NTNU – Trondheim
Norwegian University of
Science and Technology

The Performance of Market Risk Measures on High and Low Risk Portfolios in the Norwegian and European Markets.

Christian Preben Bang

Master of Science in Physics and Mathematics
Submission date: June 2012
Supervisor: Jacob Laading, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

This thesis is a study of market risk for portfolios of varying risk. The market risk measures Value at Risk and Expected Shortfall will be tested on these portfolios through a backtesting procedure on historical equity and interest rate price data from 2007 to 2012. Particular emphasis will be placed on the time-variation of the performance of the risk measures.

Assignment given: 30.01.2012

Supervisor: Jacob Laading, Department of Mathematical Sciences

Preface

The work on this thesis was carried out at the Department of Mathematical Sciences at the Norwegian University of Science and Technology (NTNU), Trondheim, during the spring semester 2012. It represent a semester load of work and leads to the degree Master of Science.

I would like to thank my supervisor, Associate Professor Jacob Laading for his constructive, relevant feedback and helpful advice.

Trondheim, June 2012

Christian Preben Bang

Abstract

A basic overview of mathematical finance and pricing theory is given. The Black-Scholes model and the LIBOR Market Model are explained, and their assumptions are discussed and tested on historical data. The normality of log-returns of stocks and forward rates is tested for different time periods, and is found to be varying greatly over time. The models are calibrated using the Exponentially Weighted Moving Average (EWMA) method and implemented to perform a backtest against historical data of two risk measures, Value at Risk and Expected Shortfall. The backtesting is done on five portfolios of varying risk, in the European and Norwegian markets. Three unleveraged portfolios consisting of bonds and stocks in different proportions, and two leveraged portfolios consisting of stocks and interest rate caps respectively are considered.

The performance of the risk measures is found to be not satisfactory for all portfolios, but performance is better for riskier portfolios and assets. Variation of performance over different time periods is found. The periods of worst performance are those of turbulent market conditions, notably in late 2008. These periods are found to loosely correspond to the time periods in which log-returns of equity and forward rates are least normal.

A sensitivity analysis of performances to the weighting parameter in the EWMA is done. The sensitivity is found to be substantial for all portfolios except for the portfolios holding stocks in the Norwegian market.

En grunnleggende oversikt over matematisk finans og prissettingsteori fremføres. Black-Scholes-modellen og LIBOR Market Model presenteres og modellenes antagelser testes på historiske data. Normaliteten av log-avkastning av aksjer og forward-renter testes for ulike tidsperioder, og stor tidsvariasjon blir funnet. Modellene kalibreres ved hjelp av Exponentially Weighted Moving Average (EWMA) og implementeres i den hensikt å utføre en backtest mot historiske data av to risikomål, Value at Risk og Expected Shortfall. Backtestingen gjøres på fem porteføljer med ulik risiko, i det europeiske og norske markedet. Tre ubelånte porteføljer bestående av obligasjoner og aksjer i ulike proporsjoner og to belånte porteføljer bestående av henholdsvis aksjer og interest rate caps blir vurdert.

Risikomålenes ytelse er dårlig for samtlige porteføljer, men ytelsen er bedre for høyrisikoporteføljer og -verdipapirer. Ytelsesvariasjon over ulike tidsperioder påvises. Periodene med dårligst ytelse er perioder preget av turbulente markeder, som i slutten av 2008. Disse periodene kan løst sies å tilsvare de tidsperiodene hvor log-avkastningen av aksjer og forward-renter fjerner seg mest fra normalfordelingen.

Det gjøres en sensitivitetsanalyse av ytelsesresultatene mhp. vektingsparameteren i EWMA-metoden. Sensitiviteten er betydelig for alle porteføljer unntatt for porteføljene som holder aksjer i det norske markedet.

Contents

1 Introduction

1.1 Background

The science of creating mathematical models of financial assets is relatively new. Shreve [8] names Harry Markowitz the founder of the mathematical theory of finance with his 1952 Ph.D. thesis *Portfolio Selection*. However, it was Robert Merton that in 1969 brought stochastic calculus into the study of finance, and with Fischer Black and Myron Schole's work on the fair price of a stock option, mathematical finance was born. Today, mathematical finance is a large field of scientific study that help us understand and improve the financial system and inspires the creation of new mathematics, operating in the cross-section between applied mathematics and financial economics.

The ongoing financial crisis was triggered by the bursting of the American housing bubble in 2007, resulting in real-estate based securities dropping in value and in turn causing financial distress for banks and illiquid credit markets. In 2008, as banks and financial institution went bankrupt, was subject to government takeover or received bailouts, the crisis spread to international financial markets. Plummeting investor confidence and declining credit availability resulted in stock markets crashing and economic contraction in many countries.

Bailouts in the financial and other sectors and the reduction in economic activity resulted in a worsening of government finances, especially in the "peripheral Euro zone countries" Greece, Ireland, Spain and Portugal. The fear of government bankruptcies, and a following collapse of the Euro, is still hindering economic growth and investor confidence. We do not yet know the full consequences of the now five years old financial crisis. The story of the financial crisis makes apparent the need for development in the field of risk management.

The field of mathematical finance can be loosely divided into two separate fields: the theory of the pricing of financial derivatives and the theory of risk and portfolio management. In this thesis, both fields play a role but we focus mainly on the latter. We go through each step in the process of financial risk modelling, with the main goal of measuring the performance of the popular risk measures Value at Risk and Expected Shortfall on different types of portfolios.

Consider we hold a portfolio, and we want information about the market risk associated with holding it. How much money might we lose in a week, for instance, due to changes in asset prices? To get this information we need to model the time evolution of the portfolio's asset.

First of all, we need models for the value of the basic or *underlying* assets, typically stocks and interest rates. These models should incorporate the most important characteristics of real world financial assets. Notably the probability distribution of asset price changes (we want our asset price models to have the same probabilistic behaviour as real world asset prices) and the exclusion of arbitrage opportunities (we do not want our models to allow for trading strategies that give risk free profits, since this is a marginal phenomenon in real markets).

The assumptions of our models should be tested against real world historical data. If we find that the assumptions of our model does not correspond with reality, they need to be improved. Of course, like in all mathematical modelling, there is a trade-off between having models that are easy or even possible to work with and having models that correspond with reality. A good model should be simple and easily understood, but not so simple that it departs too much from the phenomenon we are indeed modelling.

We also need to estimate the parameters of our models, i.e. calibrate our model. For the model to have any predictive power, it needs to reflect market conditions, so we calibrate our model based on the past and current market movements.

For risk management purposes we want to have an idea of the probability distribution of the value of the collection of financial assets held, the *portfolio*. To achieve this, we run multiple simulations forward in time and get the simulated probability distribution of future portfolio value. This simulation involves two steps.

First, we simulate the underlying assets forward to some future time using their real, as assumed by our models, probabilistic behaviour. We are in this step working with the real probability measure $\mathbb{P}$, so it is sometimes referred to as "$\mathbb{P}$-world" simulation.

Secondly, we need to value our portfolio at this future time, within each simulation. This can be a hard problem in itself, since our portfolio may include *derivative* financial assets, like stock options or interest rate agreements whose value depend on the underlying assets, but often in non-trivial ways. This step may include simulations, if we do not have other, faster numerical methods or closed form solutions for the derivatives' dependence on the underlying assets. When valuing our portfolios, it is most often useful to work with models under some probability measure $\mathbb{Q}$, like the risk-neutral martingale measure so this step is sometimes referred to as "$\mathbb{Q}$-world" simulation.

Based on the simulated probability distribution for future portfolio value, we calculate the risk measures which give information about the market risk associated with the portfolio.

In this thesis we go through the steps outlined above and try to give an overview of the pitfalls and possibilities in each step, but our main purpose is to investigate to what degree these risk measures can be trusted. We do this by a *backtesting* procedure. We simulate future values of our portfolio from some date in the past for which we have price data, as if the following period's prices are unknown, to get a simulated probability distribution. We do this for each day in our historical data and compare it with what actually happened, thus testing the performance of our risk measures.

1.2 Thesis outline

This thesis consist of three parts. In the first part, section 2,3, and 4, we introduce some needed financial and mathematical concepts. In the second part, sections 5 and 6, we introduce the Black-Scholes model and the LIBOR Market Model, and test their assumptions against historical data. In the final part, section 7 and 8, we explain the specifics of our implementation of the backtesting procedure and present the results.

2 Financial concepts

2.1 Financial assets

A financial asset is a contract that gives the holder ownership over future (expected or certain) cash flows. The (expected) present value of these cash flows determines the value of the asset. Typical assets are stocks and bonds and derivative assets, like stock options or interest rate derivatives.

In this thesis we model all underlying asset values and rates as Ito processes, i.e. stochastic processes governed by the equations

$$dS(t) = \mu(t, S)dt + \sigma(t, S)dW(t), \quad t > 0 \tag{2.1}$$

$$S(0) = s_0 \tag{2.2}$$

where $S(t)$ is the asset value at time t, the adapted stochastic processes μ and σ are functions of S and t and W is a *Brownian Motion.* Note that σ and W may be vectors, i.e we may have more than one source of randomness, and μ and σ are allowed to be depending on t and S. The dt term is called the drift term and the dW term is called the diffusion term.

2.2 Equity (Stocks)

A stock is a financial contract representing the holder's ownership of a share of a company. The stockholder thus have ownership of future cash flows generated by the company.

Some stocks are traded on a stock exchange, providing tremendous liquidity. That is, exchange traded stocks can be bought and sold in a fraction of a second. A stock index is a value weighted average of a group of stocks. In this thesis we consider the OSEBX and Eurostoxx 50 indices.

2.3 Fixed Income (Bonds)

A bond is a financial contract where one party pays a sum of money today in exchange for fixed cash flows from the counterparty in the future. Bonds are issued by governments, corporations, or other types of organizations as a means to raise money. The cash flows of a bond may be spread out equally over the bonds life-time, may include a lump payment at maturity or use some other form of payment plan.

The simplest type of bond is the *zero-coupon bond* (often called a *zero*) which pays the holder a lump sum at maturity T with no pre-maturity payments (coupons). I shall use the notation $Z(t, T)$ for the time t value of a zero-coupon bond that pays 1 at time T.

2.4 Yield

The yield $Y(t,T)$ of the zero-coupon bond is defined by the following relation

$$Z(t,T) = e^{-Y(t,T)(T-t)}$$

It is simply the constant (continuously compounded) rate at which your money is growing if you pay Z(t,T) at time t and get 1 at time T. Rearranging, we get the yield

$$Y(t,T) = -\frac{\log(Z(t,T))}{(T-t)}$$

Note that if you own a zero-coupon bond V that pays, for instance, 100 NOK at time T, i.e. $V(t,T) = V(T,T)Z(t,T) = 100\text{NOK} \times Z(t,T)$ we get the yield of V by substituting $Z(t,T) = \frac{V(t,T)}{V(T;T)}$

2.5 Continuous forward rate

The *continuous forward rate* $f(t,T)$ is not a financial entity. It is a theoretical concept. It is the instantaneous continuously compounded rate agreed on at time t for borrowing at time T. It is defined by the following relation

$$Z(t,T) = e^{-\int_t^T f(t,u)du}$$

Note that, at time t, the forward rate $f(t,T)$ is known for all T, so the forward rate curve $\{f(t,u) : u \geq t\}$ is a deterministic function. We model the forward rate curve as evolving in t, i.e. we have one forward rate curve for each value of t. The spot rate $r(t)$ for borrowing at time t relates to the forward rate in that $r(t) = f(t,t)$

2.6 Simple rates

Consider the contract where one party (the lender) agrees to pay a notional amount N today (time t) to another party (the borrower) who will pay back

$$N(1 + \delta F(t,T))$$

at time T, where $\delta = T - t$ when time is measured in units of a year. The (annual) rate $F(t,T)$ is the *simple (spot) rate* for borrowing from time t to time T.

2.7 The forward rate agreement

The forward rate agreement is a contract where the parties agree at time t (today) on a rate $L(t,T_1,T_2)$ for borrowing in some future time period, T_1 to T_2. That is, with a notional amount of N, they agree at time t that the lender pays N at time T_1 and the borrower pays back

$$N(1 + \delta L(t, T_1, T_2))$$

at time T_2, where $\delta = T_2 - T_1$.

The (annual) rate $L(t, T_1, T_2)$ is called the *simple forward rate* We see that the simple rate $F(t, T) = L(t, t, T)$ and that the continuous forward rate is the limit of the simple forward rate,

$$\lim_{T_{i+1} \to T_i^+} L(t, T_i, T_{i+1}) = f(t, T_i).$$

2.8 The LIBOR rate

The LIBOR (London Interbank Offered Rate) is an important type of (spot and forward) simple rate. The LIBOR rate is an average of the rates paid by major London banks when borrowing money from each other. It is quoted in the major currencies with different debt maturities. The LIBOR is a much used benchmark rate in financial contracts.

2.9 Derivatives

A *derivative* is a financial contract whose value depends on the value of underlying assets or rates, typically stock values or simple (forward or spot) rates. On of the simplest, and most widely traded, derivatives is the European call option, which gives to the holder a payoff of

$$max(S(T) - K, 0)$$

at some time T for some stock S and some fixed *strike* K. In other words, it gives a payoff of $S(T) - K$ at time T if the stock price S is larger than the strike K at that time, and gives zero payoff otherwise.

2.10 The cap/floor

Consider a set of equally spaced maturity dates $T_i = \delta i,\ i = 0, 1, ..., M + 1$. A caplet is a derivative contract defined by some floating interest rate $L(t, T_n, T_{n+1})$, with the maturity date T_n, a notional amount N, and a fixed strike rate K. The buyer of the cap receives payment from the seller if the floating rate is above the strike rate at maturity. The payoff function is

$$g(L(T_n, T_n, T_{n+1})) = N\delta(L(T_n, T_n, T_{n+1}) - K)^+ \tag{2.3}$$

where $\delta = T_{n+1} - T_n$ is the fraction of a year corresponding to the floating rate period. This payment is made at time T_{n+1}. If for instance the caplet is written on a 3-month LIBOR we would have $\delta = 0.25$. In other words, a caplet is a European call option on some forward rate.

A cap is a collection of caplets with different consecutive maturities. Consider an investor borrowing \$100 for 3 months at a time at the LIBOR rate. If this investor also buys a cap on 3-month LIBOR with strike rate K, notional of \$100 and maturities 3,6,9,... months from contract start, the sum effect of the portfolio is that she never pays a higher rate than K. That is, she has placed a cap on her interest rate and thus reduced her interest rate risk.

In the same way, a floorlet is a European put option on some floating rate, and a floor is a collection of floorlets which can be used to reduce the interest rate risk for an investor lending money.

2.11 The interest rate swap

In general a swap is simply a contract that exchanges one cash flow for another. The typical interest rate swap is a contract where one party, called the payer, agrees to pay a fixed rate interest (the "fixed leg") on some notional and the other party, called the receiver, pays some floating rate (e.g. LIBOR) interest (the "floating leg") on the same notional. The swap rate is the specific fixed rate that makes the value of the swap equal to zero (for both parties).

To be more precise, we follow the considerations of Glasserman [3] and consider a set of equally spaced dates $T_i = i\delta,\ i = 0, ..., M+1$. Consider a swap with payment dates $T_{n+1}, ..., T_{M+1}$ on some notional N with fixed rate R. At each payment date T_i the payer pays δNR and the receiver pays $\delta NL(T_i, T_i, T_{i+1})$. The value of the swap at time T_n from, the payer's point of view, is the difference between the value of the floating leg (which he receives) and the fixed leg (which he pays).

To simplify calculations, we include a fictitious payment at T_{M+1} of N that each party pays to the other. In that way, the value of the fixed and floating leg is the same as the value of a bond with fixed and floating rate coupons respectively, and a face value of N. The cash-flows of the floating bond can be replicated with an investment at T_n of N. Lending N between T_i and $T_i + 1$ creates a cash-flow of $\delta NL(T_i, T_i, T_{i+1})$ while keeping the original N (which create the last face value cash-flow at T_{M+1}). Therefore, the value of the floating bond must be N. The value of the fixed bond is simply the present value of the cash-flows, i.e. the cash-flows discounted by the zero coupon bonds,

$$\delta NR \sum_{n+1}^{M+1} Z(T_n, T_i) + NZ(T_n, T_{M+1}). \tag{2.4}$$

So the value of the swap from the payers point of view is

$$V = N - \delta NR \sum_{n+1}^{M+1} Z(T_n, T_i) + NZ(T_n, T_{M+1}) \tag{2.5}$$

The swap rate for this swap with payment dates $T_{n+1}, ..., T_{M+1}$, which we denote $S_n(T_n)$, is the fixed rate R that makes $V = 0$, so we have the swap rate

$$S_n(T_n) = \frac{1 - Z(T_n, T_{M+1})}{\delta \sum_{i=n+1}^{M+1} Z(T_n, T_i)} \tag{2.6}$$

2.12 The swaption

A swaption with maturity T is a derivative that gives the holder the right but not the obligation to enter into a swap at (European) or before (American) the swaption maturity T. Consider the swaption with maturity T_n on entering into the swap we discussed in the last section (payment dates $T_{n+1}, ..., T_{M+1}$), as the payer (pays fixed rate). If R is the fixed rate in the swap, we have that, at the option maturity T_n, the value of the swap is

$$V(T_n) = N - \delta N R \sum_{i=n+1}^{M+1} Z(T_n, T_i) + N Z(T_n, T_{M+1}) \tag{2.7}$$

The holder of the swaption will only exercise if $V(T_n) > 0$ so the payoff function is

$$P = \max(V(T_n), 0) \tag{2.8}$$

which is equal to 0 if $R > S_n(T_n)$ and equal to $V(T_n)$ if $R < S_n(T_n)$. Thus we have

$$P = \max(S_n(T_n) - R, 0)\delta \sum_{i=n+1}^{M+1} Z(T_n, T_i). \tag{2.9}$$

We see that the swaption has the payoff of a call option on the swap rate.

2.13 Arbitrage

Given a market, i.e. some set of investment opportunities with values $I_1(t), I_2(t), \ldots, I_m(t)$, being adapted processes defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, an arbitrage opportunity is any investment choice $a = (a_1, ..., a_m)$ such that the value of the investment $X(t) = a_1 I_1(t) + \ldots + a_m I_m(t)$ satisfies

- $X(0) = 0$
- $\mathbb{P}\{X(T) \leq 0\} = 1$
- $\mathbb{P}\{X(T) > 0\} => 0$

for some time $T > 0$.

