Pricing options under stochastic volatility

by

Rudolf Gerrit Venter

Submitted in partial fulfillment of the requirements for the degree

Magister Scientiae in Mathematics of Finance

in the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

January 2003

Summary

In this dissertation some of the real world deviations from the assumptions made in the Black-Scholes option pricing framework is investigated. Special attention is paid to volatility, the standard deviation of stock price returns. Unlike the assumption of constant volatility of increments in Brownian motion, volatility in the market is stochastic. Market models allowing for stochastic volatility are no longer complete as in the Black-Scholes framework. Options in incomplete markets are harder to price since investors demand higher returns for taking additional risk.

Duan (1995) proposed an option pricing measure for incomplete markets, due to stochastic volatility, called the Local Risk-Neutral Valuation Relationship (LRNVR). Under the LRNVR, the local risk neutral measure (Q) is equivalent to the real world measure (P), the conditional expected return under the Q measure equals the risk-free rate and the conditional one period ahead variances under both measures are equal, P almost surely. The LRNVR holds for consumers with familiar utility functions.

Stock returns are assumed to follow a Generalized Autoregressive Conditional Heteroscedastic (GARCH) process. This process is a discrete time statistical time series that is calibrated over stock returns. In this dissertation the LRNVR and related option pricing methodology is comprehensively investigated.

Warrants traded on the JSE Securities Exchange violates the Black-Scholes assumptions in two additional ways, short selling is restricted and the market is somewhat illiquid. One of the results of these violations is that the standard deviation and the implied volatility, volatility implied by the market price of the option, are out of sync. The implied volatility tends to be higher than the volatility of stock market returns.

In this dissertation the GARCH option pricing process is applied to the implied volatility of the warrant instead of the stock price process, as done by Duan. This method compares well with the use of implied volatility to price warrants.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Acknowledgements

I would like to thank my supervisors, Professor Frik van Niekerk of the University of Pretoria and Johan Verwey of Gensec Bank for their support and guidance. My thanks to Professor Dawie de Jongh for advice and references to literature is also due. I am also grateful to my family for their motivation and support.

Contents

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Glossary of notation

Glossary of frequently used notation:

$(\Omega, \mathcal{F}, P)$, 7
$a.e.$, 9
$A(L)$, 38
$B(L)$, 38
cdf, 14 (Cumulative distribution function)
$cor[X, Y]$, 13
$cov[X, Y]$, 12
$E(e^{tX})$, 17
$E[X]$, 9
$E[X \mid \Phi]$, 10
$F(x)$, 14 (Cumulative distribution function)
$f(x)$, 14 (Probability density function)
$\chi^2(v)$, 20
$L^1(\Omega, \mathcal{F}, P)$, 9
$M_X(t)$, 17
$N(\mu, \sigma^2)$, 16
pdf, 14 (Probability density function)
$Std[X]$, 10
σ_t^2, 38 (GARCH process)
$u(x)$, 66
$Var[X]$, 10
$Var[X \mid H]$, 10

Chapter 1

Introduction

Three categories of financial models prevail in the market[1]. They are the following:

1. **Structural models.** Simplifying assumptions about the underlying market processes and market equilibrium are made to infer equilibrium prices and thus the relationships between underlying instruments and their contingent claims (i.e. options). The Black-Scholes[2] formula is the most famous structural model. The Black-Scholes formula is the result of a method called *risk-neutral* (or arbitrage) pricing. A result of the risk-neutral pricing is that we can infer a unique, correct price of a contingent claim given its underlying stock price. Any other option price would lead to an arbitrage opportunity.

2. **Statistical models.** These models rely on empirical data and their co-dependencies. Fewer assumptions, if any, are made concerning the structure of the market. Examples of statistical models in financial mathematics are the capital asset pricing model and time series processes. Financial time series are used to describe data, to obtain insight into their dynamic patterns and to forecast out-of-sample returns. The *Generalized Autoregressive Conditional Heteroscedastic* (GARCH) process is a famous time series used to model the conditional variance of a process.

