to 2012 and the period from 2007 to 2012 separately. As mentioned above in part 4 this is done because of poor data with low trading from 2003 through 2006. Our results from this correlation analysis show almost no correlation between the NVIX and the OBX when looking at the whole period. However, when only looking at the period with the good data (2007-2012) we find a negative correlation of -0.8298, meaning that for this period the NVIX and OBX have a

strong negative linear relationship, which is consistent our expectations.

Secondly, we do an econometric analysis where we test to see if the NVIX does affect or predict OBX volatility. We run the regression for the same two periods as in the correlation analysis and the result is an adjusted R-squared of 0.4584 and 0.5147 for the period from 2003 to 2012 and from 2007 to 2012 respectively. This means that the latter period have a better explanatory power. For both periods we find that we can reject the null hypothesis that NVIX does not predict or affect OBX volatility. The null hypothesis can be rejected at a significance level of 1%.

After these analyses we wanted to see whether the NVIX could be used as a measure of fear or not, in doing so we replicated the method of Whaley (2008) who did this test for the VIX. The variables in this test are first difference OBX and a dummy variable which takes the OBX first difference value if the OBX is going down and 0 otherwise. This is to test whether the NVIX move more when the OBX declines as opposed to when it rises. When running this test we got an adjusted R-squared of 0.086, which is very low and indicating that important variables have been left out. Another problem is the OBX dummy variable which is only significant at a 10% level. If we look beyond this however, we find that if the OBX rise by 1% the NVIX falls by -0.916% and if the OBX falls by 1% the NVIX will rise by 1.332%. This is clearly asymmetric and a good indication that the NVIX is more sensitive to measuring investors fear of downside as opposed to investors greed.

After finishing our work with the NVIX we decided to attempt measuring investor sentiment using a different approach. This approach was to make a NFEARS index, using Google search words. After making our NFEARS index we found that it had no explanatory power on the OBX, and therefore we decided to not take this research any further. We also tried to include a survey based measure of investor sentiment called the “Trendindikator”, but due to insufficient data, we were not able to conclude anything from this.

In conclusion, the NFEARS is not a good measure for fear and ambiguity in the Norwegian stock market to this date, the data is simply too poor. We do however believe that the NVIX is a good measure for fear and ambiguity in the Norwegian stock market and this is the method we would use to measure investor sentiment. That being said, we think that the NFEARS method has potential, and with better data this could be something for future research.

7.0 References

Books.

Ackert, L. F., & Deaves, R. (2010). Behavioral finance. Psychology, decision making and markets. South-Western, Cengage Learning.

Bodie, Z., Kane, A., & Marcus, A. J. (2009). Investments. McGraw-Hill.

Gilboa, I. (2009). Theory of decision under uncertainty. Cambridge university press.

Gollier, C. (2004). The economics of risk and time. The MIT press.

Greene, W. H. (2003). Econometric Analysis. Prentice Hall.

Hull, J. C. (2011). Options, futures, and other derivatives. Pearson education limited.

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgement under uncertainty: Heuristics and biases. Cambridge university press.

Neumann J. V., & Morgenstern O. (1944). Theory of games and economic behavior. Princeton university press.

Snyder, C. & Nicholson, W. (2008). Microeconomic Theory, Basic principles and extensions. South-Western, Cengage Learning.

Stock, J. H., & Watson, M. W. (2007). Introduction to econometrics. Pearson Addison Wesley.

Varian, H. R. (1993). Microeconomic analysis. New York: W.W. Norton & Company Inc.

Periodicals.

Baker, M., & Wurgler, J. (2007, June). Investor sentiment in the stock market. *Journal of economic perspectives*, American economic association. Volume. 21. No. 2. pp. 129-152

Ellsberg, D. (1961, November). Risk, Ambiguity, and the Savage Axioms. *The quarterly journal of economics*, 75, pp. 643-659

Gennaioli, N., & Shleifer, A. (2010, November). What comes to mind. *The quarterly journal of economics*, 75, pp. 1399 - 1433

Gjerde, Ø., & Sættem, F. (1999, January). Causal relations among stock returns and macroeconomic variables in a small, open economy. *Journal of international Financial Markets, Institutions and Money*. Volume. 9. No. 1. pp. 61-74

Goldstein, D. G., & Taleb, N. N. (2007, March 28). We don't quite know what we are talking about when we talk about volatility. *Journal of portfolio management*. Vol. 33. No. 4.