In other words, an arbitrage opportunity is a portfolio needing no initial capital, with no risk of loosing money, and a possibility of earning money. Simply put, arbitrage is riskless profit, money for nothing. Arbitrage opportunities are rare in real financial markets because they disappear when people start trading them, since trading adjust the prices. We therefore want models of financial markets to exclude arbitrage opportunities.

2.14 The efficient market hypothesis

The *efficient market hypothesis (EMH)* says that all available information is already incorporated into the price of a financial asset, hence we cannot use this information to predict future returns. Put in other words, an investor cannot consistently, i.e. over time, earn excess return on his investment. That is, he cannot beat the market portfolio. In making this more precise we need to specify what kind of information we believe is incorporated in the prices. One generally speaks of three different versions of the EMH, which differ in the type of information that is hypothesized to be incorporated into the asset price.

Weak form efficiency says that all information about past prices is incorporated into the current price. That is, in the long run we cannot earn excess returns by any trading strategy based on historical price information.

Semi-strong form efficiency says that all public information (not only past price history) is near instantly incorporated into the price of an asset, so no excess returns can be earned in the long run by a trading strategy based on public information.

Strong form efficiency says that all information, public and private, is incorporated into the price of an asset, so no-one can earn excess returns in the long run.

The random characteristic of asset prices' evolution in time is explained by this instant incorporation of new information into the asset price. The price at some moment of a stock is the "correct price" since it is the value at which market participants are wiling to buy it, given the current information about the company, its business sector, the world economy etc. In other words, the stock price is a reflection of all available information. When new information arrives, as new events in the world unfold, the price adjusts accordingly. And this happens "instantaneously" i.e. very fast, according to the EMH.

There is a rich literature on testing the different forms of the EMH. One can test for autocorrelation in return time series, one can test different trading rules (e.g. always buying stock after a negative financial announcement from the company) to see if they perform significantly better than the market portfolio, and a multitude of other tests.

3 Modelling and pricing theory fundamentals

We are seeking mathematical models for the value of the different types of financial contracts. If we can build arbitrage-free models that hopefully coincides with reality, we can price these contracts and then quantify risk and manage portfolios. Since prices of financial assets and interest rates change randomly in time (for reasons discussed in the market efficiency section) our models need to incorporate this randomness. Asset prices and interest rates will therefore be modelled as stochastic processes, more specifically Ito processes. With this modelling choice, we can use past and present prices of assets and interest rates to estimate model parameters. This is discussed in section 5. We then have a working model for the *underlying* processes. However, we would like to be able to price financial contracts for which we do not have historical quotes and for which we have no obvious way of determining the value process dynamics. For the purpose of pricing such contracts we shall need one of the cornerstones of mathematical finance, namely *pricing theory.*

3.1 Pricing theory

Consider a market $\mathcal{M}$ consisiting of N underlying assets or rates in a time period $[0,T]$. The uncertainty of the future values of the assets will be modelled through a d-dimensional Brownian motion $W(t) = (W_1(t),...,W_d(t))$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We let $\{\mathcal{F}_t : 0 \leq t \leq T\}$ be a filtration on the space such that $\mathcal{F}_t$ is the σ-algebra $\sigma(W(s) : 0 \leq s \leq t)$ generated by the d-dimensional Brownian motion $W(t)$, and we let $(S_1,...,S_N) : [0,T] \times \Omega \to \mathbb{R}^N$ be the assets' value processes. As mentioned, in this thesis underlying asset values will be modelled as Ito processes, so we have the set of SDEs

$$dS_i(t) = \mu_i(t,S)dt + \sigma_i(t,S)dW(t), \qquad 0 < t \leq T, \quad i = 1,...,N \tag{3.1}$$

$$S_i(0) = s_{i,0}, \qquad i = 1,...,N \tag{3.2}$$

where μ_i is a scalar and σ_i a d-dimensional vector.

We want to find a general pricing formula for a derivative whose value depends on one or more of these underlying assets or rates. The idea is to find this price by finding a combination of the underlying assets (for which we have a model) that have the same value as the derivative at all times. We can then find the price of the derivative by by exploiting this connection.

Definition 1 *A* ***portfolio*** *is a process* $\mathbf{a} = (a_1, a_2, ..., a_N) : \Omega \times [0,T] \to \mathbb{R}^N$ *denoting how many of each asset held by an investor. That is, the investor holds* a_i *of asset* S_i*. The* ***portfolio value process*** $V(t)$ *is thus defined by*

$$V(t) = \sum_{i=1}^{N} a_i S_i(t), \quad 0 \leq t \leq T. \tag{3.3}$$

The portfolio is called ***self-financing*** *if*

$$dV(t) = \sum_{i=1}^{N} a_i dS_i(t) \quad \forall 0 \leq t \leq T, \tag{3.4}$$

in other words if the infinitesimal portfolio value change is due to changes in asset prices, not due to adding or removing capital from the portfolio.

Definition 2 *An* ***equivalent martingale measure*** *is a probability measure* $\mathbb{Q}$*, equivalent to* $\mathbb{P}$*, such that all assets prices* S_i *are martingales under* $\mathbb{Q}$*.*

Definition 3 *A* ***numeraire*** $N(t)$ *is a value process of a portfolio in the market* $\mathcal{M}$ *such that* $N(t) \geq 0$ *for all* $t \geq 0$*.*

A numeraire can typically be one of the asset prices S_i. For a chosen numeraire $N(t)$, the new *discounted* assets $\frac{S_i}{N}$ constitute a new market $\mathcal{M}_N$. An important type of asset used in a lot of pricing is the *risk neutral asset* $B(t)$ which has initial value 1 and grows with a constant *risk free rate of return* r. That is,

$$B(t) = e^{rt}, \quad 0 < t \leq T \tag{3.5}$$

$$B(0) = 1 \tag{3.6}$$

If we use the risk neutral asset B as numeraire we obtain the market $\mathcal{M}_B$. The equivalent martingale measure in this market is called the *risk neutral measure*.

Theorem 1 ***(1st fundamental theorem of asset pricing)*** *If a market* $\mathcal{M}$ *has an equivalent martingale measure, the market has no arbitrage opportunities.*

See [8] for a proof.

Definition 4 *A* ***contingent claim*** *is a random variable* $X : \Omega \to \mathbb{R}$ *representing a cash-flow at some time* T*.*

A claim is said to be ***replicable*** *if there exist a self-financing portfolio* $\mathbf{a}$ *such that its time* T *value*

$$V(T) = \sum_{i=1}^{N} a_i(T) S_i(T) = X \tag{3.7}$$

almost surely.

The cash-flows/payoffs of a derivative can be represented by a contingent claim (or a combination of several). A replicable contingent claim can therefore be thought of as a derivative whose time T value is almost surely equal to the time T value of some portfolio of underlying assets.

We now denote by $X(t)$ the time t value of the replicable contingent claim, so that $X(T) = X$. If there was some time t where $V(t) \neq X(t)$ we would have an obvious arbitrage opportunity by buying the cheaper and selling the more expensive to get a risk-less profit, since the cash-flows from the replicating portfolio and the claim are almost surely equal at time T. Thus, we have that $V(t) = X(t)$ for all $0 \leq t \leq T$

If there exist an equivalent martingale measure in the market $\mathcal{M}_N$, $\tilde{\mathbb{P}}$ that makes discounted asset prices $\frac{S_i(t)}{N(t)} = D(t)S_i(t)$ martingales, we know that

$$V(t) = \frac{1}{\tilde{D}(t)}\tilde{\mathbb{E}}[V(T)D(T)|\mathcal{F}(t)] \tag{3.8}$$

$$= \frac{1}{\tilde{D}(t)}\tilde{\mathbb{E}}[X(T)D(T)|\mathcal{F}(t)] \tag{3.9}$$

$$\tag{3.10}$$

from the definition of martingales and using that $V(T) = X(T)$. Here $\tilde{\mathbb{E}}$ denotes expectation with respect to the measure $\tilde{\mathbb{P}}$. Since $X(t) = V(t)$ for all $0 \leq t \leq T$, because of the exclusion of arbitrage opportunities, we have that

$$X(t) = \frac{1}{\tilde{D}(t)}\tilde{\mathbb{E}}[X(T)D(T)|\mathcal{F}(t)]. \tag{3.11}$$

We have achieved a pricing formula for the contingent claim. The task of finding the time t price $X(t)$ of some contract is reduced to calculating the expectation, which can be done in a number of ways, discussed in the following sections. It should be noted that $X(t)$ is only uniquely defined if there exists a unique equivalent martingale measure.

Theorem 2 *If the market $\mathcal{M}_N$ has a unique equivalent martingale measure $\tilde{\mathbb{P}}$, then every contingent claim X is replicable.*

See [8] for a proof. A market where every claim is replicable is called a *complete* market.

3.2 Monte Carlo simulation

We have seen that the value of an asset is the expected discounted value of future cash flows under the proper probability measure. Now, how can we use this to actually price traded assets? How do we calculate this expectation? One possibility is to try deriving an analytical expression from the definition of expectation, be it an exact solution or some good enough approximation. This approach, however, is unsuccessful for all but the simplest types of derivatives, so we need to find numerical solutions. There are two general ways of finding numerical solutions. One is to solve a PDE, provided by the

Feynman-Kac theorem (see appendix) or otherwise. The other way is to approximate the expectation through *Monte Carlo (MC) simulation.*

MC simulation is a numerical method for approximating the value of expectation integrals. Say you want to find the expected value of some function g of a random variable X with some density function f. MC-simulation uses the fact that if you draw n independent random samples $X_1, X_2, \ldots, X_n$ of the random variable then

$$g(\bar{X}) = \frac{g(X_1) + g(X_2) + \ldots + g(X_n)}{n} \approx \mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx \tag{3.12}$$

Further, we know that the standard deviation of $g(\bar{X})$ is approximately $\frac{\sigma}{\sqrt{n}}$ where σ is the standard deviation of one instance of $g(X)$. We can therefore estimate the error in our estimation of $\mathbb{E}[g(X)]$ by estimating the standard deviation of $g(\bar{X})$, $\frac{s}{\sqrt{n}}$, where

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (g(X_i) - g(\bar{X}))^2 \tag{3.13}$$

is the standard unbiased estimator for variance. We then have an approximate $(1-\alpha)$ confidence interval for $\mathbb{E}(g(X))$,

$$g(\bar{X}) \pm z_{\alpha/2} \frac{s}{\sqrt{n}} \tag{3.14}$$

I shall call

$$z_{\alpha/2} \frac{s}{\sqrt{n}} \tag{3.15}$$

the $(1-\alpha)$ *standard error* and hence

$$\left(z_{\alpha/2} \frac{s}{\sqrt{n}} \right) / g(\bar{X}) \tag{3.16}$$

the $(1-\alpha)$ *relative standard error.*

These powerful observations follows from the Central Limit Theorem.

For an example, consider we want to price a European call option written on the stocks S under the risk neutral measure. Under this measure any traded derivative (like an option) is a martingale when discounted by the factor e^{-rt}. That is, the discounted value process of the option is $e^{-rt} V(t, S(t))$ and we therefore know that

$$V(0, S(0)) = \mathbb{E}^*[e^{-rT} V(T, S(T))] = e^{-rT} \mathbb{E}^*[V(T, S(T))], \tag{3.17}$$

where the $*$ means that the expectation is taken with respect to the risk-neutral measure. Assuming the stock price S is a function of only one random number, $W(T)$, as is done in the famous Black-Scholes model to be introduced later, the payoff of the option at maturity is a function g of one random number. Therefore this expectation is defined to be

$$V(0) = e^{-rT} \int_{-\infty}^{\infty} g(x) f(x) dx \tag{3.18}$$

which, as explained above is approximately equal to the mean payoffs of the option (discounted by e^{-rT}) from n simulations of the stock prices evolution.

The weakness of MC simulation is that it converges slowly. However, the computational complexity increase more slowly with dimension than "regular" integral solvers. Therefore, we should ideally use some regular numerical method for low-dimensional problems. An example of a low-dimensional problem is to find the price of a European option as above. The payoff depends on only one random number $S(T)$, the price of the underlying at maturity so this is a one-dimensional problem and some other method than MC should ideally be used. Other types of options, however, may depend on the price of the underlying at many times prior to maturity, $S(T_1), S(T_2), ..., S(T_d)$ so these problems are many-dimensional. In these cases, MC simulation is often our best choice.

3.3 The Delta-Gamma approximation

Another way of pricing derivatives is to use the *delta-gamma approximation* for the derivative's dependence on the underlying. Assume that the derivative value $V(t, S)$ depends on the underlying asset value S. Given the value of the underlying and derivative at time t, the change in derivative value from t to $t + \delta t$ is approximately

$$V(t + \delta t, S(t + \delta t)) - V(t, S(t)) \approx \frac{\partial V}{\partial t}(t)\delta t + \frac{\partial V}{\partial S}(t)\delta S + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(t)(\delta S)^2 \tag{3.19}$$

where $\delta S = S(t + \delta t) - S(t)$. We see that the derivative price change is a function of the derivatives delta $\frac{\partial V}{\partial S}$ and gamma $\frac{\partial^2 V}{\partial S^2}$, hence the name "delta-gamma approximation". This approximaton comes from Ito's lemma. In the limit $\delta t \to dt$ the left and right hand side converge, i.e.

$$dV = \frac{\partial V}{\partial t}(t)dt + \frac{\partial V}{\partial S}(t)dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(t)(dS)^2 \tag{3.20}$$

where $(dS)^2 = d[S, S](t)$ is the infinitesimal change in quadratic variation. In the case of the Black-Scholes model for a stock, for instance, we have $dS^2 = \sigma^2 S^2 dt$. For the delta-gamma approximation to be useful, it must of course be computationally simpler to find the delta and gamma values than to do a MC simulation.

3.4 The Black caplet formula

Assume a set of *tenor* dates $0 = T_0 < T_1 < ... < T_{M+1}$, and let the forward rates $L(t, T_i, T_{i+1})$ be denoted by $L(t, T_i)$ The Black caplet formula gives the time t price of a caplet that pays

$$g(L) = \delta(L(T_n, T_n) - K)^+ \tag{3.21}$$

at time $T_{n+1} = T_n + \delta$. The caplet value is

$$V(t, T_n) = Z(t, T_{n+1}) [L(t, T_n)N(d_+) - KN(d_-)] \tag{3.22}$$

where

$$d_{\pm} = \frac{1}{\sqrt{\int_t^{T_n} \sigma^2(T_n - s)ds}} \left[\log \frac{L(t,T_n)}{K} \pm \frac{1}{2} \int_t^{T_n} \sigma^2(T_n - s)ds \right]. \tag{3.23}$$

and $N(\bullet)$ is the cumulative distribution function for the standard normal distribution $N(0,1)$. The formula is dependent on the process $L(t,T)$ following a log-normal walk with deterministic volatility σ. For our purposes we will assume that σ is stationary (depends on time only through the difference $T-t$). We shall further assume that σ is piecewise constant left continuous between the time to maturities $T_n - T_{n-1}, T_n - T_{n-2}, ..., T_n - 0$. We then have that

$$\int_t^{T_n} \sigma^2(T_n - s)ds \tag{3.24}$$

$$= \sigma^2(T_n - T_{k-1})(T_k - t) + \sigma^2(T_n - T_k)(T_{k+1} - T_k) + ... + \sigma^2(T_n - T_{n-1})(T_n - T_{n-1}) \tag{3.25}$$

$$= \sigma^2(T_n - T_{k-1})(T_k - t) + \delta \left(\sigma^2(T_n - T_k) + ... + \sigma^2(T_n - T_{n-1}) \right) \tag{3.26}$$

where T_k is the next immediate tenor date to t, and where we in the last equality have assumed that the dates $T_0, T_1, ..., T_{M+1}$ are equally spaced with $\delta = T_i - T_{i-1}$.

4 Financial risk measures

4.1 What is financial risk?

Financial risk is simply the risk of financial losses, of losing money. We distinguish between different types of risk by the reasons why the loss occurred. We have

- Market risk- the risk of financial loss due to jumps in market prices of held assets.
- Credit risk - the risk of financial loss due to debtors not meeting their obligations and defaulting.
- Operational risk - the risk of financial loss due to the operations of a company. Include legal risk, production failure risk, risk of fraud, etc.
- Liquidity risk - the risk of financial loss due to illiquid markets, i.e. that an asset cannot be traded fast enough to prevent a loss or achieving the required return.

In this thesis the object of study is the market risk of portfolios of financial assets.

4.2 Risk measures

It is of great importance for financial institutions to be able to quantify the market risk associated with their portfolios. We therefore need some good risk measures that provide this information. What constitutes a good risk measure? [1] presents some desirable properties of a risk measure and call risk measures that have these properties "coherent".

Let $X \in \mathcal{G}$ where $\mathcal{G}$ is a set of stochastic processes, like return processes of a portfolio. Then a coherent risk measure $\rho : \mathcal{G} \to \mathbb{R}$ must be

- **Translation invariant**: For all $X \in \mathcal{G}$ and all $\alpha \in \mathbb{R}$, $\rho(X + \alpha r) = \rho(X) - \alpha$. Adding wealth invested without risk to the portfolio reduces risk.
- **Subadditive**: For all X_1 and $X_2 \in \mathcal{G}$, $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$ (The risk of two portfolios combined cannot be greater than the sum of the individual risks of the portfolios).
- **Positive homogenous**: For all $X \in \mathcal{G}$ and all $\lambda \geq 0$, $\rho(\lambda X) = \lambda \rho(X)$ (Scaling our portfolio simply scales the risk)
- **Monotone**: For all X and $Y \in \mathcal{G}$ such that $X \geq Y$ in all states of the world, $\rho(X) \leq \rho(Y)$ (If one portfolio always has better returns than another its risk is smaller).

4.3 Value at Risk

Given a profit-and-loss probability distribution for a portfolio over some period of time - which might be estimated in a variety of ways - the *Value at Risk (VaR)* risk measure is defined in the following way:

Given a probability level x percent, and a time period τ, the VaR is the amount of money (or the proportion of portfolio value) such that the probability of a loss greater than the VaR is x percent.

For instance, if you have estimated that the probability of a loss in time τ equal to or greater than 10NOK is 5 percent. Then the 5 percent VaR over time τ is 10NOK.