3. **Combination of structural and statistical models.** This category of models combines the above categories of models. The GARCH option pricing model under the *local risk-neutral valuation relationship* (LRNVR), discussed in this dissertation, is the combination of GARCH literature and risk-neutral valuation.

[1] See 'Risk Management' by Crouhy, Galai and Mark [9].

[2] The Black-Scholes model was developed by Black and Scholes (1973) and Merton (1973).

Under risk-neutral pricing, the price of a contingent claim is independent of the risk preference and utility functions of buyers and sellers, hence there exists a unique and correct option price. The cost of this model is the simplifying assumptions. Some of the crude assumptions made in the Black-Scholes model are:

1. Stock prices are *lognormally* distributed, thus the continuously compounded stock returns are normally distributed.
2. The mean and volatility under this distribution are constant.
3. The risk-free interest rate is constant or a known function of time.
4. Delta hedging is done continuously (short selling is allowed and securities are perfectly devisable).
5. No transaction costs on the underlying.
6. No arbitrage opportunities.

Empirical evidence shows that none of these assumptions are valid. In this dissertation the assumption of constant volatility is abandoned, for (conditional) *stochastic volatility.*

Volatility has many definitions. It is generally seen as the standard deviation of a random process (i.e. the stock returns process). In the Black-Scholes framework, implied volatility can be inferred from the market price of the option and the underlying. Conditional volatility can be seen as a measure of risk. This is because levels of trade tend to increase in uncertainty in the stock, sector or market in general and hence the standard deviation or price fluctuations increase.[3].

In this dissertation, volatility is seen as the standard deviation of a stochastic process. Implied volatility comes into play in later chapters where the GARCH option pricing model is applied to JSE Exchange traded warrants.

1.1 The Problem of Stochastic Volatility

The Black-Scholes model is a complete market model. A market model is *complete* if and only if all contingent claims are *replicable.* Equivalently, under no arbitrage conditions, a market model is complete if and only if there exists a unique risk-free probability measure.

If stochastic volatility is introduced into a market model, it is no longer complete[4]. This is because there are too much variability in the stock price

[3]For a thorough discussion on market volatility, see Poon & Granger [29].

[4]See Fouque et al [17].

which cannot be hedged away completely, since there are no instruments in the market which is perfectly correlated with the individual stock's volatility. Equivalently there doesn't exist a unique risk-neutral probability measure.

A consequence of stochastic volatility is that the price of the contingent claim depends on the risk preference and utility of investors. This complicates computation of the price of the contingent claim.

1.2 A Proposed Solution

The aim of this dissertation is to discuss a solution too the problem of option pricing in incomplete markets, due to stochastic volatility. The LRNVR was introduced by Jin-Chuan Duan [10] in 1995. Duan proved that the measure

$$dQ = e^{-(r-\rho)T}\frac{U'(C_t)}{U'(C_{t-1})}dP$$

satisfies the LRNVR. In this measure, r is the risk-free interest rate, ρ is an impatience factor and U' is the first derivative of the utility function of consumption C_t at time t. The measure Q is called the local risk-neutral measure.

The volatility process in this dissertation is the GARCH process introduced by Engle (1982) and Bollerslev (1986) [6]. The GARCH process is a discrete time process of the changing variance of the returns of an underlying instrument. This process captures phenomena of returns series coined "stylized facts". These phenomena are heavy-tails[5] of distributions, volatility clustering[6] and mean reversion[7]. GARCH processes have been extended to capture another stylized fact called the leverage effect[8]. Such GARCH processes are called asymmetric GARCH processes.

The GARCH parameters are derived from actual market prices. The stock price, at expiry of a European option, is forecasted with the GARCH process under the local risk-neutral measure. This forecast is done with Monte Carlo simulations.

In this dissertation the GARCH option pricing method is applied to South African put warrants.

[5] Excess kurtosis above that of the normal distribution.

[6] Volatility levels tend to cluster together at the same levels for a certain duration, after which it clusters together at another level.

[7] Volatility levels tens to revert back to a certain long-term level after a shock. The reversion to this level is not neccesarily immediate.

[8] The market tends to react more drastcally to bad news than good news.