Jorion, P. (1995, June). Predicting volatility in a foreign exchange market. *The journal of finance*. Volume 50. No. 2. pp. 507-528

Kahneman, D., & Tversky, A. (1979, March). Prospect Theory: An analysis of decision under risk. *Econometrica*. Volume 47. No. 2. pp 263-292.

Tetlock, P. C. (2007, June). Giving content to investor sentiment: the role of media in the stock market. *The journal of finance*. Volume 62. No. 3. pp. 1139-1168

Tversky, A. Kahneman, D. (1981, January 30). The Framing of decisions and the psychology of choice. [Electronic version]. *New series*, Vol. 211, No. 4481. pp. 453-458

Internet periodicals.

Barber, B. M., & Odean, T., (2008). All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Retrieved February 23, 2012, from http://faculty.haas.berkeley.edu/odean/Papers%20current%20versions/AllThatGlitters_RFS_2008.pdf

Bordalo, P., Gennaioli, N., & Shleifer, A., (2010, March. Revised 2012, January). Saliency theory of choice under risk. Retrieved February 18, 2012, from http://www.economics.harvard.edu/faculty/shleifer/files/Saliency_QJE_Jan18_final.pdf

Gennaioli, N., & Shleifer, A. (2008, September 17). Local thinking. Retrieved February 18, 2012, from <http://econ.as.nyu.edu/docs/IO/9363/083w-shleifer.pdf>

Gilboa, I. (2009, May). Decision Theory. Retrieved January 20, 2012, from http://www.tau.ac.il/~igilboa/pdf/Gilboa_PSEn_Decision_Theory.pdf

Gjerde, Ø., Knivsflå, K. H., & Sættem, F. (2000, October). The stock market and investment in the small and open Norwegian economy. Retrieved May 1, 2012, from http://brage.bibsys.no/nhh/bitstream/URN:NBN:no-bibsys_brage_22634/1/R55_00.pdf

Karni, E. (2005, November 9). Savages subjective expected utility model. Retrieved January 21, 2012, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.117.541&rep=rep1&type=pdf>

Whaley, R. E. (2008, November 6). Understanding VIX. Retrieved April 3, 2012, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1296743

Internet.

About. Trading options as insurance. Retrieved March 9, 2012, from <http://daytrading.about.com/od/options/qt/OptionsInsuranc.htm>

Behaviouralfinance. Prospect Theory. Retrieved February 20, 2012, from <http://prospect-theory.behaviouralfinance.net/>

CBOE. The CBOE volatility index – VIX. Retrieved March 9, 2012, from <http://www.cboe.com/micro/vix/vixwhite.pdf>

Econport. Decision-making under uncertainty – basic concepts. Retrieved February 26, 2012, from http://www.econport.org/econport/request?page=man_ru_basics4

Eia. Europe brent spot price FOB. Retrieved April 4, 2012, from <http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RB RTE&f=D>

Gallup. In U.S., 54% have stock market investments, lowest since 1999. (04.20.2011) <http://www.gallup.com/poll/147206/stock-market-investments-lowest-1999.aspx>

Google. About Google trends. Retrieved April 28, 2012, from <http://www.google.com/intl/en/trends/about.html#1>

Google insights. Google innsikt i søk. (Google insights). <http://www.google.com/insights/search/>

Google insights. Search word example “økonomisk krise”. (Economic crisis). <http://www.google.com/insights/search/#q=%C3%B8konomisk%20krise&geo=NO&cmpt=q>

Huffingtonpost. Stock market plunges begin to feed economic fear. (08.21.2011). http://www.huffingtonpost.com/2011/08/21/stock-market-recession_n_932389.html

Investopedia. Noise trader. Retrieved April 28, 2012, from <http://www.investopedia.com/terms/n/noisetrader.asp#axzz1u5P8BajL>

Mataf. Historiske norske kroner / US Dollar (NOK/USD). Retrieved April 4, 2012, from <http://www.mataf.net/no/histo/NOK-USD>

Newsweb. OBX vekter 17.04.2012 (OBX weights 04.17.2012). Retrieved April 17, 2012, from <http://www.newsweb.no/newsweb/search.do?messageId=302703>

Norges bank. NIBOR. Daily observations of nominal rate. Retrieved April 4, 2012, from <http://www.norges-bank.no/en/price-stability/interest-rates/nibor-nominal-interest-rate-daily-observations/>