Definition 5 *Let $V(t)$ be the portfolio value. Let t and $t+\tau$ be todays date and a later date respectively, let $\alpha \in (0,1)$ be some confidence level and let $L(t,\tau) = V(t) - V(t+\tau)$ be the (random) loss over the time period. Then the* VaR of the portfolio at the confidence level α *is the smallest number l such that the probability that the loss $L(t,\tau)$ exceeds l is not larger than α.*

$$VaR_{\alpha}(L(t,\tau)) = -\inf\{l \in \mathbb{R} : P(L(t,\tau) > l) \leq \alpha\} \tag{4.1}$$

Why is VaR a widely used risk measure? Why not just use the portfolio volatility as our risk measure? Since the VaR number is based on models whose parameters are calculated from current market conditions, it simply tells us that "markets just turned volatile so now you have more risk". But this is information about the past and present and we must be careful not to believe that VaR gives us information about the future. VaR simply tells us what we already know (our portfolio's volatility) in a new way. There are several reasons for using VaR to provide this information rather than some other measure, like portfolio variance.

One reason is simply that VaR is easy to understand. You do not need to be privy to the nuts and bolts of the mathematical models or the trading strategies to understand what the VaR number means. In addition, VaR can be aggregated. A financial institution can simply add the 1% VaR of different operations and trading desks to get the 1% VaR of the whole institution. For these reasons, VaR is a useful managerial tool for achieving an overview of an institution's financial risk.

It is important to note that the VaR does not give any information about the maximum plausible loss. Rather than viewing VaR as indicative of our worst possible losses, it should be viewed as the worst possible loss within the range of predictability. In other words, losses exceeding the VaR are impossible to predict and can be extremely large.

An undesirable property of the VaR is that it violates the sub-additivity property, and is therefore not a coherent risk measure. The related risk measure Expected Shortfall however, which we introduce next, *is* a coherent risk measure and may therefore be a better choice in some situations.

4.4 Expected Shortfall

Expected Shortfall (ES) is in the same family of risk measures as VaR, but whereas VaR is the α% percentile of the loss distribution, ES is the average over the worst α% of losses.

Definition 6 *Let $V(t)$ be the portfolio value. Let t and $t+\tau$ be todays date and a later date respectively, let $\alpha \in (0,1)$ be some confidence level and let $L(t,\tau) = V(t) - V(t+\tau)$ be the (random) loss over the time period. Then the* ***ES of the portfolio at the confidence level*** *α is the mean loss over the worst α% losses.*

$$ES_\alpha(L(t,\tau)) = \frac{1}{\alpha}\int_0^\alpha VaR_{\alpha'}(L(t,\tau))d\alpha' \tag{4.2}$$

VaR gives no information about what might happen in that last quantile, the worst 1% or 5% of cases. ES, on the other hand, seem to provide this. An important point, however, is that the output number from the model, e.g. the ES, is only as good as the model is close to reality. And the model is only as good as the data we use to calibrate it. Since extreme losses are rare, we have correspondingly few data points, and therefore large uncertainty in our modelling of the probability distribution.

As an extreme example, which applies to the considerations around VaR as well, consider a stock that over the last 10 years have had daily returns in the interval $[-3\%, 3\%]$ except for one special day where the return was -20%. Since we have around 2500 data points in the normal return interval (-3% to 3%) we know a great deal about the probabilities of returns within this range, assuming that historical return distribution is a good indicator for future return distributions. But we have no information about probabilities of larger gains or losses other than that they are indeed possible.

When calculating the ES, you use, implicitly or explicitly, some kind of empirical probability distribution for the worst 1% of cases. If we only have a few data points in that range, we are doing a reckless interpolation or extrapolation on our data. This is to be kept in mind when using ES. It does not necessarily provide more information about market risk than VaR.

4.5 Risk measure calculation

The VaR and ES of a portfolio are calculated from some loss probability distribution over some time period. In our calculations, we concern ourselves with the five day (weekly) loss of our portfolios. This five day loss distribution can be found/approximated in different ways. One simple way is to directly use the historical data, by calculating the loss our portfolio would suffer over each five day period in the historical data. Another approach is to use MC simulation of our portfolio value to achieve a five day loss distribution.

In this thesis we use the simulation approach. We simulate, using the models introduced in the next section, the evolution of portfolio value over five trading days. We do 1000 simulations, which gives us a simulated distribution of portfolio loss. Once we have the simulated distribution of portfolio loss, we pick the 5% quantile directly from the simulated distribution. We thus calculate ES as the arithmetic mean over the worst α%

of losses, and report VaR as the least bad of these. The details of this procedure are presented in section 7.3.

5 Models

5.1 The Black-Scholes model

The most popular model for stock price evolutions is the *Black-Scholes (BS) framework*. The model consist of a stock and a risk free asset (the same as discussed in section 3.1) whose prices at time t are denoted by $S(t)$ and $B(t)$ respectively. These two investment opportunities' values are governed by the equations

$$dS(t) = \mu S(t)dt + \sigma S(t)dW_t, \qquad S(0) = S_0 \tag{5.1}$$
$$dB(t) = rB(t)dt, \qquad B(0) = 1 \tag{5.2}$$

where r, μ, and σ are constant and W is a Brownian motion. In addition, the Black-Scholes framework makes the following assumptions

- There are no arbitrage opportunities. Therefore, any risk-less investment portfolio will have the same rate of return r.
- You can buy and sell (including short selling) any amount, including fractions, of the investment opportunities.
- There are no transaction costs
- The stock does not pay dividend

With the BS model at our disposal, we can price and *hedge* (create a replicating portfolio using stock and money market account) derivatives based on the stock price. Consider the European call option earlier mentioned, with maturity T, strike K and thus the payoff function

$$V(T, S) = max(S(T) - K, 0).$$

Letting the options value at time t be $V(t, S)$ we have from the preceding the pricing formula that

$$V(t, S) = N(t)\tilde{\mathbb{E}}[V(T, S)/N(T)|\mathcal{F}(t)] = N(t)\tilde{\mathbb{E}}[\max(S(T) - K, 0)/N(T)|\mathcal{F}(t)], \tag{5.3}$$

for some numeraire-measure pair $(N(t), \tilde{\mathbb{P}})$, so one way of finding $V(t, S)$ is to estimate the expectation like previously discussed. Since the risk free asset $B(t)$ is part of the Black-Scholes market, we can use this as a numeraire, so $N(t) = B(t) = e^{rt}$. We then need to find the expectation under the risk neutral measure, i.e. we need to find the dynamic of the stock price under this measure. This can be shown to be

$$dS(t) = rS(t)dt + \sigma S(t)dW_t, \qquad S(0) = S_0. \tag{5.4}$$

However, Wilmott ([12]) presents a different way of finding the price without directly calculating the expectation, where we find a PDE without using the Feynman-Kac theorem. By Ito's lemma we have that

$$dV = \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dW$$

Now consider a portfolio that consists of one option and some amount a of the stock. The value of this portfolio is $\Pi(t) = V(t) + a(t)S(t)$ and we have that in a small time step t

$$d\Pi = dV + adS$$

Here we have assumed that a does not change ($da = 0$) over a small time step, that is change in portfolio value comes only from changes in the portfolio asset values and not from rearranging postitions. We have

$$\begin{aligned} d\Pi &= \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dW + a\left(\mu S dt + \sigma S dW\right) \\ &= \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + a\mu S\right) dt + \sigma S\left(\frac{\partial V}{\partial S} + a\right) dW \end{aligned}$$

With the choice $a = -\frac{\partial V}{\partial S}$, the random part of $d\Pi$ is eliminated. That is, we have a risk-less portfolio, so we must have that

$$d\Pi = \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \mu S\frac{\partial V}{\partial S}\right) dt = r\Pi dt = r(V - \frac{\partial V}{\partial S}S)$$

We now have a PDE for the option value,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{5.5}$$

called the Black-Scholes equation.

In the derivation we did not assume anything about the function $V(t, S)$ except that it depends on t and S. This equation is therefore valid for all derivatives whose value is a function of t and S. It is through the initial and boundary conditions alone that the characteristics of the option is brought into the calculation.

For a European call or put the Black-Scholes equation has analytical solutions. However, for slightly more complicated derivatives we need to do numerical approximations.

5.2 The LIBOR Market Model

The LIBOR Market Model (LMM) models the term structure of interest rates through simple forward rates. Consider a set of maturity dates $0 = T_0 < T_1 < ... < T_{M+1}$, and let $\delta_i = T_{n+1} - T_i$ be the differences between these dates. The simple forward rate $L_n(t)$ is the interest rate agreed upon at time t for borrowing money from T_n to T_{n+1}. That is, the borrower agrees at time t to get 1 at time T_n and pay back $1 + \delta_n L_n(t)$ at time T_{n+1}. A simple replication argument shows that

$$1 + \delta_n L_n(t) = \frac{Z_n(t)}{Z_{n+1}(t)} \tag{5.6}$$

where $Z_n(t)$ is the time t price of a zero-coupon bond with maturity T_n. Thus

$$L_n(t) = \frac{Z_n(t) - Z_{n+1}(t)}{\delta_n Z_{n+1}(t)}, \quad 0 \leq t \leq T_n, \quad n = 0, 1, ..., M \tag{5.7}$$

For $t = T_n, n = 0, ...M$, (5.7) can be inverted so that we get

$$Z_n(T_i) = \prod_{j=i}^{n-1} \frac{1}{1 + \delta_j L_j(T_i)}, \quad n = i + 1, ..., M + 1 \tag{5.8}$$

But for general t between tenor dates, the forward rates give us no information about the discount rate, so we have

$$Z_n(t) = Z_{\eta(t)}(t) \prod_{j=\eta(t)}^{n-1} \frac{1}{1 + \delta_j L_j(t)}, \quad 0 \leq t < T_n \tag{5.9}$$

where $\eta(t)$ is the index of the next tenor date after t. That is, $\eta(t) = i$ s.t. $T_{i-1} \leq t < T_i$.

We are looking to model these simple forward rates, L_n. We let L_n, $n = 1, ...M$ be the forward rates with settling date T_n and maturity T_{n+1}. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a filtered probability space and $T > 0$ be some final time. The M forward rates will be modelled by the M-dimensional stochastic process $(L_1, ..., L_M) : [0, T] \times \Omega \to \mathbb{R}^M$. We want a log-normal model of the forward rates, so our starting point is to say that

$$\frac{dL_n}{L_n} = \mu_n(t)dt + \sigma_n(t)^\top dW(t), \quad 0 \leq t \leq T_n, \quad n = 1, ..., M \tag{5.10}$$

where W is a d-dimensional standard Brownian motion in $\mathbb{P}$. μ_n and $\sigma_n = (\sigma_{n1}, ..., \sigma_{nd})^\top$ might still be L-dependent.

The model cannot allow arbitrage, so we need to find conditions on μ_n so as to make this happen. We know that if the value of traded assets in an economy, discounted by some numeraire, are martingales under the corresponding (to the numeraire) probability measure then the economy is arbitrage free. We see here a possible path for creating arbitrage "freeness" in our model. Forward rates, however, are not traded assets but we know do their relation to zero-coupon bonds, which are traded assets, through Eq. (5.6). We can therefore proceed by checking what the conditions will be on the μ_n of our forward rates when we impose the condition that zero prices must be martingales under some measure. The following development and resulting dynamics is due to Glasserman (see [3] for a more complete account).

5.2.1 Spot measure

Consider the following trading strategy. Start with one unit in the money market account, and use them to buy $\frac{1}{Z_1(0)}$ zero coupon bonds maturing at T_1. Then, at time T_1 get the money $(\frac{1}{Z_1(0)})$ from the maturing zeros and use it to buy $\frac{1}{Z_1(0)Z_2(T_1)}$ and then continue doing this at each tenor date. The time t value of this portfolio is

$$Z^*(t) = Z_{\eta(t)}(t) \prod_{j=0}^{\eta(t)-1} [1 + \delta_j L_j(T_j)] \tag{5.11}$$

The spot measure P^* is the equivalent martingale measure that makes tradeable assets' values, dicounted with $Z^*(t)$, martingales. In other words, from eq's (5.9) and (5.11) we have that

$$D_n(t) = \frac{Z_n(t)}{Z^*(t)} = \left(\prod_{j=0}^{\eta(t)-1} \frac{1}{1 + \delta_j L_j(T_j)} \right) \prod_{j=\eta(t)}^{n-1} \frac{1}{1 + \delta_j L_j(t)}, \quad 0 \leq t \leq T_n \tag{5.12}$$

must be a martingale under the spot measure, and thus have the dynamic

$$\frac{dD_{n+1}}{D_{n+1}} = \nu_{n+1}(t)^T dW(t) \tag{5.13}$$

By way of Ito's formula for d dimensions and some algebra we get through eq. (5.12) that

$$\nu_{n+1}(t) = - \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t)}{1 + \delta_j L_j(t)} \sigma_j(t). \tag{5.14}$$

Further it can be shown that $\mu_n = \sigma_n^T \nu_{n+1}$ and thus that

$$\mu_n(t) = \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t) \sigma_n(t)^T \sigma_j(t)}{1 + \delta_j L_j(t)} \tag{5.15}$$

under the spot measure (see [3] for details). The LMM dynamic under the spot measure is therefore

$$\frac{dL_n(t)}{L_n(t)} = \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t) \sigma_n(t)^T \sigma_j(t)}{1 + \delta_j L_j(t)} dt + \sigma_n(t)^T dW^*(t) \tag{5.16}$$

where $W^*(t)$ is a Brownian motion under the spot measure P^*.

5.2.2 Forward measure

Instead of using the spot measure, that makes asset values discounted by $Z^*(t)$ martingales, we can use the *forward measure*. This measure is the measure that makes traded assets, in this case zero coupon bonds, be martingales when discounted by the price of a zero coupon bond with some future maturity. That is, the forward measure P^{M+1} makes the discounted zeros

$$D_n(t) = \frac{B_n(t)}{B_{M+1}} = \prod_{j=n+1}^{M} (1 + \delta_j L_j(t)) \tag{5.17}$$

martingales. A similar derivation as for the spot measure can be done under the forward measure. That is, the drift $\mu_n(t)$ that makes $D_n(t)$ a martingale can be shown to be

$$\mu_n(t) = -\sum_{j=n+1}^{M} \frac{\delta_j L_j(t)\sigma_n(t)^T \sigma_j(t)}{1 + \delta_j L_j(t)} \tag{5.18}$$

under the forward measure P^{M+1} (again, see [3] for details). The LMM dynamic under the forward measure is therefore

$$\frac{dL_n(t)}{L_n(t)} = -\sum_{j=n+1}^{M} \frac{\delta_j L_j(t)\sigma_n(t)^T \sigma_j(t)}{1 + \delta_j L_j(t)} dt + \sigma_n(t)^T dW^{M+1}(t) \tag{5.19}$$

where $W^{M+1}(t)$ is a Brownian motion under the forward measure P^{M+1}.

5.2.3 Correlations and volatilities

For an alternative, but equivalent, formulation of the LMM dynamics, let $B(t) = (B_1, ..., B_M)(t)$ be a general M-dimensional Brownian motion, i.e for $t_j > t_i$

$$B(t_j) - B(t_i) \sim N(\mathbf{0}, \rho) \tag{5.20}$$

where $\rho = (\rho_{ij})_{i,j=1,...,n}$. Let

$$\frac{dL_i}{L_i} = \mu_i(t)dt + \hat{\sigma}_i(t)dB_i(t), \quad 0 \le t \le T_i, \quad i = 1, ..., M \tag{5.21}$$

where $\hat{\sigma}_i$ is a scalar. This can be shown to be an equivalent formulation of the LMM. We do this by calculating $\frac{dL_i}{L_i} \cdot \frac{dL_j}{L_j}$, that is the instantaneous covariation between $\frac{dL_i}{L_i}$ and $\frac{dL_j}{L_j}$, for the two different formulations. From eq. (5.10) we get

$$\frac{dL_i}{L_i} \cdot \frac{dL_j}{L_j} = \sigma_i^\top \sigma_j = \sum_{k=1}^{d} \sigma_{ik}\sigma_{jk} \tag{5.22}$$

and from eq. (5.21) we get

$$\frac{dL_i}{L_i} \cdot \frac{dL_j}{L_j} = \hat{\sigma}_i \hat{\sigma}_j \rho_{ij}. \tag{5.23}$$

Therefore, we have

$$\hat{\sigma}_i \hat{\sigma}_j \rho_{ij} = \sigma_i^\top \sigma_j, \tag{5.24}$$

a relationship between the parameters of the two types of formulations of the model which is useful in the calibration and implementation of the model. With this formulation, the dynamic under the spot measure can be written as

$$\frac{dL_n(t)}{L_n(t)} = \sum_{j=\eta(t)}^{n} \frac{\delta_j L_j(t) \rho_{jn} \hat{\sigma}_j(t)}{1 + \delta_j L_j(t)} \hat{\sigma}_n(t) dt + \sigma_n(t) dB_n^*(t) \tag{5.25}$$

where $B^*(t)$ is the general Brownian motion discussed above under the spot measure. The dynamic under the forward measure can be written as

$$\frac{dL_n(t)}{L_n(t)} = -\sum_{j=n}^{M} \frac{\delta_j L_j(t) \rho_{jn} \hat{\sigma}_j(t)}{1 + \delta_j L_j(t)} \hat{\sigma}_n(t) dt + \sigma_n(t) dB_n^{M+1}(t) \tag{5.26}$$

where $B^{M+1}(t)$ is the general Brownian motion discussed above under the forward measure.

So far we have discussed the theoretical side of the BS and LMM models. To use these models in simulations we need to decide on a specific discretization (see section 7.5) and we need to estimate the volatilities of and correlations between the rates. This estimation is the topic of the next section.

5.3 Model calibration

We want to estimate the model parameters so that our models best explain reality. This *model calibration* can be done in a variety of ways. We can use some form of historical volatility estimation, we can get implied volatilities and covariances from the current prices of derivatives by using the Black model in reverse, or we can combine these and other ideas to do an optimization of the parameters with error in derivatives prices (model price vs. real price) as objective function.