1.3 Description of South African Derivative Instruments and Experiment

There are two markets where financial derivatives are traded in South Africa. The one market is the warrants market of the JSE Securities Exchange (JSE) and the other is the South African Futures Exchange (SAFEX). The SAFEX exchange was bought by the JSE on the 1^{st} of July 2001.

Equity options on SAFEX are traded on a limited number of stocks and on some index futures. The SAFEX market tends to be illiquid. In illiquid markets the spread between bid and offer prices tends to be wider than that of a more liquid markets.

On the JSE, warrants[9] are traded. A warrant is an option issued, like a stock, by financial institutions on equities, certain interest rate instruments and some indices. This means that a market player must own a warrant to sell it, thus no short selling is allowed. The warrants market is more liquid than the SAFEX options market, but because no short selling is allowed, there are no way to gain from overpriced warrants. In this market, only market equilibrium (supply and demand) controls price levels. The result is that the implied volatility levels of warrants tend to be higher than the volatility of stock prices. See figure 1.1.

In this dissertation the GARCH option pricing method is applied to equity European put warrants on the JSE. Approximately 30% of traded warrants are European put warrants. The warrants market was selected because it's more liquid than the SAFEX option market. In more liquid markets, option prices reacts more rapidly to changes in the price of the underlying, thus the testing of the GARCH option pricing method is easier to do.

In Duan's 1995 paper the GARCH process is calibrated to the returns series of the underlying equity or index with the maximum likelihood method. Since the implied volatility of warrants are higher than the historical standard deviation of the underlying equity, the GARCH process in this dissertation is fitted to the implied volatility of the warrant.

1.4 Outline of the Dissertation

In the following chapter, essential background to probability theory is discussed. This discussion includes some measure theoretical background, stochastic mathematics and discussions on the normal distribution.

In chapter 3, basic concepts of time series are introduced. Autoregressive Moving Averages time series are the main topic of discussion. Univariate volatility processes literature is reviewed and investigated in section 4 which

[9]Warrants on the JSE must not be confused for an option issued by a company on its own stock which is available in some countries.

Figure 1.1: The moving 30-day standard deviation against the implied volatility of the warrant: 3SAPIB on Sappi. The breaks in the implied volatility graph is due to market illiquidity. The intrinsic value of the replicating portfolio is more than the value of the option.

builds on the ARMA discussion. The most important univariate volatility process is the (vanilla) GARCH process. Other important GARCH processes are also investigated.

Risk-neutral valuation is the basis of modern option pricing. Risk-neutral valuation and continuous time finance is discussed in chapter 5. This discussion leads to the pricing of options in incomplete markets and the LRNVR investigated in chapter 6.

Chapter 7 is about the application of the LRNVR to option pricing. Delta hedging under LRNVR is also investigated.

Monte Carlo simulations and optimization forms part of chapter 8 where the implementation of GARCH option pricing is discussed.

Results are given in chapter 9 and the conclusion follows in chapter 10. Related literature is discussed in section 11.

Part I

Background

Chapter 2

Some Probability Essentials

2.1 Introduction[1]

In this chapter some of the essential background to probability theory is given. Although the background is basic, very few mathematicians, statisticians or probability theorists would be familiar with all the concepts presented.

In section 2.2 the basic concepts concerning a probability space is briefly stated. Moments are discussed in section 2.3.

Cumulative distribution functions and partial density functions are discussed in section 2.4. Some of the main theorems of this section is stated and proved. In section 2.5 the moments and other issues regarding the normal distribution is specified.

A short detour is taken in section 2.6 where returns series are discussed. The section ends with section 2.7 where some important hypothesis tests are discussed.

2.2 Probability Space

2.2.1 Probability Space

The triple $(\Omega, \mathcal{F}, P)$ is called a probability space. The set Ω is a non-empty set, $\mathcal{F} = \{\mathcal{F}_t\}_{t \in N, t \geq 0}$ is filtration of $\sigma-$algebras $\mathcal{F}_t$ defined on Ω and P is a probability measure on $\mathcal{F}$.

A function $Z_t : \Omega \to \mathbb{R}$, on the probability space, is called a stochastic process.