Nornettbloggen. En av fire vil ha Facebook-aksjen. (One out of four wants the Facebook-stock. (05.18.2012) <http://nordnetbloggen.no/en-av-fire-vil-ha-facebook-aksjen/>

Oslo børs. www.oslobors.no

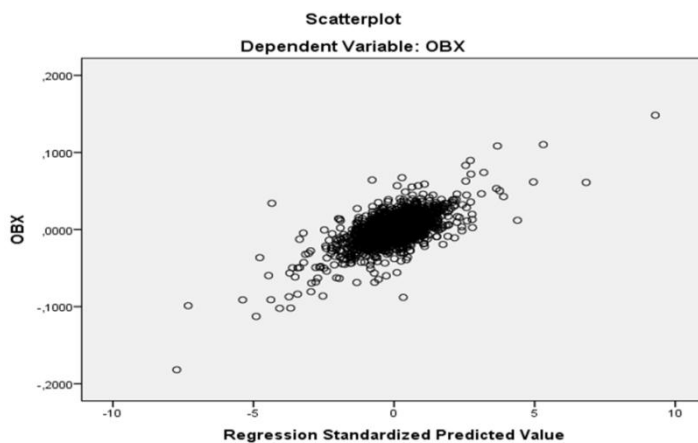
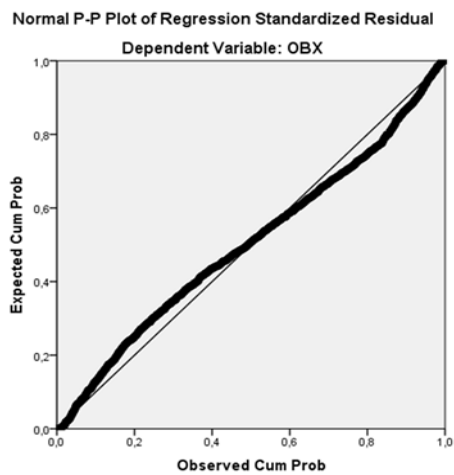
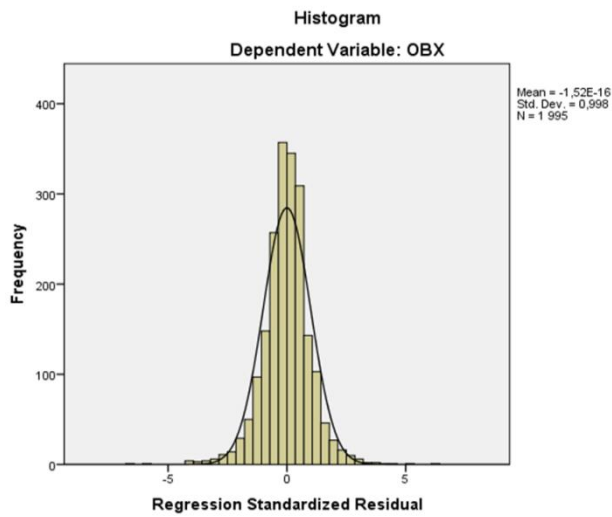
Uchicago. Tulipmania. Retrieved May 6, 2012, from http://penelope.uchicago.edu/~grout/encyclopaedia_romana/aconite/tulipomania.html

Vpsinfo. Ownerships distribution in listed companies. Retrieved May 18, 2012, from <http://vpsinfo.manamind.com/sectorstats/personStatList.do?f=a&d=&p=&l=en&m=p>