5.3.1 Historical volatility

One way of estimating the model parameters is by calculating historical volatilities and correlations. These are most easily estimated by a *Simple Moving Average (SMA)*. Assume we have daily quotes of N simple forward rates and a stock index $L_1, ..., L_M, S$

for the last D trading days. That is, we have the data set $\{L_{i,j}\}_{i=1,...,M,j=1,...,D}$ and $\{S_j\}_{j=1,...,D}$. For each rate (index i) and day (index j) calculate the log-return $r_{i,j} = log\left(\frac{\Delta L_{i,j}}{L_{i,j}}\right)$ $(\Delta L_{i,j} = L_{i,j} - L_{i,j-1})$ and calculate log-returns of the stock index in the same way. We then have the data set of returns $\{r_{i,j}\}_{i=1,..,M+1,j=1,...,D-1}$ (the index $i = M + 1$ now denotes the stock index). Now to get an estimate of today's volatilities and correlations pick some day range R, 100 days say, and then calculate, by the standard unbiased estimator, the (annualized) covariance matrix of the returns from the last R days, i.e.

$$\hat{\Sigma}_{ab} = (252)\frac{1}{R-1}\sum_{k=0}^{R-1}(r_{a,D-k} - \bar{r}_a)(r_{b,D-k} - \bar{r}_b), \quad a, b = 1, ..., M+1 \tag{5.27}$$

where $\bar{r}_i$ is the average return of rate i over the last R days

$$\bar{r}_i = \frac{1}{R}\sum_{k=0}^{R-1} r_{i,D-k}, \quad i = 1, ..., M+1 \tag{5.28}$$

From the covariance matrix, we have the volatilities

$$\hat{\sigma}_i = \sqrt{\hat{\Sigma}_{ii}}, \quad i = 1, ..., M+1 \tag{5.29}$$

and the correlations

$$\hat{\rho}_{ij} = \frac{\hat{\Sigma}_{ij}}{\sqrt{\hat{\Sigma}_{ii}\hat{\Sigma}_{jj}}}, \quad i, j = 1, ..., M+1 \tag{5.30}$$

How can we improve this method? We are calculating the historic volatility, but since real market volatilities are not constant we cannot know that the past R day's average volatility is a good estimator for today's volatility. We have a conflict between wanting R as small as possible, so as to not use returns too far into the past, and wanting R as large as possible, so as to not make the variance of the estimator $\hat{\Sigma}_{ab}$ too large. This can be improved by using a more sophisticated historical average method that gives less weight to returns on days further into the past. One such method is the *Exponentially Weighted Moving Average (EWMA)*. This moving average is weighted so that data points in the near past are given more weight than those further back. That they specifically are exponentially weighted gives the EWMA the nice property that it can be updated by a recursive formula. That is, today's EWMA is a function of yesterday's EWMA.

Assume we are at the n-th day day of the historic returns. We now use all of the n returns, but with decaying weights so we use the covariance matrix estimator (annualized)

$$\hat{\Sigma}^*_{ab,n} = (252)\frac{1-\lambda}{1-\lambda^n}\sum_{k=0}^{n-1}\lambda^k(r_{a,n-k} - \bar{r}_{a,n})(r_{b,n-k} - \bar{r}_{b,n}), \quad a, b = 1, ..., N \tag{5.31}$$

where $0 < \lambda < 1$ determines the rate of decay of weights, and $\bar{r}_{i,n} = \frac{1}{n}\sum_{k=1}^{n} r_{i,k}$. A smaller λ means that the weights decay faster.

First of all, we see that this weighting is acceptable in that their sum $\frac{1-\lambda}{1-\lambda^n}\sum_{k=0}^{n-1}\lambda^k = \frac{1-\lambda}{1-\lambda^n}\frac{1-\lambda^n}{1-\lambda} = 1$ (sum of geometric series). Secondly, we see that

$$\begin{aligned}
\text{EWMA}(n) &= \frac{1-\lambda}{1-\lambda^n}\sum_{k=0}^{n-1}\lambda^k x_{n-k} \\
&= \frac{1-\lambda}{1-\lambda^n}x_n + \lambda\frac{1-\lambda}{1-\lambda^n}\sum_{k=1}^{n-1}\lambda^{k-1}x_{n-k} \\
&= \frac{1-\lambda}{1-\lambda^n}x_n + \lambda\frac{1-\lambda^{n-1}}{1-\lambda^n}\frac{1-\lambda}{1-\lambda^{n-1}}\sum_{k=0}^{n-2}\lambda^k x_{n-1-k} \\
&= \frac{1-\lambda}{1-\lambda^n}x_n + \lambda\frac{1-\lambda^{n-1}}{1-\lambda^n}\text{EWMA}(n-1).
\end{aligned}$$

We can therefore update the covariance matrix estimator recursively through the formula

$$\hat{\Sigma}^*_{ab,n} = (252)\frac{1-\lambda}{1-\lambda^n}(r_{a,n} - \bar{r}_{a,n})(r_{b,n} - \bar{r}_{b,n}) + \lambda\frac{1-\lambda^{n-1}}{1-\lambda^n}\hat{\Sigma}^*_{ab,n-1}, \ a, b = 1, ..., M+1. \tag{5.32}$$

Like before, we have the volatilities

$$\hat{\sigma}_{i,n} = \sqrt{\hat{\Sigma}^*_{ii,n}}, \quad i = 1, ..., M+1 \tag{5.33}$$

and the correlations

$$\hat{\rho}_{ij,n} = \frac{\hat{\Sigma}^*_{ij,n}}{\sqrt{\hat{\Sigma}^*_{ii,n}\hat{\Sigma}^*_{jj,n}}}, \quad i, j = 1, ..., M+1 \tag{5.34}$$

.

5.3.2 Implied volatility

Another way of estimating volatilities and correlations is by using today's prices of derivatives. The idea is to use the Black-Scholes formula for stocks, and the Black formula for caplets in reverse. That is, if we know today's price of for instance a European call option written on the stock S, we can find the implied volatility of S by solving the Black-Scholes formula for the volatility. Likewise, if we know today's price of a caplet written on some rate L, we can find the implied volatility of L by solving the Black formula for the volatility.

If we want to find implied correlations between rates we need to know the price and have a closed form solution for the price of some derivative that depends on more than one rate.

One well known problem in implied volatility estimation is that the implied volatility varies with the strike price of the derivative from which the volatility is derived. For stock options, this phenomenon is known as "volatility smile". This is because implied volatilities typically are higher for stock options that are "in the money" (for a call option: current stock price higher than strike) or "out of the money" (current stock price lower than strike) and typically lower for "at the money" options (current stock price equal to strike), so that a graph of implied volatility vs. option strike price looks like a smile.

The estimation of volatilities and covariance matrices through this method is a field of study in itself and we shall not go into further details here.

5.3.3 Optimization

Another possible calibration approach is to use some sort of optimization scheme. With a given set of volatilities and correlations, we can calculate prices of derivatives. If we know the actual market prices of these derivatives, we can do an optimization with the volatilities and correlations as target variables and price errors, that is the difference between market prices and model prices, as object function (to be minimized). To do this, we need to create an error function by a sensible combination and weighting of the price errors of the individual derivatives. This optimization approach can obviously be done in countless different ways, depending on the type of optimization and the choice of error function. We shall not discuss this approach further here.

6 Data analysis and model assumption tests

The data set considered in this thesis consist of OSEBX (Oslo Børs Hovedindeks) and Eurostoxx 50 closing price quotes for each trading day in the period 02.01.2007 - 29.02.2012, along with swap rate quotes in the Norwegian (NOK) and European (EUR) interbank market for the same period.

Since our term structure data is in the form of par swap rates, we have used a bootstrapping method to build a zero curve (and hence a forward rate curve) with equally spaced tenor dates. This is explained in section 7.1. Using the LMM notation we get from this procedure a forward rate curve for each day of historical data consisting of the 19 rates $L_1, ..., L_{19}$ with $\delta_i = T_{i+1} - T_i = 0.25$, i.e. the tenor dates are equally space with 3 months between. Many of the resulting forward rates are perfectly correlated (discussed in section 7.3), so the statistics and tests on the fixed income data below are done on the forward rates L_1, L_4, and L_8.

Table 6.1: *Means and standard deviations of daily log-returns of Eurostoxx 50 (E) and L_1,L_4, and L_8 for European market, annualized at 252 trading days in a year.*

	E	L_1	L_4	L_8
Mean	-0.0969	-0.2059	-0.2546	-0.1434
St.D.	0.2775	0.2368	0.4388	0.2705

Table 6.2: *Means and standard deviations of daily log-returns of OSEBX (O) and L_1,L_4, and L_8 for Norwegian market, annualized at 252 trading days in a year.*

	O	L_1	L_4	L_8
Mean	-0.0072	-0.0946	-0.0946	-0.0743
St.D.	0.3177	0.3508	0.2880	0.1796

We begin by looking at some simple statistics of our data sets. In tables 6.1 and 6.2 the standard deviations of log-returns over the whole period, which estimate volatility in log-normal models, of stock index and rates are shown. In figures 6.1 and 6.2 the daily volatilities, as found using the EWMA estimator are shown.[1]

The Norwegian stock market is overall more volatile than the European. Further, the 3 month rate L_1 is substantially more volatile in the Norwegian market than in then European, and the 1 and 2 year rates L_4 and L_8 are substantially more volatile in the European market. So the Norwegian capital market have larger volatility for short term debt, a property not shared by the European market in which the 1 year rate has nearly twice the volatility of the 3 month and 2 year rates.

The time period from 24.07.2008 (day 400) to 07.05.2009 (day 600) is characterised by

[1] To make comparison between figures and results easier, in all figures we represent the days of our historical data by their ordering. In other words, day 1 represents the date 02.01.2007 and day 1327 represents the date 29.02.2012.

very high volatilities in both markets, with a spike around November-December 2008 (day 500). This period was characterised by plummeting stock markets and interest rates, bankrupt financial institutions, and great uncertainty. We see that the European and Norwegian market volatilities largely agree, but it seems the European stock market was "hit harder", compared to the 2008 peak, by the two later volatility peaks around 28.04.2010 - 07.07.2010 (day 850-900) and 24.06.2011-23.01.2012 (day 1150-1300).

In table 6.3 we have the correlations of log-returns and stock index over the whole period. Not surprisingly, the correlations between rates, and between rates and equity is substantial. Further we see that in both markets the stock index correlates more with long term rates than short term rates, and that the correlations are overall larger in the European market than in the Norwegian.

Table 6.3: *Correlations between daily log-returns of stock index and L_1,L_4, and L_8 for European market (left) and Norwegian market (right). (Eurostoxx 50 = E, OSEBX = O)*

	E	L_1	L_4	L_8
E	1			
L_1	0.2224	1		
L_4	0.3818	0.7408	1	
L_8	0.4471	0.4336	0.6659	1

	O	L_1	L_4	L_8
O	1			
L_1	0.1553	1		
L_4	0.2322	0.3864	1	
L_8	0.3035	0.3103	0.5133	1

6.1 Equity, Black-Scholes

The assumptions of the BS and LMM models must be tested against real world historical data. For the BS model we had that

1. There are no dividends. This is obviously not true for stocks in general, but the stock indices are dividend adjusted so this is indeed true for our historical stock prices.

2. There are no transaction costs. Again, this is obviously not true, so continuous hedging schemes are not possible in real life trading.

3. We can buy and sell any fraction of any asset (perfect divisibility). There is no obvious way of achieving the equivalent of buying one half of a stock, let alone $\frac{\pi}{246}$ of one, or any other conceivable fraction. However, the people applying hedging strategies are generally not in the business of buying one or two stocks or options, rather thousands at a time, so there will be little error in assuming perfect divisibility.

4. The returns of stock are normally distributed. This is one of the big debates of mathematical finance. On the one hand the assumption of normal returns is the natural choice. On the other hand it is simply empirically wrong.

Firstly, why is normally distributed returns the natural choice? This is a direct consequence of the Central Limit Theorem (CLT) (see appendix A.10). The CLT says that the sum of n iid. random variables is a normally distributed variable. Since log-returns

Figure 6.1: *Daily volatilities for forward rates and equity in the European market, calculated with the EWMA estimator.*

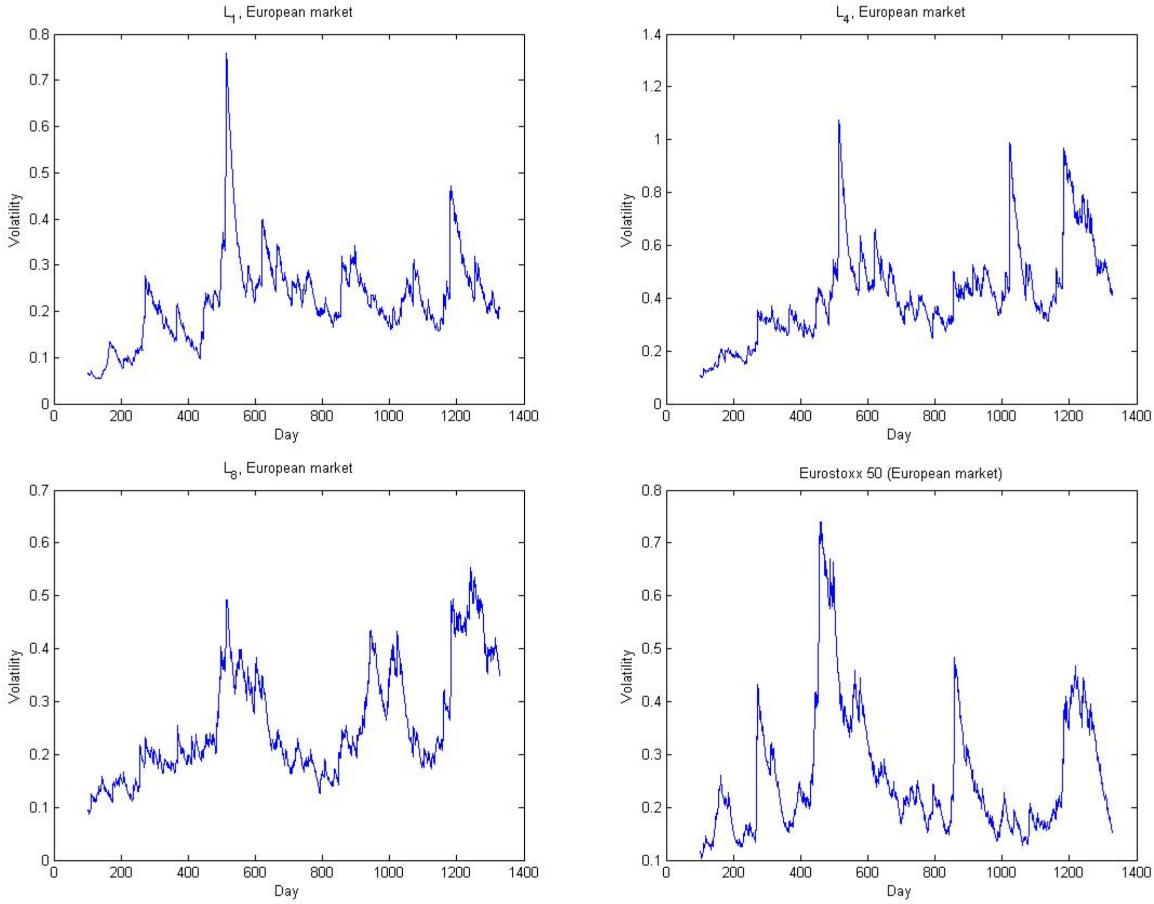


Figure 6.2: *Daily volatilities for forward rates and equity in the Norwegian market, calculated with the EWMA estimator.*

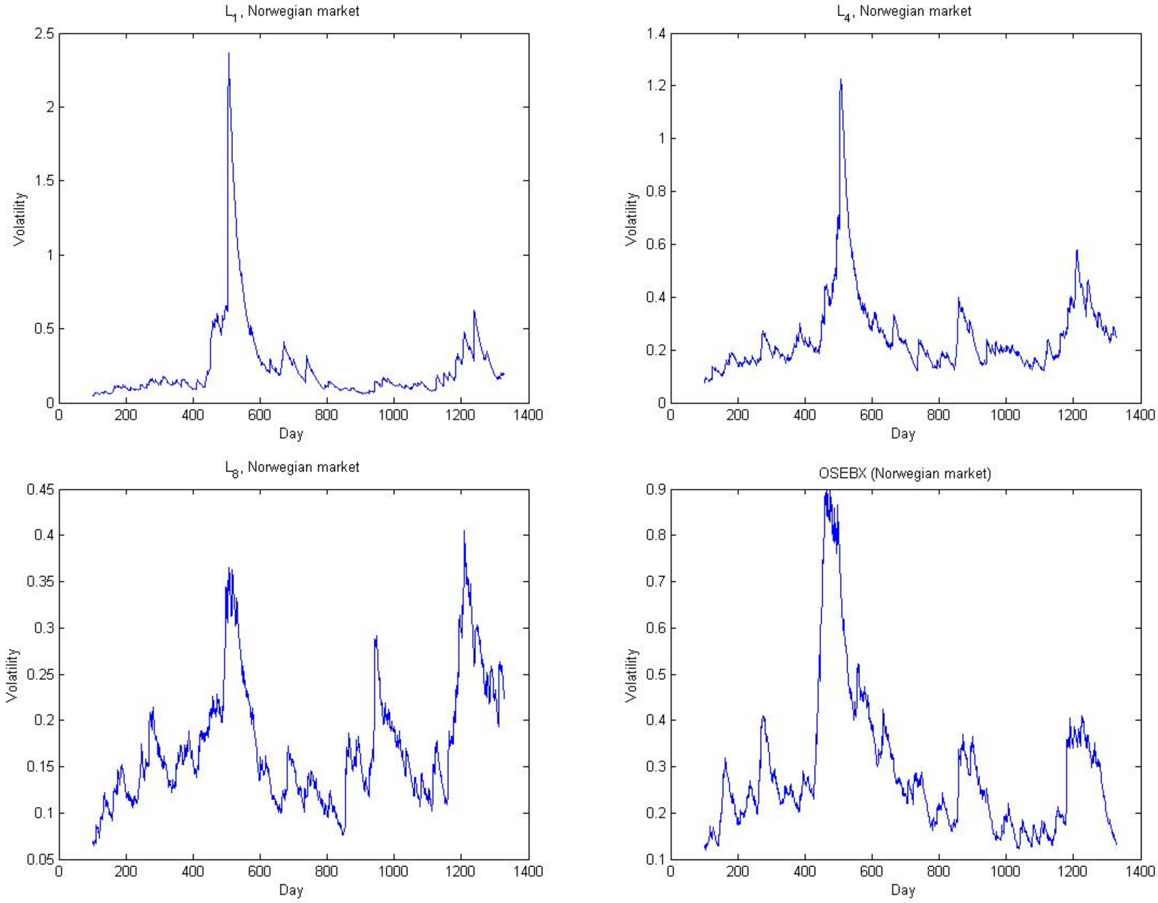
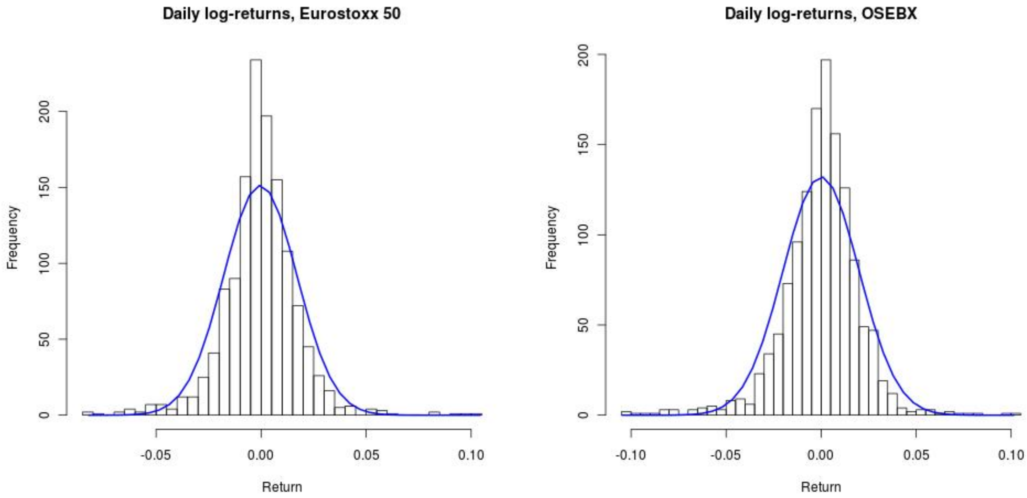