2.2.2 $\sigma-$algebra

A family of subsets $\mathcal{F}$ of a set Ω is called a $\sigma-$algebra if the following holds:

[1]For further discussions on probability theory and measure theoretical aspects see [31], [7], [3] and [27]. [4], [13], [17], [26] and [32] are also useful.

1. $\emptyset \in \mathcal{F}$
2. If $X \in \mathcal{F}$ then $\Omega \backslash X \in \mathcal{F}$
3. If (X_n) is a sequence of sets in Ω then $\bigcup\limits_{n=1}^{\infty} X_n \in \mathcal{F}$.

2.2.3 Borel Sets in $\mathbb{R}$

The Borel sets (one thing) is the smallest $\sigma-$algebra generated by all the open sets in $\mathbb{R}$.

2.2.4 Filtration

$\mathcal{F} = \{\mathcal{F}_t\}_{t\in N, t\geq 0}$ is a filtration of $\sigma-$algebras with the following properties

1. $\mathcal{F}_0$ contains all null sets
2. $\mathcal{F}_t = \cap_{s:s>t}\mathcal{F}_s$ for $t \geq 0$

Property 2 is called right continuous. A filtration with these properties is said to satisfy the "usual conditions".

2.2.5 Measurability and Adaptedness

Definition 2.2.1 *A stochastic process, $Z_t : \Omega \to \mathbb{R}$, is said to be measurable with respect to a σ-algebra $\mathcal{F}$ if*

$$\{Z_t \in B\} \in \mathcal{F}$$

for every Borel set $B \in B(\mathbb{R})$.

Definition 2.2.2 *A stochastic process, Z_t, is said to be adapted to a filtration $\mathcal{F}$ if Z_t is $\mathcal{F}_t$ measurable for all $t \in \mathbb{R}$.*

Remark 2.2.3 *Throughout this dissertation only real-valued stochastic processes defined on $(\Omega, \mathcal{F}, P)$ will be considered, that is $X : \Omega \to \mathbb{R}$.*

Remark 2.2.4 *A stochastic process at a specific time is often referred to as a random variable.*

Remark 2.2.5 *Take note that a stochastic process/random variable is defined in terms of a probability space.*

2.2.6 Almost everywhere

Definition 2.2.6 *Two functions, f and g are equal almost everywhere (sometimes called almost surely) if*

$$f(x) = g(x)$$

for all $x \notin N \in \mathcal{F}$ where $P(N) = 0$. Almost everywhere is abbreviated by a.e.

Definition 2.2.7 *A sequence of functions (f_n) converges to f almost everywhere if there exists a set $N \in \mathcal{F}$ with $P(N) = 0$ such that $f(x) = \lim f_n(x)$ for all $x \notin N$.*

2.3 Moments and Stationarity

2.3.1 Expected Value

Definition 2.3.1 *A random variable $X \in \Omega \to \mathbb{R}$ is said to be integrable if*

$$\int_\Omega |X|\, dP < \infty$$

The family of integrable random variables are denoted by $L^1(\Omega, \mathcal{F}, P)$ or in this dissertation L^1 for short.

Definition 2.3.2 *For any $X \in L^1(\Omega, \mathcal{F}, P)$,*

$$E[X] := \int_\Omega X dP$$

is called the expected value of X.

Remark 2.3.3 *The expected value of a random variable from a symmetric distribution is often called the mean or average.*

Remark 2.3.4 *For a probability space with density function f and integrable Borel function $h : \mathbb{R} \to \mathbb{R}$,*

$$E[h(X)] = \int_{\mathbb{R}} h dP_X = \int_{-\infty}^{\infty} h(x) f(x)\, dx$$

Proof. See Brzezniak et al. [7]. ■

2.3.2 Conditional Expectation

We can call the filtration $\mathcal{F}_t \subset \mathcal{F}$, the $\sigma-$algebra that contains all the information available to an investor at time t. A σ-algebra can also be a condition in a conditional expectation.