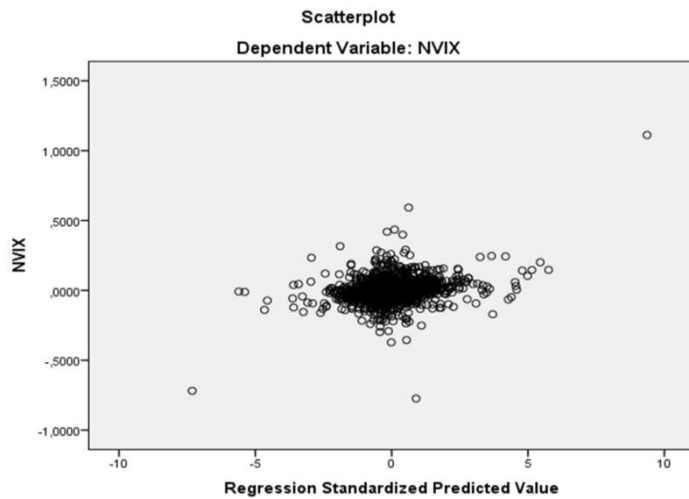
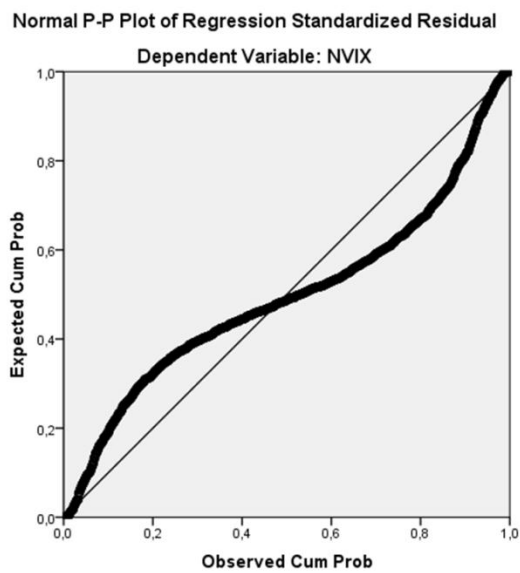
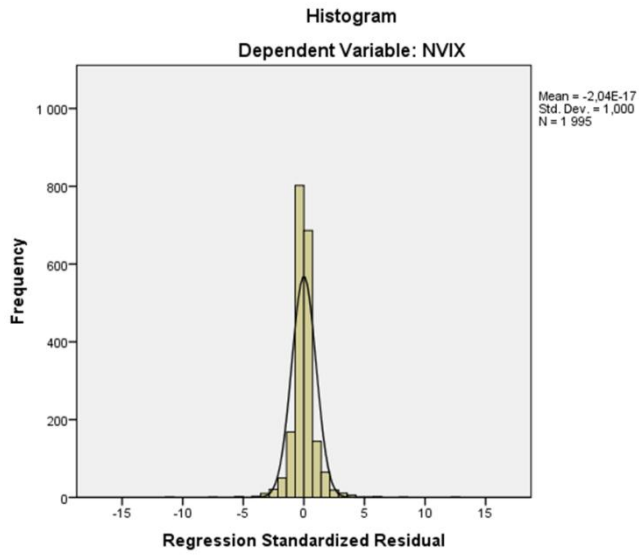
8.0 Appendix

8.1 Histograms, P-P plots and scatterplots retrieved from SPSS

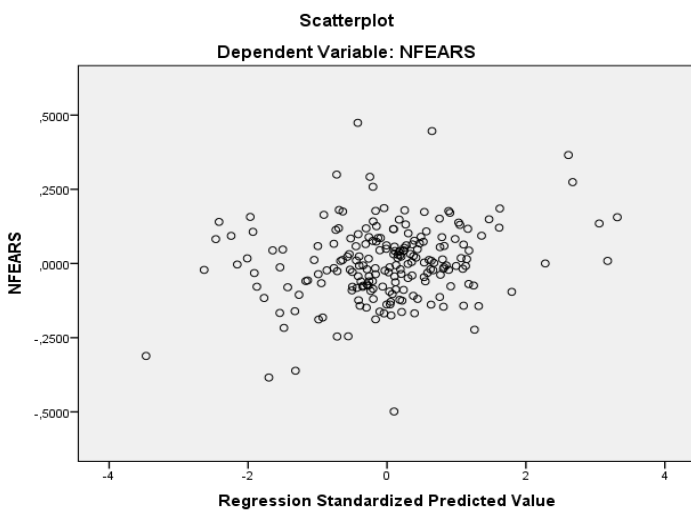
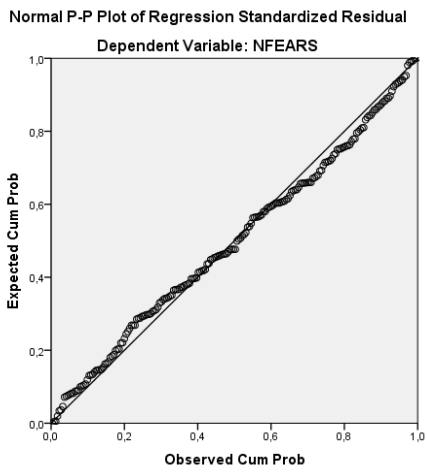
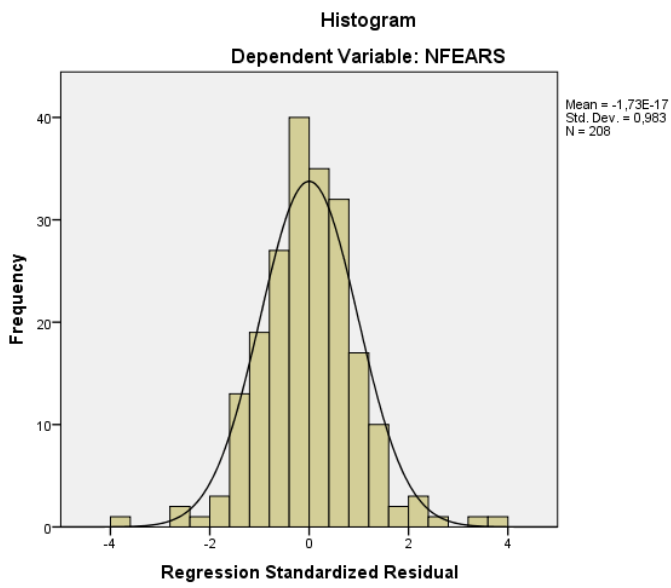
Histogram, P-P plot and Scatterplot for OBX