Figure 6.3: *log-returns of the Eurostoxx 50 and OSEBX indices for the whole period, along with the corresponding normal distributions.*



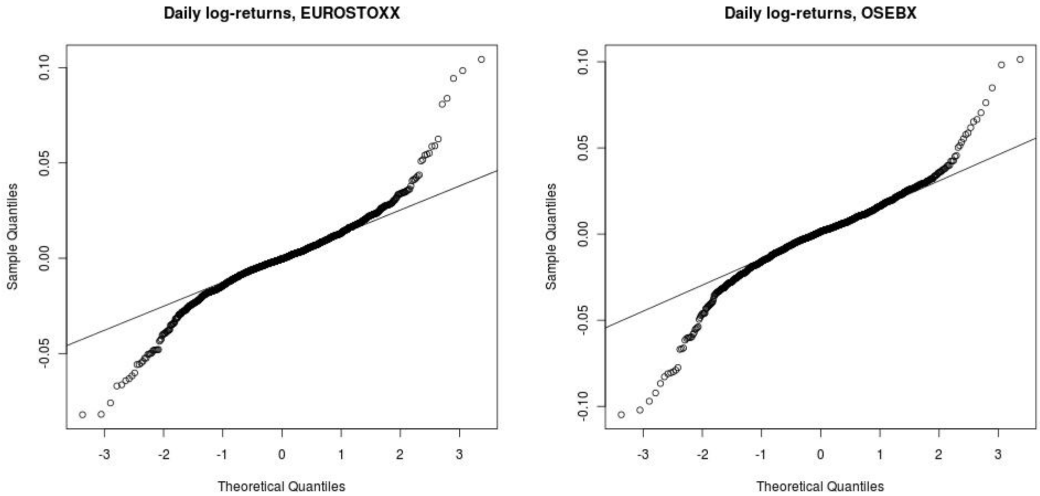
are time-additive, a daily return is the sum of the return over smaller time-periods. Letting the number of these smaller time periods go to ∞ we achieve the limit in the CLT and the returns are thus normally distributed. And from the EMH we assume that the asset price moves are completely random and independent of earlier price moves. Therefore, the assumption of normally distributed returns is really a consequence of the efficient market hypothesis (EMH). If strong form EMH was true, that all information is instantly incorporated into the asset price, we would have normally distributed returns.

The problem, of course, is that real life stock returns often depart from normality. When comparing a histogram of log-returns of daily stock price changes with a Gaussian density function with the same mean and standard deviation (see figure 6.3) it seems the normal distribution underestimates the probability of the smallest returns (tall spike around the mean) and the largest returns (fat tails) and overestimates the probability of intermediate returns.

A QQ-plot is a plot where the theoretical quantiles of the normal distribution are plotted against the sample quantiles of our data set. QQ-plots of daily returns over the period are shown in figure 6.4. If the observed returns were normal, they would lie along the straight line.

It seems therefore, that stock returns have non-normal behaviour. However, these historical data include different types of markets. Perhaps non-normal returns is a characteristic of turbulent markets. It is possible that we have some periods with close to normal returns and other "less normal" periods. We would like to get some idea of which periods are "most normal", i.e. which periods have returns that are closest to being normally distributed. To achieve this, we need to quantify the idea of "closeness to normality". This will be done with hypothesis testing for 100 daily returns at a time. That is, we

Figure 6.4: *QQ plots of log-returns of the Eurostoxx 50 (left) and OSEBX (right) indices for the whole period.*



calculate for each consecutive 100 day period some test statistics and plot the result. In doing this we can see in which periods normality is a good and, more importantly, in which periods normality is a bad assumption.

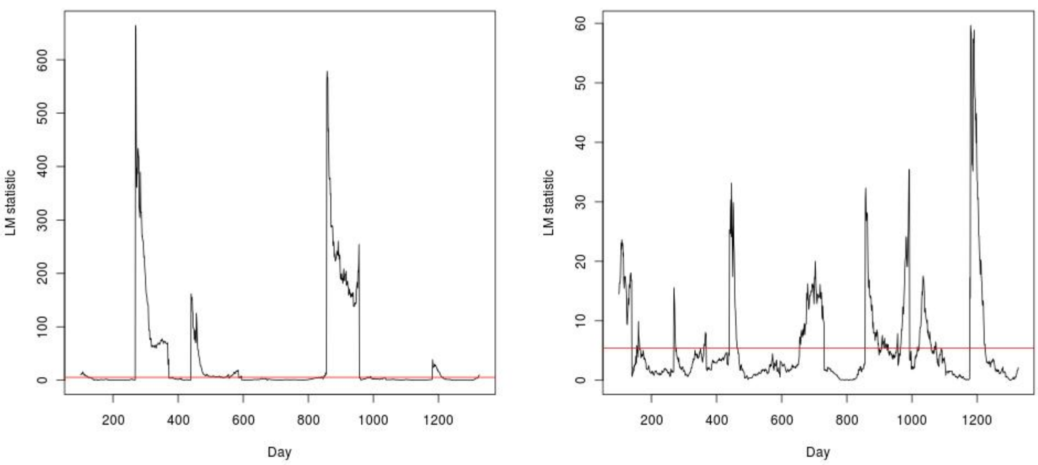
6.1.1 Jarque-Bera Lagrange Multiplier test

The Jarque-Bera Lagrange multiplier (LM) test for normality is a popular normality test of economic time series. For a collection of n returns $r_1, ...r_n$, the test statistic is defined as

$$LM = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right) \tag{6.1}$$

where $S = \frac{\frac{1}{n}\sum(r_i-\bar{r})^3}{(\frac{1}{n}\sum(r_i-\bar{r})^2)^{3/2}}$ is the sample skewness and $K = \frac{\frac{1}{n}\sum(r_i-\bar{r})^4}{(\frac{1}{n}\sum(r_i-\bar{r})^2)^2}$ is the sample kurtosis. Under the null hypothesis that returns are normally distributed, which implies that the true skewness and kurtosis are 0 and 3 respectively, LM is χ^2 distributed with 2 degrees of freedom in the asymptotic limit $n \to \infty$. However, for small sample sizes like $n = 100$ the distribution of LM is not well approximated with the χ^2 distribution. In [13], a distribution table for LM is created by Monte Carlo simulation, using 10^7 replications. We calculated the LM statistic for each consecutive 100 day period in the log-return time series for OSEBX and Eurostoxx 50. In figure 6.5 we have plotted these LM values, along with the critical value at the 5% significance level, extracted from the distribution table. We see that both the OSEBX and Eurostoxx log-returns are in most 100 day periods deemed normal by this test, but for the 100 day periods including extreme returns we have vary tall spikes in the LM statistic. A single extreme return have a large influence on the LM statistic.

Figure 6.5: *The Jarque-Bera Lagrange Multiplier normality test statistic for consecutive 100 day periods along with critical value at 5% significance level. Eurostoxx 50 (left) and OSEBX (right). On the x-axis, the 'Day' number refers to the last day in the 100 day period for which the test statistic is calculated.*



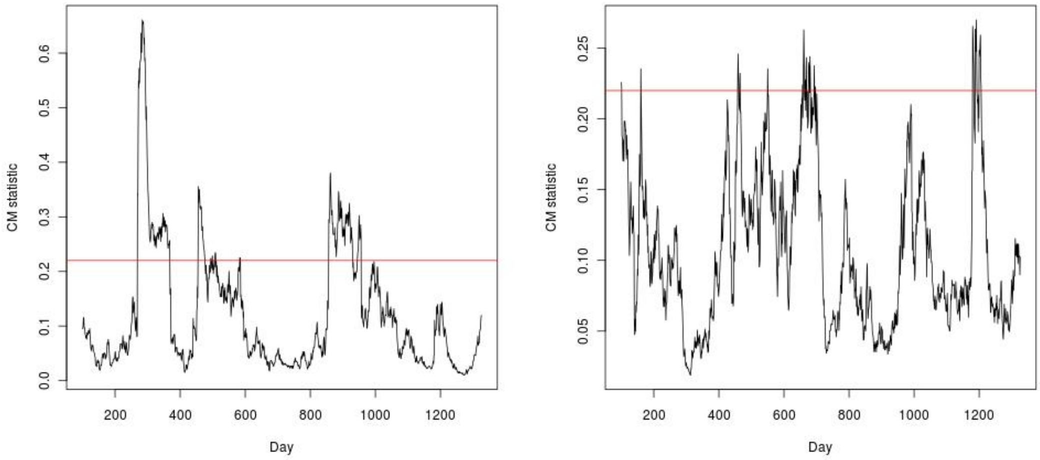
6.1.2 Cramer-von Mises test

The Cramer-von Mises test is a more general test for the closeness of an *empirical distribution* to some theoretical distribution. We define the CM test statistic as

$$\mathrm{CM} = \frac{1}{12n} + \sum_{i=1}^{n} \left[\frac{2i-1}{2n} - F(r_i) \right] \tag{6.2}$$

where $r_1, ..., r_n$ are the log-returns for some time period. We calculated the Cramer-von Mises (CM) test statistic for each consecutive 100 day period. In figure 6.6 the CM test statistic is plotted along with the critical value at the 5% significance level, taken from [10]. Comparing this with the results for the LM statistic, the CM statistic is less strict in that a single extreme return has less influence. For the Eurostoxx returns we see that the two test generally agree on which time periods are normal. For the OSEBX returns we see that the two tests agree on what periods are less normal than others but the CM rejects the null hypothesis of normality in fewer of the 100 day periods.

Figure 6.6: *The Cramer-von Mises test statistic for the consecutive 100 day periods along with critical value at 5% significance level. Eurostoxx 50 (left) and OSEBX (right). On the x-axis, the 'Day' number refers to the last day in the 100 day period for which the test statistic is calculated.*



6.2 Forward rates, LMM

Most of the assumptions of the Black-Scholes model are assumptions of the LMM as well. We have perfect divisibility and no transaction costs or taxes. Further, we have normality of returns, but in the LMM, the log-returns/changes are assumed to be *multivariate* normal. The same conceptual discussion of normality, as for the Black-Scholes equity case, is valid here. However, we now need to be able to test for *multivariate normality.*

Can we find some multivariate equivalent of the 100 day moving average of "closeness to normality" introduced in the preceding section? One idea would be to check for normality in the log-changes of the forward rates one by one, since the rates being separately normal is a necessary condition for them to be multivariate normal. However, individual normality is obviously not a sufficient condition so we would like some multivariate test.

6.2.1 Mardia's test

We can generalize to more dimensions the quantities of skewness and kurtosis. Just like the d-dimensional mean vector $\vec{\mu}$ is the generalization of the mean μ and the $d \times d$ covariance matrix Σ is the generalization of the variance σ^2, we have d-dimensional skewness and kurtosis. Mardia's test is, like the Jarque-Bera test in the one-dimensional case, a test of skewness and kurtosis but in this case the multidimensional generalizations. We test the normality of the log returns of the forward rates L_1, L_4, and L_8. Letting $L_{i,j}$ denote the forward rate L_i at the j-th day of our historical data, we therefore have the sample vectors $\mathbf{r_1}, \mathbf{r_2}, ..., \mathbf{r_{1326}}$ where

$$\mathbf{r_i} = \begin{pmatrix} \log \frac{L_{1,i}}{L_{1,i-1}} \\ \log \frac{L_{4,i}}{L_{4,i-1}} \\ \log \frac{L_{8,i}}{L_{8,i-1}} \end{pmatrix} \tag{6.3}$$

Using [6], for a sample of n p-dimensional sample vectors $\mathbf{r_i}$ we first define

$$g_{ij} = (\mathbf{r_i} - \mathbf{\bar{r}})^T \mathbf{\Sigma}^{-1} (\mathbf{r_j} - \mathbf{\bar{r}}), \tag{6.4}$$

where $\mathbf{\Sigma} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{r_i} - \mathbf{\bar{r}})(\mathbf{r_i} - \mathbf{\bar{r}})^T$ is the covariance matrix. We then define the one-tailed test statistic

$$b_{1,p} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}^3 \tag{6.5}$$

and the two-tailed test statistic

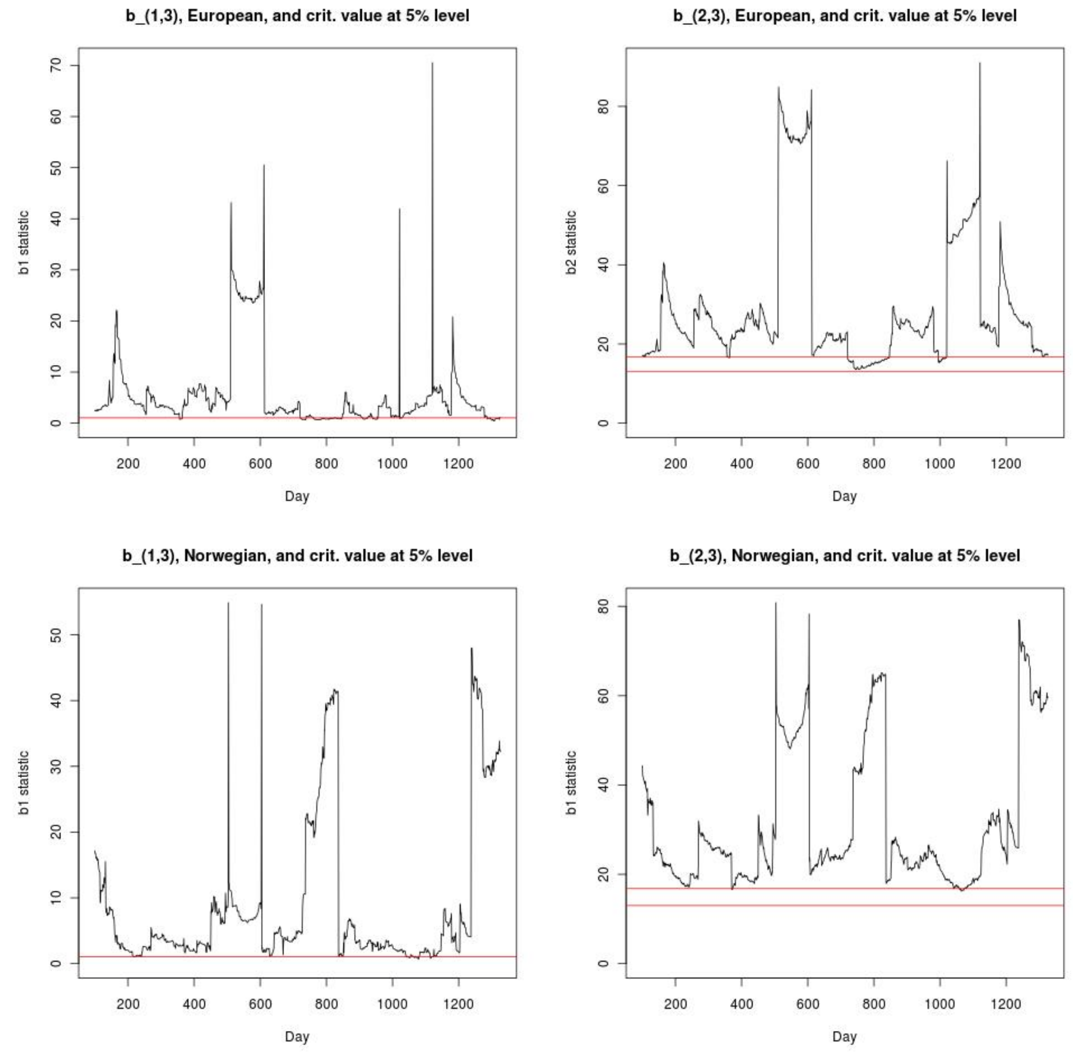
$$b_{2,p} = \frac{1}{n} \sum_{i=1}^{n} g_{ii}^2 \tag{6.6}$$

In [6], tables of critical values for $b_{1,p}$ and $b_{2,p}$ are given for different values of n,p, and different levels of significance.

For our calculations, we have $p = 3$ and $n = 100$, since we calculate the test statistics for each 100 day period. In figure 6.7, the test statistics for each 100 day period is plotted along with the critical value(s) at the 5% significance level. We see that the skew and kurtosis of forward rates are significantly non-normal, by Mardia's test, for most time periods. The forward rates are least normal in loosely the same periods as for the stocks returns, notably in 2008 (day 500 to day 600).

It should be noted that the normal distribution is not uniquely defined through kurtosis and skewness. That is, a data set that passes a normality test based on these statistics is not necessarily normal.

Figure 6.7: *Plotted values for each day of the Mardia test for 3-dimensional multivariate normality of log-returns of forward rates L_1, L_4, and L_8. Test for skewness ($b_{1,3}$) and kurtosis ($b_{2,3}$). European and Norwegian markets. On the x-axis, the 'Day' number refers to the last day in the 100 day period for which the test statistic is calculated.*



7 Implementation

7.1 Bootstrapping, building the zero curve

Our data set consist of par swap rates with 5 different maturities. The quoted swap rates are the fixed rate on a quarterly swap of the quoted maturity. For example, the 1 year swap rate is the fixed rate in a fixed for floating swap agreement with cash-flows every three months, lasting for 1 year (payments at 3,6,9 and 12 months). We need to build a forward rate curve using only this information. We have equation (5.7) for getting forward rates from the zero curve. The problem is therefore reduced to building an equally spaced zero curve from the par swap rate data. We know the relationship between swap rates and zeros from equation (2.6). Changing notation so that $S(T_n)$ denotes the quoted swap rate with maturity T_n we have

$$S(T_n) = \frac{1 - Z(0, T_n)}{0.25 \sum_{i=1}^{n} Z(0, T_i)} \tag{7.1}$$

where the maturities are $T_i = 0.25i$.

With this notation we have, for each trading day of the historical data, the rates

$$S(T_1), S(T_4), S(T_8), S(T_{20}), S(T_{40}).$$

I shall call $T_1, T_4, T_8, T_{20}, T_{40}$ the 'familiar maturities' and the other T_is 'unfamiliar'. The first equation ($n = 1$) is

$$S(T_1) = \frac{1 - Z(0, T_1)}{0.25 Z(0, T_1)} \tag{7.2}$$

which gives us

$$Z(0, T_1) = \frac{1}{1 + 0.25 S(T_1)}, \tag{7.3}$$

so we immediately have the first discount factor. However, the $n = 4$ equation is

$$S(T_4) = \frac{1 - Z(0, T_4)}{0.25(Z(0, T_1) + Z(0, T_2) + Z(0, T_3) + Z(0, T_4))} \tag{7.4}$$

so it is clear that we do not have enough information to build the complete zero curve $Z(0, T_i)_{i=1,\ldots,40}$. We therefore need to apply some kind of *bootstrapping* scheme to the data. We do this by assuming some interpolation rule to determine the zeros for unfamiliar maturities. In other words, we assume that we can express the zeros of unfamiliar maturities as a function of zero values, z and x, of the closest preceding and succeeding familiar maturities, respectively. So we have

$$S(T_4) = \frac{1 - x}{0.25(z + f_2(z, x) + f_3(z, x) + x)}, \tag{7.5}$$

where in this case $z = Z(0, T_1)$ and the unknown $x = Z(0, T_4)$

Now, since we have already found the zero $Z(0, T_1)$, $z = Z(0, T_1)$ is known. The right hand side of (7.5) is therefore a function of x, once we have decided on what interpolation functions f_2 and f_3 to use. That is, by finding the x value that makes the right hand side equal the known value $S(T_4)$, we have found a value for $Z(0, T_4)$ and hence, through the functions f_2 and f_3, a value for $Z(0, T_2)$ and $Z(0, T_3)$.

What kind of interpolation should we use? Since this point in the process involves quite a bit of guesswork we follow the simplest path, namely linear interpolation. In other words we use the interpolation functions

$$f_i(x) = z + \frac{x - z}{\tau(+, i) - \tau(-, i)}(T_i - \tau(-, i)) \tag{7.6}$$

where $\tau(-, i)$ and $\tau(+, i)$ is the closest preceding and succeeding familiar maturity to T_i, respectively (the maturities corresponding to the zeros z and x). The function τ is defined so that if T_i is one of the familiar maturities, $T_i = \tau(+, i)$. This interpolation choice gives us

$$S(T_4) = \frac{1 - x}{0.25(z + z + \frac{x-z}{T_4-T_1}(T_2 - T_1) + z + \frac{x-z}{T_4-T_1}(T_3 - T_1) + x)}, \tag{7.7}$$

For the sake of general applicability, we should find x through some sort of numerical scheme. (We have done this using the spreadsheet software LibreOffice Calc). With the linear interpolation choice, of course, we can also solve the equation analytically. Some simple algebra gives us

$$Z(0, T_4) = x = \frac{1 - 0.25S(T_4)Z(0, T_1)(3 - \frac{T_2-T_1}{T_4-T_1} - \frac{T_3-T_1}{T_4-T_1})}{1 + 0.25S(T_4)(1 + \frac{T_2+T_3-2T_1}{T_4-T_1})} \tag{7.8}$$

We have now found values for the first four zeros and we continue in the same way to find the rest of the zero curve, starting from equation (7.1), inserting known zeros and interpolation functions and solving for the zero of longest maturity. Such a simple bootstrapping procedure should, however, be used with caution. At the long end of the zero curve the quoted swap rate maturities are far apart so the interpolations are spanning long time periods. We also know that the short end of the term structure sometimes have characteristics that are not well captured by piecewise linearity. Improving the bootstrapping is a field of study in itself, but we shall not proceed further here.

For the purposes of our simulations, we applied the proposed scheme to the par swap rates for maturities up to 5 years. The result is a zero curve for quarterly maturities from 3 months up to 5 years.

7.2 Portfolios

To main object of this thesis is to test the different financial risk measures by doing a backtesting procedure on the historical data for some different types of portfolios. We shall explore whether there are performance differences of the risk measures between some low risk, unleveraged portfolios and high risk, leveraged portfolios. We shall consider five portfolios in which a wealth of 100 are leveraged and invested in different ways. The portfolios will be varying from unleveraged and near risk-free to highly leveraged and very risky.

Conservative portfolio This portfolio has placed the whole wealth in 1 year zeros

Moderate portfolio This portfolio is a 50/50 divide between stocks and 3 year zeros.

Equity portfolio This portfolio has placed the whole wealth in equity.

Leveraged equity portfolio This portfolio borrows 400 (achieving a debt to value of 80%) by short selling 3 year zeros and spend the resulting 500 on equity.

Leveraged interest rate derivatives portfolio This portfolio borrows 400 (achieving a debt to value of 80%) by short selling 3 year zeros and spend the resulting 500 on caps (starting from T_1, four quarterly caplets) with strike at the current rate

$$L_0(0) = \frac{1 - Z_1(0)}{\delta_0 Z_1(0)} \tag{7.9}$$

This portfolio thus bets on a rise in interest rates.

Noting that we have four traded assets, namely

- 1 year zero
- 3 year zero
- Stock
- Cap

we can summarize the portfolio holdings in table form, as shown in table 7.1.

Table 7.1: The value of holdings of the four assets in each portfolio.

	1 year zero	3 year zero	Stock	Cap
Conservative portfolio	100	0	0	0
Moderate portfolio	0	50	50	0
Equity portfolio	0	0	100	0
Lev. equity portfolio	0	-400	500	0
Lev. IR deriv. portfolio	0	-400	0	500

We have not included the special kind of risks associated with holding a leveraged portfolio. In our set-up, a leveraged portfolio will be heavily correlated with the corresponding unleveraged portfolio since the value of a loan, i.e. a bond, are subject to only small changes compared to equity or derivatives and the costs of financial distress are not accounted for. So we expect little difference in VaR and ES performance between the moderate portfolio, the equity portfolio and the leveraged equity portfolio. The debt to value of 80% in the leveraged portfolios is therefore arbitrarily chosen and is risk measure-wise simply corresponding to holding some proportion of bonds to equity/IR derivatives.

7.3 Modelling scheme

Because of the interpolation, many of the resulting forward rates will be perfectly correlated. More specifically, the rates L_1, L_2, L_3 are perfectly correlated, as are the rates L_4, L_5, L_6, L_7, and the rates $L_8, ..., L_{20}$. The choice of modelling all of them is one of convenience rather than necessity. We shall therefore in our modelling use three factors, i.e. three Brownian motions, corresponding to these three groups of rates. As mentioned in section 6, for the purposes of data analysis we therefore investigate the rates L_1, L_4, and L_8 only. The covariance matrix to be estimated is therefore the 4×4 matrix of covariances of the log-returns of these three rates and the log-returns of the stock market index.

The VaR and ES time period of choice is 5 days, and we calculate the 5% and 1% VaR and ES. All portfolios are statically weighted (in number of assets held, not in asset value terms) over these 5 days.

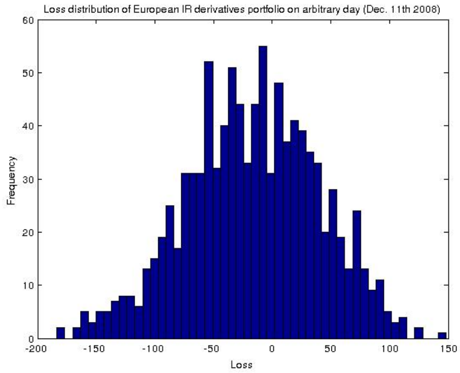
We will, at each day in our historical data, consider the risk of the portfolio as if it is initialized on that day. In other words, we will at each day $d_i, i = 1, ..., N$ initialize and value our portfolio at its starting point, with value of 100, holding the above discussed bonds, derivatives, and stock so that $T_0 = 0$ is set to be on the day d_i.

We will at each date d_i value the portfolio $V_r(d_i)$ and then forecast with 1000 simulations the portfolio value $V_f(d_{i+5})$ one trading week (5 days) into the future to get a simulated distribution of the portfolio loss $L_f(d_i) = V_r(d_i) - V_f(d_{i+5})$. From this we calculate the 1% and 5% VaR and ES. We then find the real portfolio value $V_r(t_{i+5})$ at t_{i+5} based on the historical data, i.e. what actually happened to the forward rates and equity at t_{i+5}. We then have the real loss $L_r(t_i) = V_r(t_i) - V_r(t_{i+5})$ to be compared with the VaR and ES numbers. We begin our simulations at the 101st day of the historical data, on May 24th 2007. The preceding 100 days are used only for calculating historical volatility at May 24th.

For each day d of our historical data, beginning on the 101st day and ending on the 1322nd, and each portfolio we do the following

- Calculate the covariance matrix, based on the EWMA of all preceding historical data
- Value the assets today (derivatives pricing/"$\mathbb{Q}$-world") and calculate the number held of each asset.

Figure 7.1: *Example of simulated 5 day loss distribution, European IR derivatives portfolio on December 11th 2008.*



- Do 1000 simulations of the underlyings (forward rates and stock index) 5 days into the future under the real probability measure ("$\mathbb{P}$-world").

  For each simulation:

  - Value the four assets, and hence the portfolio at the forecasted 5 days ahead point (derivatives pricing/"$\mathbb{Q}$-world").
  - Calculate the forecasted 5 day loss.

- Using the historical data for the underlyings five days ahead ($d + 5$), value the portfolio (derivatives pricing/"$\mathbb{Q}$-world").
- Calculate the real 5 day loss.

The VaR and ES are calculated from the loss distribution resulting from the 1000 simulated 5 day losses. An example of one such loss distribution can be seen in figure 7.1.

For the EWMA estimate of the covariance matrix, we used $\lambda = 0.94$ as proposed in [7].

7.4 Volatility structure

We use the EWMA method of historical volatilities to calculate the covariance matrix (see section 5.3.1). We assume a stationary volatility structure. $\sigma_i(t) = \sigma(T_i - t)$ We also assume constant volatilities between tenor dates. That is, $\sigma_n(t) = \sigma(T_n - t) = \sigma(T_i)$, for $(T_n - t)$ in the interval $[T_i, T_{i+1})$. For example, the rate $L_1(0.1) = L(0.1, 0.25, 0.5)$, i.e. the rate at time $t = 0.1$ for borrowing from time $T_1 = 0.25$ to $T_2 = 0.5$ has the volatility and correlations like the rate $L_1(0) = L(0, 0.25, 0.5)$, for which we have historical data.

As previously mentioned, since many of the rates are perfectly correlated in the historical data, we use four factors, i.e. a four-dimensional multivariate Brownian motion. One for the rates L_1, L_2, L_3, one for the rates L_4,...,L_7, one for the rates $L_8, ..., L_{20}$, and one for the stock S. Therefore, $W(t_{i+1}) - W(t_i) = (W_1(t_{i+1}) - W_1(t_i), ..., W_4(t_{i+1}) - W_4(t_i))$ has the distribution $\sqrt{t_{i+1} - t_i}Z$, where $Z \sim N(\mathbf{0}, \rho)$ and ρ is the correlation matrix of L_1, L_4, L_8, S.

7.5 Model discretization

For the discretization of the BS model, with the set of simulation dates $t_1, t_2, ...$ we use an Euler scheme on $\log S$ and get

$$\hat{S}(t_{i+1}) = \hat{S}(t_i)exp\left(\left[\mu(\hat{S}(t_i), t_i) - \frac{1}{2}\hat{\sigma}^2\right][t_{i+1} - t_i] + \sqrt{t_{i+1} - t_i}\hat{\sigma}^2 Z_4\right) \tag{7.10}$$

where $\hat{\sigma}$ is the volatility of the stock and Z_4 is the fourth element of the random vector $Z = (Z_1, ..., Z_4) \sim N(\mathbf{0}, \rho)$

For the LMM discretization , with the set of simulation dates $t_1, t_2, ...$, we use an Euler scheme on $\log L_n$ and get

$$\hat{L}_n(t_{i+1}) = \hat{L}_n(t_i)exp\left(\left[\mu_n(\hat{L}_n(t_i), t_i) - \frac{1}{2}\hat{\sigma}_n^2\right][t_{i+1} - t_i] + \sqrt{t_{i+1} - t_i}\hat{\sigma}_n^2 Z_\alpha\right), \quad n = 1, ..., M \tag{7.11}$$

where $\hat{\sigma}_n$ is the volatilities of the second LMM formulation in section 5.2.3 and Z_α is the α-th element in the random vector $Z = (Z_1, ..., Z_4) \sim N(\mathbf{0}, \rho)$, where $\alpha = 1, 2, 3$ depending on the rate so that

- $n = 1, 2, 3 \to \alpha = 1$
- $n = 4, 5, 6, 7 \to \alpha = 2$
- $n = 8, ..., 20 \to \alpha = 3.$

These two formula are used to create forward rate paths $(L_n(0), L_n(t_1), ...)$, $n = 1, ..., 20$ and stock paths $(S(0), S(t_1), ...)$

7.6 Forward "$\mathbb{P}$-world" simulation

To calculate VaR and ES we need the *real* probability distribution of portfolio value at the future time. Therefore, simulating forward in time we need to use the *real statistical* probability measure, i.e. the *real* drift of the stock indices and forward rates. Note that "real" in this refers to drift under the real probability measure, not drift corrected for inflation.

Estimating the real drift of equity is in the Black-Scholes framework a question of estimating μ in eq. (5.1). Firstly, we obviously cannot use the mean return $\bar{r}$ over any short historical period of time, like a year. If the stock market has crashed in that time period, we will have $\bar{r} < 0$, and a $\mu < 0$ does not make sense. If we could use the latest historical data to calculate μ in this way, that would be a direct contradiction of the weak form efficient market hypothesis. The μ in the Black-Scholes model does not signify the specific trend of a stock or index. The rate of return μ is the expected return coming part from the time value of money, i.e. expected risk free return r and the risk premium $\lambda\sigma$. We cannot, expect in a very few historical cases, have a negative r and we can certainly not have a negative risk premium. The only sensible way of estimating μ therefore, is to look at a large time-horizon, typically decades, to see the *general* trend of the stock or market index.

Now, through the relation $\mu = r + \lambda\sigma$ and the knowledge that r changes through time, what we really should be looking at is the general trend of the risk premium $\lambda\sigma$. If we have data for the short rate for near risk free fixed income products, like government bonds, in the same long time period we can use this as a proxy for r and thus calculate the historical risk premium. In this thesis, however, we are limited by the data, which is recent, so we shall use the naive method of letting μ be as expected by the long term trend of stock markets, which is about 10% (for the American stock market) depending somewhat on the source. This method is overly simplistic. We therefore restrict ourselves to calculating short term VaR and ES numbers, daily and weekly. Over such short periods, the drift term is insignificant compared to the diffusion term. It could even be argued that the drift term could be neglected all together, so we shall not press the issue of real drift of the stock indices any further.

We also need to estimate the real drift of forward rates, i.e. estimate each μ_n in equation 5.10. Interest rates have very different drift characteristics than stocks. Interest rates are mean reverting, so do for obvious reasons not have expected exponential growth, like stocks do. The potential drift of interest rates will be short term phenomena, and thus extremely hard to estimate in any meaningful way. For our purposes of calculating short term VaR and ES, the real drift is, like for equity, negligible, so we will simply assume zero real drift.

7.7 Derivatives pricing/"$\mathbb{Q}$-world simulation"

In one replication of the real world simulation, if a stock or forward rate increase by some amount, how has the price of our derivatives changed? How do we find the new portfolio value? Since prices of derivatives may depend on the underlying in non-trivial ways, this may be a very hard problem. Intuitively, we simply need to revalue our portfolio, using some available method like MC simulation. A double MC simulation, however, is very time consuming. E.g. if we use 10000 replications for the real-world forward simulation and 10000 replications for pricing a derivative (under some suitable measure) within each replication, we need 100 million replications, which even for modern computers will take far too long. There are ways to do these simulations more efficient. One can for instance use some kind of variance reduction technique in the MC simulation to speed things up. But for backtesting on long time series, it will still be infeasible to use MC simulation. The delta-gamma approximation might be useful, but this depends on the estimation of the derivative's delta and gamma being simpler to calculate than simulating the derivative's change in value through MC simulation. For vanilla derivatives, like caps and swaptions, we can use the Black formula for valuation. In our simulations, only vanilla derivatives are traded so, for the sake of simulation efficiency, this is the approach used to value the cap in the interest rate derivatives portfolio.