Histogram, P-P plot and Scatterplot for NVIX



Histogram, P-P plot and Scatterplot for NFEARS



8.2 SAS Codes

```
proc sql; create table table1 as
select *
from DATA
order by 'aggr date'n desc;
quit;
```

```
/*Checking for duplicates*/
```

```
proc sql; create table test as
select 'aggr date'n, symbol, count(symbol)
from table1
group by 'aggr date'n, symbol;
quit;
```

```
/*Finding if the transactions are Put or Call
```

```
    Finding number of days from start to end date*/
```

```
proc sql; create table table2 as
select *, 'aggr date'n as aggrdate, day('aggr date'n) as day,
substr(symbol,5,1) as put_temp,
    case when calculated put_temp in
('A','B','C','D','E','F','G','H','I','J','K','L') then 1 else
0 end as
    Call,
    intck('day', 'aggr date'n, 'tradethru'n) as DayDiff
from table1
where calculated daydiff > 7
order by 'aggr date'n;
quit;
```

```
/*Counting number of transactions per transactiondate per
number of days from start to end*/
```

```
proc sql; create table table2 as
select *, count(daydiff) as count /*, 'aggr date'n, sum(ask-
bid) as SumDiff*/
from table2
group by aggrdate, daydiff;
quit;
/*
proc sql; create table table3 as
select *
from table2
where count >=7;
```



```

quit;*/

/*Unique identification for each row*/

data table3;
set table2;
obs = _n_;
run;

proc sort data=table3; by aggrdate daydiff; quit;

data tableX;
set table3;
retain daydiff2 temp_var;
by aggrdate;
    if first.aggrdate then do;
        temp_var = 1;
        daydiff2 =daydiff;
    end;
    else do;
        if daydiff = daydiff2 then do;
            temp_var = temp_var;
            daydiff2=daydiff;
        end;
        else do;
            temp_var = temp_var+1;
            daydiff2 = daydiff;
        end;
    end;
run;

data tableX (drop = daydiff2);
set tableX;
run;

/*Finding the 2 first end "transaction days"*/

proc sql; create table tableX as
select *
from tableX
where temp_var < 3;
quit;

/*Finding the average for transactions*/
/*Finding the Strike price*/

```

```

proc sql; create table tablex_1 as
select *, (bid+ask)/2 as snitt, substr(symbol,6,9) as
strikeprice,
    case when daydiff >= 0 then 930 end as M_current_day,
    case when daydiff >= 0 then 510 end as M_settlement_day,
    case when daydiff >= 0 then daydiff-2 end as M_other_days,
    case when daydiff >= 0 then (daydiff-2)*1440 end as
M_other_minutes,
    case when daydiff >=0 then (((daydiff-
2)*1440)+930+510)/525600 end as minutes
from tablex;
quit;

```

```

data tablex_1;
set tablex_1;
if put_temp = '' then do;
    call = .;
end;
run;

```

/*Finding all CALL*/

```

proc sql; create table call as
select *
from tablex_1
where call = 1;
quit;

```

/*Finding all PUT*/

```

proc sql; create table put as
select *
from tablex_1
where call = 0;
quit;

```

/* Finding NEAR OVERVIEW*/

```

proc sql; create table call_near_overview as
select a.aggrdate, substr(a.symbol,6,9) as symbol_call,
a.snitt as snitt_call, substr(b.symbol,6,9) as symbol_put,
    b.snitt as snitt_put, snitt_call-snitt_put as
snitt_diff, a.minutes as T1
from call a
    left join put b
        on a.aggrdate=b.aggrdate and a.strikeprice=b.strikeprice
where a.temp_var = 1 and b.temp_var=1;
quit;

```

```

proc sql; create table call_near_overview as
select *, case when snitt_diff < 0 then snitt_diff*-1 else
snitt_diff end as snitt_diff_abs
from call_near_overview;
quit;

data call_near_overview (drop= snitt_diff);
set call_near_overview;
symbol_call_int=symbol_call*1;
symbol_put_int=symbol_put*1;
intrest=0.0038;
F1=symbol_call_int+CONSTANT('E')**(intrest*T1)*(snitt_call-
snitt_put);
if aggrdate le 16868 then do;
    K1=floor(F1/10)*10;
end; else do;
    K1=floor(F1/10)*10+floor(mod(F1,10)/5)*5;
end;
if symbol_call_int > K1 then OptionType='Call'; else
if symbol_call_int < K1 then OptionType='Put'; else
OptionType='K1';
if optiontype='Put' then Mid_Quote_Price=Snitt_Put; else
if optiontype='Call' then Mid_Quote_Price=Snitt_call; else
if optiontype='K1' then
Mid_Quote_Price=mean(snitt_put,snitt_call);

drop symbol_call symbol_put;
run;

proc sort;
by aggrdate symbol_call_int;
run;

data call_near_overview_;
set call_near_overview;
lag_K=lag(symbol_call_int);
run;

proc sort data=call_near_overview_;
by aggrdate descending symbol_call_int;
run;

data call_near_overview_;
set call_near_overview_;
lead_K=lag(symbol_call_int);
run;

proc sort data=call_near_overview_;
by aggrdate symbol_call_int;
run;

```

```

data call_near_overview_;
set call_near_overview_;
by aggrdate;
if first.aggrdate then Delta_K=lead_K-symbol_call_int; else
if last.aggrdate then Delta_K=symbol_call_int-lag_K; else
Delta_K=(lead_K-lag_K)/2;
Contr_by_strike=Delta_K/symbol_call_int**2*CONSTANT('E')**(int
rest*T1)*Mid_Quote_Price;
temp_var=2*Contr_by_strike/T1;
run;

```

```

proc univariate data=call_near_overview_ noprint;
by aggrdate;
output out=temp1 sum=sum_contribution1;
var temp_var;
run;

```

```

/*Finding smallest absolute value per day*/

```

```

proc sort data= call_near_overview;
by aggrdate snitt_diff_abs;
quit;

```

```

data near_overview_lowest;
set call_near_overview;
by aggrdate snitt_diff_abs;
if first.aggrdate;
var1=1/t1*(f1/k1-1)**2;
keep aggrdate f1 K1 T1 var1;
run;

```

```

data total_near;
merge near_overview_lowest temp1;
by aggrdate;
signal=sum_contribution1-var1;
run;

```

```

/* Finding NEXT OVERVIEW*/

```

```

proc sql; create table call_next_overview as
select a.aggrdate, substr(a.symbol,6,9) as symbol_call,
a.snitt as snitt_call, substr(b.symbol,6,9) as symbol_put,
      b.snitt as snitt_put, snitt_call-snitt_put as
snitt_diff, a.minutes as T2
from call a
      left join put b
      on a.aggrdate=b.aggrdate and a.strikeprice=b.strikeprice