The valuation of the zeros at the forecasted 5 days ahead point within one simulation is done through equation 5.9 so that letting $\tilde{t} = \frac{5}{252}$ be the 5 days ahead point we have the value of the 1 year zero

$$Z_4(\tilde{t}) = Z_1(\tilde{t}) \prod_{j=1}^{3} \frac{1}{1 + \delta_j L_j(\tilde{t})} \tag{7.12}$$

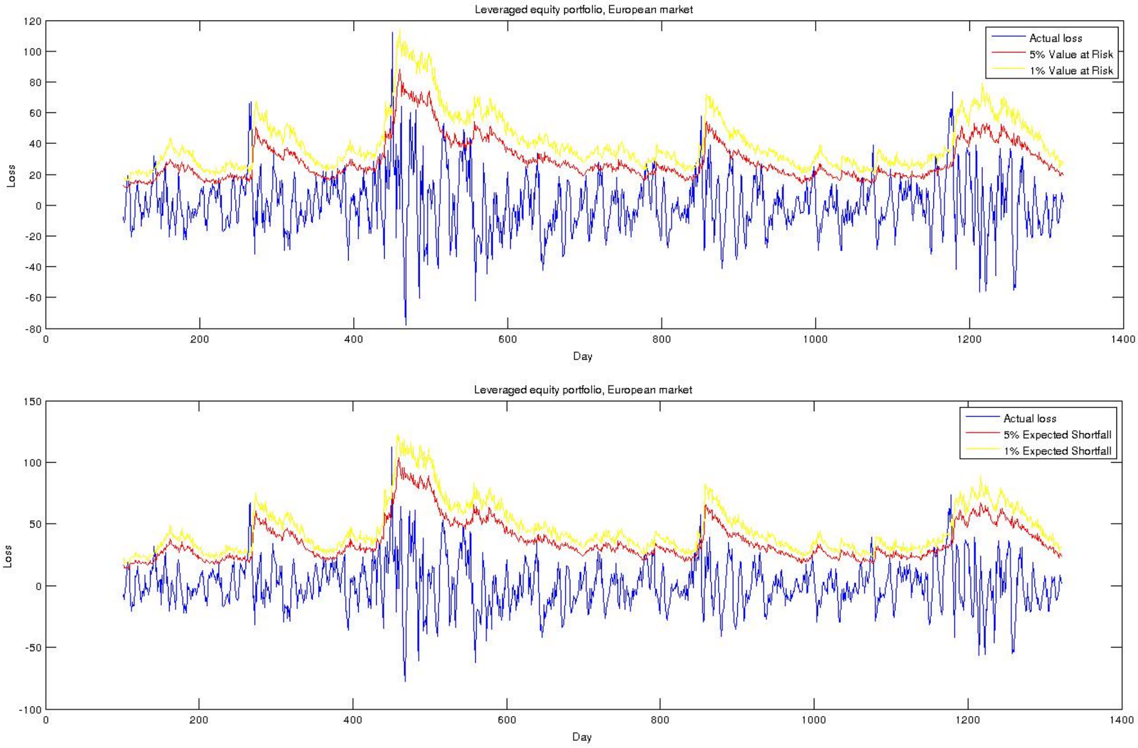
and the 3 year zero

$$Z_{12}(\tilde{t}) = Z_1(\tilde{t}) \prod_{j=1}^{11} \frac{1}{1 + \delta_j L_j(\tilde{t})} \tag{7.13}$$

We have simulated values for the L_j. The missing part of these equations is the $Z_1(\tilde{t})$, i.e. the discount factor from time $T_1 = 0.25$ back to time $\tilde{t} = \frac{5}{252}$. Our model does not provide these discount factors, so we use the assumption that the 3-month discount factor, from time $T = 1$ back to time $t = 0$ does not change and do the linear interpolation

$$Z_1(\tilde{t}) = Z_1(0) + \frac{Z_1(T_1) - Z_1(0)}{T_1 - 0}\tilde{t} = Z_1(0) + 0.25(1 - Z_1(0))\tilde{t} \tag{7.14}$$

Figure 8.1: *Actual 5 day loss along with 5% and 1% 5 day VaR (top) and ES (bottom) numbers for each day for the Leveraged Equity portfolio in the European market.*



8 Results

The performance of the risk measure can be measured by the percentage of days our portfolios' actual loss is larger than the VaR and ES numbers. The actual loss should exceed the 5% VaR on around 5% of the days and exceed the 1% VaR on around 1% of the days. We do not know for exactly how many days the actual loss should exceed the ES numbers, since we do not know a priori the loss distribution. It is reasonable to expect, however, that the tail of the loss distribution is left skewed. In other words, among the 1% worst losses the probability of smaller losses is greater than the probability of larger losses. So we should have that the actual loss exceed the 5% ES on less than 2.5% of the days and exceed the 1% ES on less than 0.5% of the days. In figure 8.1, an example plot of the actual loss and VaR and ES numbers is shown for the leveraged equity portfolio in the European market.

8.1 European market

The performance results for the European market are summarized in table 8.1, where the percentage of days in which the actual loss exceeds the VaR and ES numbers are

Table 8.1: *The percentage of days in the data set where the actual loss over 5 days exceeds the 5 day VaR and ES numbers, European market*

Portfolio	Conservative	Moderate	Equity	Lev. equity	Lev. IR
Exceeding 1% VaR	3.52%	2.86%	2.37%	2.37%	2.54%
Exceeding 5% VaR	8.67%	7.61%	7.28%	7.20%	7.04%
Exceeding 1% ES	2.37%	1.72%	1.39%	1.39%	1.80%
Exceeding 5% ES	4.83%	3.93%	3.76%	3.19%	3.60%

shown. We see that the risk measures perform badly overall. Furter, we see that the performance seem to be better for the riskier portfolios and better for the 5% VaR and ES than the 1% numbers, but the differences are rather small, expect for the conservative portfolio, for which performance is significantly worse.

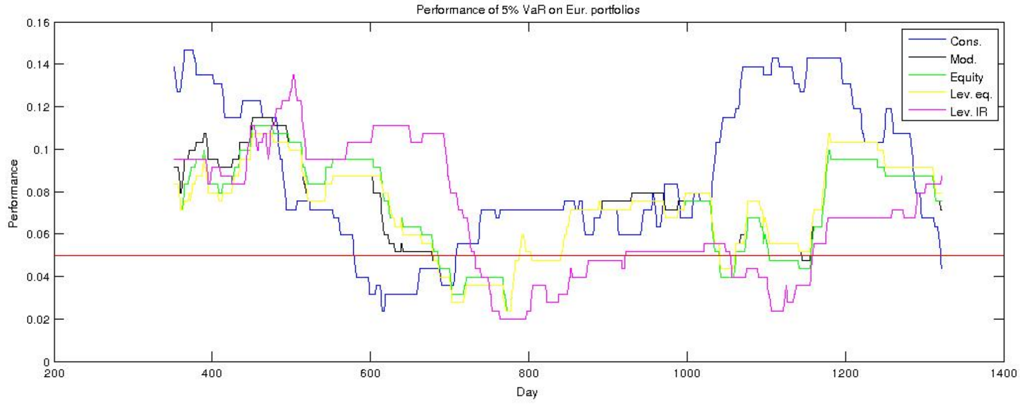
In figure 8.2 we have plotted, for each portfolio, the performance of the 5% VaR and ES for each 252 day (1 trading year) period, measured in the percentage of days where the actual loss exceeds the VaR and ES numbers. The performance is seen to be varying greatly in time. Most periods have poor performance, but for some periods we do have acceptable risk measure performance. Comparing with Mardia's test for the European market in figure 6.7, we see that they loosely agree. We have non-normal peaks in the periods ending around 10.10.2007 (day 200) and from 11/12/2008 to 07.05.2009 (day 500 to day 600) corresponding to the period of poor performance in the beginning of the time period in figure 8.2. We also have non-normal peaks towards the end of the period, in the periods ending around 24.11.2010 to 02.09.2011 (day 1000 to day 1200) corresponding to the period of poor performance in the end of the time period in figure 8.2. This theme of poor performance in the beginning and end of the time period with a better intermediate period, also agrees with the LM and CM test statistics for normality of stock returns as seen in figures 6.5 and 6.6.

8.2 Norwegian market

The performance results for the Norwegian market are summarized in table 8.2, where the percentage of days in which the actual loss exceeds the VaR and ES numbers are shown. As in the European market, the risk measure performance is bad overall and we seem to have better performance for riskier portfolios. We also see that we have similar performance for the Moderate, Equity and Leveraged Equity portfolios, and the performance on the conservative portfolio is significantly worse. In contrast to the European market, however, we have significantly better performance for the IR Derivatives portfolio.

In figure 8.3 we have plotted, for each portfolio, the performance of the 5% VaR and ES for each 252 day (1 trading year) period, measured in the percentage of days where the actual loss exceeds the VaR and ES numbers. The performance varies in time for the Norwegian market as well, with overall performance being bad, but with periods of good

Figure 8.2: *Performance of 5% VaR and ES measured as percentage of days in which actual loss exceeds VaR (top) and ES (bottom), for each 252 day period in the European market. Also plotted is a 5% (for VaR) and 2.5% (for ES) reference line. The numbers on the "Day" axis reprepsents the last day in the 252 day period.*



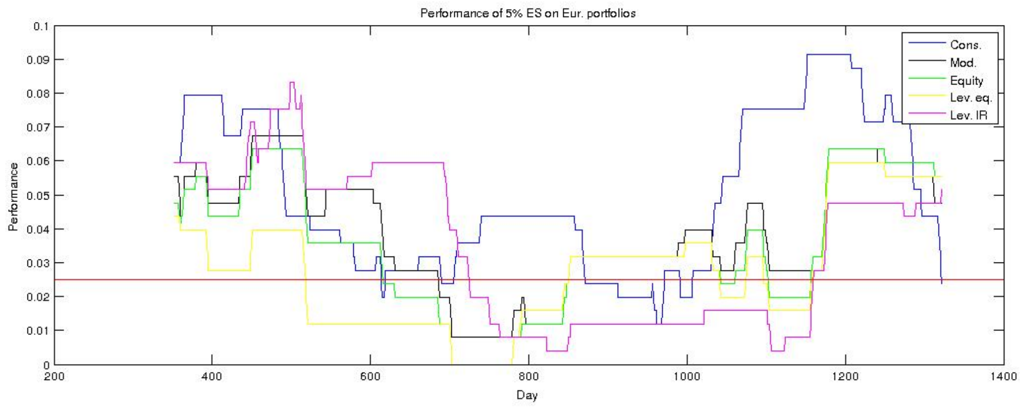


Table 8.2: *The percentage of days in the data set where the actual loss over 5 days exceeds the 5 day VaR and ES numbers, Norwegian market.*

Portfolio	Conservative	Moderate	Equity	Lev. equity	Lev. IR
Exceeding 1% VaR	4.42%	3.60%	3.52%	3.44%	1.72%
Exceeding 5% VaR	8.92%	7.61%	7.53%	7.53%	5.65%
Exceeding 1% ES	3.11%	2.37%	2.29%	2.21%	0.98%
Exceeding 5% ES	6.14%	4.58%	4.42%	4.34%	3.03%

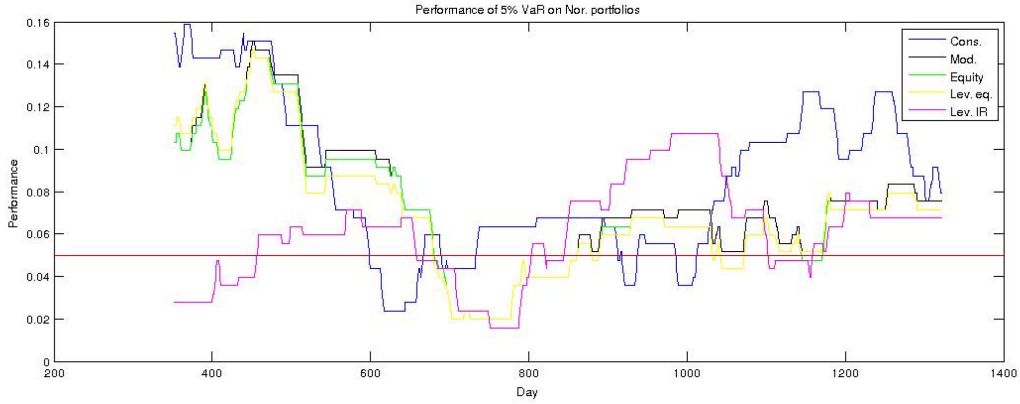
performance. Comparing with Mardia's test for the Norwegian market in figure 6.7, we see that, like for the European market, the figures loosely agree, in the same way as discussed for the European market. There are two periods of both non-normal returns and poor risk measure performance at the beginning and the end of the time period, with a more normal period in between, in which we also have better performance. For the Norwegian market however, we cannot see a clear correspondence between periods of poor performance and periods of non-normality of stock returns.

8.3 Sensitivity to λ

The hardest part of financial modelling is arguably the model calibration, i.e. the estimation of volatilities. It is therefore natural to check to what degree changes in volatility influences the result. In our approach, in which risk measure performance is measured as the percentage of days the actual loss exceeds the VaR numbers, changing the volatilities directly holds no merit, since increased volatilities will increase the VaR and ES numbers and obviously improve performance. However, the $\lambda = 0.94$ value was arbitrarily chosen, and this parameter will conceivably influence the results. We therefore did a sensitivity analysis on the results with respect to this parameter, by running the simulations for different values of λ. The resulting performances for the 5% VaR in both markets are shown in figure 8.4.

Wee see that the sensitivity to λ is different in the European and Norwegian markets. For the Conservative portfolio, the performance is monotonically increasing (i.e. a downward slope in figure 8.4) with increasing λ in both markets, but the performance increase is larger in the Norwegian market. For the Moderate, Equity, and Leveraged Equity portfolios however we see that the performance is not sensitive to λ in the Norwegian market but increases greatly with λ in the European market. As for the Leveraged IR Derivatives portfolio, the λ sensitivity have similar characteristics in both markets, with a performance peak for $\lambda \in (0.92, 0.96)$. This shows that our original choice of $\lambda = 0.94$ was indeed sensible. In the Norwegian market, however, the dip in performance for λ close to 1 is very slight compared to the European market.

Figure 8.3: *Performance of 5% VaR and ES measured as percentage of days in which actual loss exceeds VaR (top) and ES (bottom), for each 252 day period in the Norwegian market. Also plotted is a 5% (for VaR) and 2.5% (for ES) reference line. The numbers on the "Day" axis reprepsents the last day in the 252 day period.*



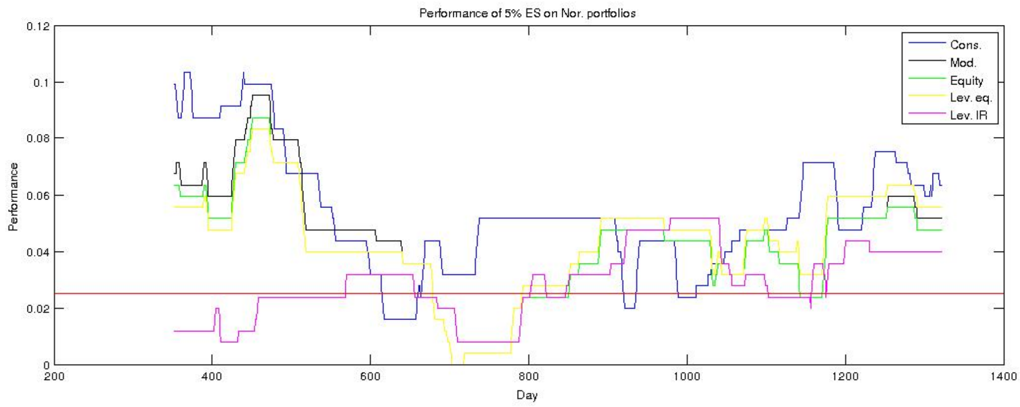
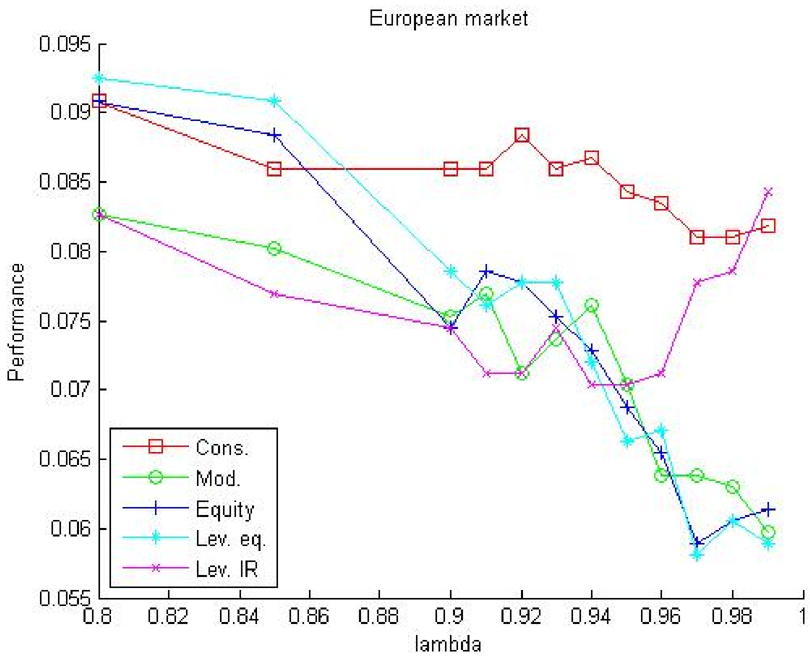
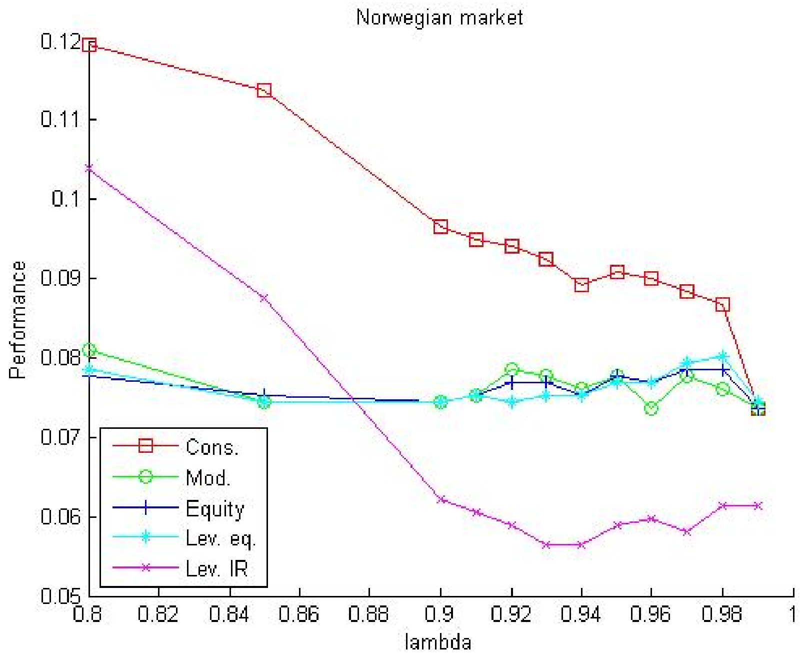


Figure 8.4: *The performance of 5% VaR in European (top) and Norwegian (bottom) market for different values of the λ parameter in the EWMA estimation of the covariance matrix.*





8.4 Discussion

There are a quite a few points to be made about the preceding results. First of all, the main purpose is to get a good idea of the expected performance of the risk measures. The performance should therefore be measured over as long time periods as possible, by the law of large numbers. If we use too short time periods, there will always be some periods with good performance and some with bad, depending on these periods including days of abnormal returns or not. In this respect, our performance results over the whole period are the most trustworthy.

For these performance numbers, we have seen that we have similar performance for the Moderate, Equity, and Leveraged Equity portfolios in both markets. This is not unexpected, since these three portfolios are qualitatively equal, i.e. they hold the same types of assets, and differ only in the amount held, positive or negative, of the 3 year zero. Further, in both markets we observe significantly worse performance on the Conservative portfolio. This portfolio holds all wealth in 1 year zeros, an asset which none of the other portfolios hold. Thus, it seems that the risk measures perform worse on the 1 year zero than the other assets.

We further saw that the performances on the IR Derivatives portfolio was similar to the performance on the Moderate, Equity, and Leveraged Equity portfolios in the European market. In the Norwegian market, however, the performance was significantly better for this portfolio. Such a divergence might be explained by that the Norwegian interest rate market is under normal market conditions so volatile that the estimated volatilities are large enough so that our risk measures perform better when we have abnormal returns. In tables 6 and 6 we saw that the volatility of the 3 month rate L_1 over the whole period is much larger for the Norwegian market than the European. And the cap held in the IR Derivatives portfolio depend on the rates L_1, L_2, L_3, and L_4, the first three of which is perfectly correlated with L_1.

Keeping in mind that larger λ means a greater weight on historical returns farther into the past, this idea may be strengthened by the fact that, in contrast to the European market, the VaR performs only slightly worse on the IR Derivatives portfolio for λ approaching 1. In general therefore, the performance is better if the volatility is drawn from the general market situation rather than what happened on the last few days.

Even though the whole period performance results are the most trustworthy, it is instructive to observe the large differences in performance for different periods. The time period in question, from early 2007 to early 2012 was a period of great distress in the financial markets. These were years of "abnormal" market conditions in which abnormal returns were experienced more often than in other historical time periods. This shows clearly the relevance of the discussions in section 4 about the problem with Value at Risk and related risk measures. It is a risk measure that performs well as long as markets behave like we want them to. However, in times of volatile markets, when it is most important to control our financial risk, they are a less reliable tool for exact risk measurement. Of course, as discussed above, the VaR is not only useful in this way.

Even though we cannot trust VaR as an exact risk measure, it is still a good indicator of market conditions. If VaR substantially increases from one day to the next, we know that our risk has increased. Since VaR is easy to understand and easily aggregated over a financial firms different operations it is therefore a useful risk control tool for managers and analysts at every level in a financial institution.

Finally, it is important to remember that the main results of this thesis is the end point of multiple steps of modelling and implementation choices. From the initial choices of how to handle the historical data, through the model calibration choices, to the specific discretization and computer implementation of the models. In each of these steps, we could have done things differently. We therefore have many sources of error and it is hard to directly see which choices influence our results the most.

Arguably one of the most important choices is the model calibration. We used historical volatilities and correlations to calibrate our models. We see in figure 8.1 that our risk measures has a delay in reacting to increased volatility. This will obviously influence the performance. We have seen that a different λ in the EWMA estimates of the covariance matrix could have increased performance for some portfolios, but overall was $\lambda = 0.94$ a good choice. A calibration method based on implied volatilities or optimization would conceivably have improved performance through making the risk measures react quicker to increased market volatilities and give more correct prices of derivatives.

For another example, the poor performance for the Conservative portfolio, i.e. the one year zero, may be due to several modelling choices. It may be due to the interpolation done when calculating the shortest discount factor when revaluing zeros. It may also be due to linear interpolation choice in the bootstrapping of the historical data in section 7.1. Or it may of course simply be a consequence of real market conditions, rather than any specific modelling choice.

These are just a couple of examples of how modelling choices may influence results. Exchanging the LMM with another interest rate model, or calculating VaR and ES in different ways are other obvious examples. To get a better picture of the importance of each modelling choice, more sensitivity analysis, like we did for the λ parameter, must be done.

9 Conclusion

The overall performances of the risk measures were found to be not satisfactory for all tested portfolios. A tendency towards the measures performing better on the riskier leveraged portfolios was seen, but this is more likely attributable to the differences of performance of the individual assets rather than to the leverage itself.

Further, the risk measures performance was found to be varying greatly in time. We found that the risk measures performed bad in loosely the same historical time periods as we found the log-returns of the stock indices and forward rates to diverge the most from the normal distribution. Specifically, in late 2008 and the following time, a period of great market turbulence, we observed both great divergence from normality of the log-returns and very poor risk measure performance.

The correlation between market turbulence, non-normality, and poor risk measure performance is hardly surprising, but strengthens the point that Value at Risk and related measures are less useful in non-normal and volatile periods, the very periods in which good risk measures are most crucial.

We found that many of the portfolios were sensitive to the parameter choice λ in the volatility estimation, especially in the European market, which shows the great influence modelling choices may have on results. The trade-off between simplicity and accuracy will, like in all mathematical modelling, always be an issue. This speaks against using Value at Risk and Expected Shortfall as precise absolute measures of risk. On the other hand, they can still be useful risk management tools, since they certainly provide information on the changes in day to day market risk and communicate this in a way that is useful and easy to grasp.

10 Further work

Many of our modelling choices can be improved upon. We have for instance used the rather simple historical volatility approach to model calibration. Improving this, by using an implied volatility approach or doing parameter optimization will probably improve the predictive ability of our models. Using a more sophisticated bootstrapping of our original data set and estimating more carefully the real drift of asset, are other natural points of improvement.

Further, we did not allow for cross-currency trading. We simulated the Norwegian and European markets separately. Since cross-currency trading is done by many, if not most, financial institutions, it would be useful to allow this in our modelling framework. This would of course complicate the models since we would need to model the exchange rate between currencies.

For the purposes of computational speed, we did only 1000 simulations. This results in relatively high standard errors in our Monte Carlo simulation. A good point of further work is to do the implementation in a computationally time-saving way, like using a variance reduction technique in the MC simulation. This would free up resources so that we could reduce the standard error and/or include exotic derivatives in our portfolios.

The way we implemented the portfolios made leverage, i.e. borrowing money against the held assets, no more than an inverted scaling of holding zero coupon bonds. Investigating more systematically the effects of portfolio leverage on risk measure performance could be useful. This would require a well defined idea of what leverage should mean within the context of these models. We could for instance incorporate the possibility of "portfolio bankruptcy" or margin calls if the portfolio has liquidity (cannot make payments) or solidity (has more debt than the total value of assets held) problems.

We have seen that risk measure performance is better for some types of assets than others in our implementation.It would be interesting to investigate whether we could find consistent ways of correcting our risk measures accordingly, to create a more trustworthy risk measure. For instance, if we find that our risk measures over time overestimate the market risk associated with holding the OSEBX index, we could correspondingly scale the risk measure for this asset. Doing the same for other assets could then give us a better risk measure, even when these assets are combined. Issues with this initial idea immediately come to mind. The correlation of different asset prices might for instance be a problem, when combining the corrected risk measures. It might still however, be an interesting matter of further study.

References

[1] Artzner, Philippe; Delbaen, Freddy; Eber, Jean-Marc; Heath, David: *Coherent Measures of Risk.* 1998

[2] Csorgo, Sandor; Faraway, Julian: *The exact and asymptotic distribution of Cramer-von Mises statistics.* Journal of the Royal Statistical Society: Series B, 58(1), 221-234, 1996

[3] Glasserman, Paul: *Monte Carlo methods in financial engineering.* Xth edition, Springer Science 2004

[4] Kohn, Robert V: *Continuous time finance notes.* Retrieved from http://math.nyu.edu/faculty/kohn 2004

[5] Pietersz, R: *The LIBOR market model.* Leiden University 2003

[6] Rencher, Alvin C.: *Methods of Multivariate Analysis.* 2nd edition, Wiley-Interscience 2002

[7] J.P.Morgan and Reuters: *Riskmetrics - technical document.* Technical report, J.P.Morgan and Reuters 1996

[8] Shreve, Steven E: *Stochastic Calculus for Finance II, Continuous-Time Models.* Xth edition, Springer Science 2004

[9] Stephens, M.A: *EDF Statistics for Goodness of Fit and Some Comparisons.* Journal of the American Statistical Association, 69, 730-737, 1974

[10] Weibull Reliability Engineering Resources: *Critical Values for Cramer-von Mises Test.* Retrieved from http://www.weibull.com/RelGrowthWeb/Appendix_B_Critical_Values_for_Cramer-von_Mises_Test.htm

[11] Wilmott, Paul: *Paul Wilmott on quantitative finance.* 2nd edition, John Wiley & Sons, Ltd 2006

[12] Wilmott, Paul; Howison, Sam; Dewynne, Jeff: *The mathematics of financial derivatives.* Press Syndicate of the University of Cambridge 1995

[13] Wuertz, Diethelm; Katzgraber, Helmut: *Precise finite-sample quantiles of the Jarque-Bera adjusted Lagrange multiplier test.* Swiss Federal Institute of Technology 2009

A Definitions and theorems

This section provides needed definitions, mostly from probability theory. Most of the definitions in this section are taken from [8].

Definition 7 *Let X and Y be two random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We say that $X = Y$* ***almost surely*** *if*

$$\mathbb{P}(X = Y) = \mathbb{P}(\omega : X(\omega) = Y(\omega)) = 1 \tag{A.1}$$

Note that $X = Y$ almost surely does not entail that $X(\omega) = Y(\omega) \forall \omega \in \Omega$, unless Ω is finite and countable.

A.1 Probability space

Definition 8 *Let Ω be a nonempty set, and let $\mathcal{F}$ be a collection of subsets of Ω. $\mathcal{F}$ is called a* ***σ-algebra*** *if:*

- *the empty set $\emptyset$ belongs to $\mathcal{F}$,*
- *whenever a set A belongs to $\mathcal{F}$, its complement A^c also belongs to $\mathcal{F}$, and*
- *whenever a sequence of sets $A_1, A_2, ...$ belongs to $\mathcal{F}$, their union $\cup_{n=1}^{\infty} A_n$ also belongs to $\mathcal{F}$.*

Definition 9 *Let Ω be a nonempty set, and let $\mathcal{F}$ be a σ-algebra of subsets of Ω. A* ***probability measure*** *is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ that, to every set $A \in \mathcal{F}$ assigns a number in $[0, 1]$, called the probability of A and written $\mathbb{P}(A)$. The function $\mathbb{P}$ must satisfy:*

- *$\mathbb{P}(\Omega) = 1$ and $\mathbb{P}(\emptyset) = 0$, and*
- *(countable additivity) whenever $A_1, A_2, ...$ is a sequence of disjoint sets in $\mathcal{F}$, then*

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) \tag{A.2}$$

Definition 10 *Let $(\Omega, \mathcal{F})$ be a measurable space, and let $\mathbb{P}$ and $\mathbb{Q}$ be probability measures on the space. We say that $\mathbb{P}$ and $\mathbb{Q}$ are* ***equivalent*** *if $\mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0$ for all $A \in \mathcal{F}$. That is, $\mathbb{P}$ and $\mathbb{Q}$ agree on which sets in $\mathcal{F}$ have zero probability.*

Definition 11 *A* ***probability space*** *is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is an arbitrary nonempty set, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and $\mathbb{P}$ is a probability measure on $\mathcal{F}$.*

A.2 Borel sets and random variables

Definition 12 *Let $\mathcal{G}$ be the smallest σ-algebra that contains all closed intervals of $\mathbb{R}$ (and all other subsets $A \subset \mathbb{R}$ necessary for $\mathcal{G}$ to be a σ-algebra). $\mathcal{G}$ is then called the* ***Borel σ-algebra of subsets of*** $\mathbb{R}$*, $\mathcal{B}(\mathbb{R})$. The sets in $\mathcal{B}(\mathbb{R})$ are called* ***Borel sets****.*

Definition 13 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A* ***random variable*** *is a function $X : \Omega \to \mathbb{R}$ such that for every Borel set $B \in \mathcal{B}(\mathbb{R})$, the subset of Ω given by*

$$\{\omega \in \Omega : X(\omega) \in B\} \tag{A.3}$$

is in the σ-algebra $\mathcal{F}$.

A.3 Filtration and information

Definition 14 *Let $(\Omega, \mathcal{F})$ be a measurable space. A* ***filtration of*** *$\mathcal{F}$ is a sequence of σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$ such that:*

- $\mathcal{F}_t \subseteq \mathcal{F}$, $\forall t$,
- $t_1 \leq t_2 \Rightarrow \mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2}$

Definition 15 *A* ***filtered probability space****, $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ along with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ of $\mathcal{F}$.*

Definition 16 *Let $X : \Omega \to \mathbb{R}$ be a random variable. The* ***σ-algebra generated by*** *X, $\sigma(X)$, is the collection of all subsets of Ω of the form $\{\omega \in \Omega : X(\omega) \in B\}$ where $B \in \mathcal{B}(\mathbb{R})$.*

Definition 17 *Let $X : \Omega \to \mathbb{R}$ be a random variable. Let $\mathcal{G}$ be a σ-algebra of subsets of Ω. If every set in $\sigma(X)$ is also in $\mathcal{G}$, we say that X is* ***$\mathcal{G}$-measurable****.*

A.4 Stochastic processes

Definition 18 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A* ***stochastic process*** *$X(t)$ is a collection of random variables $X_t : \Omega \to \mathbb{R}$, indexed by $t \in [0, T]$ for some final time T.*

Definition 19 *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space. Let $X(t)$ be a stochastic process on this space. $X(t)$ is an **adapted** stochastic process if , for all t, the random variable $X(t)$ is $\mathcal{F}(t)$-measurable.*

Definition 20 *Let $f(t)$ be a function defined for $0 \leq t \leq T$. The **quadratic variation** of f up to time T is*

$$[f, f](T) = \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2, \tag{A.4}$$

where $\Pi = \{t_0, t_1, ..., t_n\}$ and $0 = t_0 < t_1 < ... < t_n = T$.

We denote by $d[f, f]$ or df^2 the infinitesimal change in quadratic variation.

A.5 Expectations and Martingales

Definition 21 *Let X be a random variable defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The **expectation of** X is defined to be*

$$\mathbb{E}X = \int_\Omega X(\omega) d\mathbb{P}(\omega) \tag{A.5}$$

Definition 22 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, and let X be a random variable that is non-negative or integrable. The **conditional expectation of** X **given** $\mathcal{G}$ is the random variable $\mathbb{E}[X|\mathcal{G}]$ that satisfies*

- *$\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable.*
- *$\int_A \mathbb{E}[X|\mathcal{G}](\omega) d\mathbb{P}(\omega) = \int_A X(\omega) d\mathbb{P}(\omega)$ for all $A \in \mathcal{G}$.*

Note: if $\mathcal{G}$ is the σ-algebra generated by some random variable Y, we often write $\mathbb{E}[X|Y]$.

Definition 23 *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space. If the adapted stochastic process $M(t)$ satisfies*

$$\mathbb{E}[M(t)|\mathcal{F}(s)] = M(s) \text{ for all } 0 \leq s \leq t \leq T, \tag{A.6}$$

*$M(t)$ is a **martingale**.*

Note: if $\mathcal{F}(t)$ is the filtration generated by a stochastic process $X(t)$, we often write

$$\mathbb{E}[M(t)|X(s)] \tag{A.7}$$

instead of $\mathbb{E}[M(t)|\mathcal{F}(s)]$.

A.6 Brownian motion

A Brownian motion is a stochastic process , continuous in time, where an increment between two times are normal distributed with mean 0 and variance equal to the time difference.

Definition 24 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For each $\omega \in \Omega$, suppose there is a continuous function $W(t)$ of $t \geq 0$ that satisfies $W(0) = 0$ and that depends on ω. Then $W(t)$, $t \geq 0$ is a* ***Brownian motion*** *if for all $0 = t_0 < t_1 < \ldots < t_m$ the increments*

$$W(t_1) = W(t_1) - W(t_0), W(t_2) - W(t_1), \ldots, W(t_m) - W(t_{m-1})$$

are independent and each of these increments are normally distributed with

$$\mathbb{E}[W(t_{i+1}) - W(t_i)] = 0$$
$$Var[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i$$

A.7 Ito process

Definition 25 *Let $W(t)$, $t \geq 0$ be a Brownian motion, and let $\mathcal{F}(t)$, $t \geq 0$ be be the filtration generated by $W(t)$. An* ***Ito process*** *is a stochastic process of the form*

$$X(t) = X(0) + \int_0^t \mu du + \int_0^t \sigma dW(t),$$

where $X(0)$ is nonrandom and μ and σ are adapted stochastic processes. We write

$$dX_t = \mu dt + \sigma dW_t$$

A.8 Ito's lemma

If we have an Ito drift-diffusion process

$$dX_t = \mu dt + \sigma dW_t,$$

then a twice differentiable function $f(t, x)$ of real variables has the property that

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma \frac{\partial f}{\partial x} dW_t \tag{A.8}$$

This is the chain rule for differentiation of a function of an Ito process. Ito's lemma is a frequently needed tool when working with SDEs.

A.9 Feynman-Kac Theorem

The Feynman-Kac theorem provides a connection between partial differential equations and expectations.

Theorem 3 *(Feynman-Kac) Consider the stochastic differential equation*

$$dX(u) = \beta du + \gamma dW(u). \tag{A.9}$$

where $W(u)$ is a Brownian motion. Let $h(y)$ be a Borel-measurable function and let r be a constant. Fix $T > 0$, and let $t \in [0, T]$ be given. Define the function

$$f(t, x) = \mathbb{E}[e^{-r(T-t)}h(X(T))|x]. \tag{A.10}$$

where $x = X(t)$. Then $f(t, x)$ satisfies the PDE

$$f_t(t, x) + \beta(t, x)f_x(t, x) + \frac{1}{2}\gamma^2(t, x)f_{xx}(t, x) = rf(t, x) \tag{A.11}$$

and the terminal condition

$$f(T, x) = h(x) \text{ for all } x. \tag{A.12}$$

See [8] for a proof.

A.10 Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be independent random variables with the same distribution, with mean μ and variance σ^2. Then the limit

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \ldots + X_n}{n}$$

is normal distributed with mean μ and variance $\frac{\sigma^2}{n}$