```

```
where a.temp_var = 2 and b.temp_var=2;
quit;
```

```
proc sql; create table call_next_overview as
select *, case when snitt_diff < 0 then snitt_diff*-1 else
snitt_diff end as snitt_diff_abs
from call_next_overview;
quit;
```

```
data call_next_overview (drop= snitt_diff);
set call_next_overview;
symbol_call_int=symbol_call*1;
symbol_put_int=symbol_put*1;
intrest=0.0038;
F2=symbol_call_int+CONSTANT('E')**(intrest*T2)*(snitt_call-
snitt_put);
if aggrdate le 16868 then do;
    K2=floor(F2/10)*10;
end; else do;
    K2=floor(F2/10)*10+floor(mod(F2,10)/5)*5;
end;
if symbol_call_int > K2 then OptionType='Call'; else
if symbol_call_int < K2 then OptionType='Put'; else
OptionType='K2';
if optiontype='Put' then Mid_Quote_Price=Snitt_Put; else
if optiontype='Call' then Mid_Quote_Price=Snitt_call; else
if optiontype='K2' then
Mid_Quote_Price=mean(snitt_put,snitt_call);

drop symbol_call symbol_put;
run;
```

```
proc sort;
by aggrdate symbol_call_int;
run;
```

```
data call_next_overview_;
set call_next_overview;
lag_K=lag(symbol_call_int);
run;
```

```
proc sort data=call_next_overview_;
by aggrdate descending symbol_call_int;
run;
```

```
data call_next_overview_;
set call_next_overview_;
lead_K=lag(symbol_call_int);
run;
```

```

proc sort data=call_next_overview_;
by aggrdate symbol_call_int;
run;

data call_next_overview_;
set call_next_overview_;
by aggrdate;
if first.aggrdate then Delta_K=lead_K-symbol_call_int; else
if last.aggrdate then Delta_K=symbol_call_int-lag_K; else
Delta_K=(lead_K-lag_K)/2;
Contr_by_strike=Delta_K/symbol_call_int**2*CONSTANT('E')**(int
rest*T2)*Mid_Quote_Price;
temp_var=2*Contr_by_strike/T2;
run;

proc univariate data=call_next_overview_ noprint;
by aggrdate;
output out=temp2 sum=sum_contribution2;
var temp_var;
run;

/* Finding smallest absolute value per day*/

proc sort data= call_next_overview;
by aggrdate snitt_diff_abs;
quit;

data next_overview_lowest;
set call_next_overview;
by aggrdate snitt_diff_abs;
if first.aggrdate;
var2=1/t2*(f2/k2-1)**2;
keep aggrdate f2 K2 T2 var2;
run;

data total_next;
merge next_overview_lowest temp2;
by aggrdate;
sigma2=sum_contribution2-var2;
run;

data total;
merge total_near total_next;
by aggrdate;
VIX=100*sqrt((T1*sigma1*(T2*525600-43200)/(T2*525600-
T1*525600)+T2*Sigma2*(43200-T1*525600)/(T2*525600-
T1*525600))*525600/43200);
run;

```

```

/*
Proc sql; create table near_term_strike as
select*
from call
where temp_var = 1
order by 'aggrdate'n asc;
quit;

proc sql; create table next_term_strike as
select aggrdate, strikeprice,
from put
where temp_var=2
order by 'aggrdate'n asc;
quit;*/

```

8.3 STATA codes

. varbasic lnfNVIX lnfOBX

Vector autoregression

Sample: 3 - 1995	No. of obs	=	1993
Log likelihood = 7552.259	AIC	=	-7.56875
FPE = 1.77e-06	HQIC	=	-7.558436
Det(Sigma_ml) = 1.75e-06	SBIC	=	-7.540664

Equation	Parms	RMSE	R-sq	chi2	P>chi2
lnfNVIX	5	.072282	0.0892	195.1979	0.0000
lnfOBX	5	.019313	0.0013	2.639272	0.6199

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----						
lnfNVIX						
lnfNVIX						
L1.	-.3101952	.0233383	-13.29	0.000	-.3559374	-.2644531
L2.	-.1253732	.0232318	-5.40	0.000	-.1709066	-.0798398
lnfOBX						
L1.	-.6805619	.0879895	-7.73	0.000	-.8530181	-.5081058
L2.	-.2988188	.0888837	-3.36	0.001	-.4730277	-.1246099
_cons	.0007615	.0016189	0.47	0.638	-.0024113	.0039344
-----+-----						
lnfOBX						
lnfNVIX						
L1.	-.0017415	.0062356	-0.28	0.780	-.0139631	.0104801

L2.		-0.015509	.0062072	-0.25	0.803	-0.0137168	.0106149
InfOBX							
L1.		-0.0100486	.0235095	-0.43	0.669	-0.0561263	.0360291
L2.		-0.0370913	.0237484	-1.56	0.118	-0.0836374	.0094547
_cons		.0006059	.0004325	1.40	0.161	-0.0002418	.0014537

. irf graph irf