Definition 2.3.5 $(\Omega, \mathcal{F}, P)$. *The conditional expectation given a $\sigma-$algebra is: for an integrable random variable X and $\sigma-$algebra $\Phi \subseteq \mathcal{F}$*

$$E[X \mid \Phi]$$

where $E[X \mid \Phi] : \Omega \rightarrow \mathbb{R}$ is unique P a.s. satisfying

1. *$E[X \mid \Phi]$ is Φ measurable*
2. *$\int_H E[X \mid \Phi]\, dP = \int_H X dP$, for all $H \in \Phi$.*

Theorem 2.3.6 *For X, $Y \in L^1$, $a, b \in \mathbb{R}$ and $\sigma - algebra$ $\Phi \subset \mathcal{F}$ the conditional expectation has the following basic properties (all equalities a.s.):*

1. *$E[aX + bY \mid \Phi] = aE[X \mid \Phi] + bE[Y \mid \Phi]$*
2. *$E[E[X \mid \Phi] \mid \Psi] = E[X \mid \Psi]$ where $\Phi \subset \Psi$ is also a $\sigma - algebra$*
3. *$E[X \mid \Phi] = X$ if X is Φ measurable*
4. *$E[X \mid \Phi] = E[X]$ if X is independent of Φ*

Proof. See Appendix B in Oksendal [27]. ■

2.3.3 Variance, Conditional Variance and Standard Deviation

Definition 2.3.7 $(\Omega, \mathcal{F}, P)$. *A random variable $X \in \Omega \rightarrow \mathbb{R}$ is said to be square integrable if*

$$\int_\Omega X^2 dP < \infty$$

The family of square integrable random variables are denoted by $L^2(\Omega, \mathcal{F}, P)$ or L^2 for short.

Definition 2.3.8 $(\Omega, \mathcal{F}, P)$. *The variance of a square integrable random variable X is defined as*

$$\begin{aligned} Var[X] &= E\left[(X - E[X])^2\right] \\ &= E\left[X^2\right] - (E[X])^2 \\ &= \int_H X^2 dP + \left(\int_H X dP\right)^2 > 0 \end{aligned}$$

Definition 2.3.9 $(\Omega, \mathcal{F}, P)$. *The standard deviation of a square integrable random variable X is defined as*

$$Std[X] = \sqrt{Var[X]}$$

The conditional variance and its properties follows directly from that of the conditional expected value:

Definition 2.3.10 $(\Omega, \mathcal{F}, P)$. *For any square integrable random variable X, the conditional expected value of X given a $\sigma - algebra$ $H \subseteq \mathcal{F}$, is*

$$\begin{aligned} Var[X \mid H] &= E\left[(X - E[X \mid H])^2 \mid H\right] \\ &= E\left[X^2 - 2XE[X \mid H] + (E[X \mid H])^2 \mid H\right] \text{ a.s. } P \\ &= E\left[X^2 \mid H\right] - (E[X \mid H])^2 \text{ a.s. } P \end{aligned}$$

Theorem 2.3.11 $(\Omega, \mathcal{F}, P)$. *For a square integrable random variable Z and $a, c \in \mathbb{R}$, the conditional variance of $K = aX + c$, given a $\sigma - algebra$ $H \subseteq \mathcal{F}$, is*

$$Var(K \mid H) = a^2 Var(X \mid H) \text{ a.s. } P$$

Proof. *Equalities almost surely*

$$\begin{aligned} & Var[K \mid H] \\ = {} & E\left[(K)^2 \mid H\right] - (E[K \mid H])^2 \\ = {} & E[a^2X^2 + 2acX + c^2 \mid H] - (E[aX \mid H] + c)^2 \\ = {} & a^2E[X^2 \mid H] + 2acE[X \mid H] + c^2 - a^2(E[X \mid H])^2 \\ & -2ac(E[X \mid H]) - c^2 \\ = {} & a^2E[X^2 \mid H] - a^2(E[X \mid H])^2 \\ = {} & a^2 Var(X \mid H). \end{aligned}$$

■

Theorem 2.3.12 $(\Omega, \mathcal{F}, P)$.*For two square integrable random variable X and Y and $a, b, c \in \mathbb{R}$, the conditional variance of $Z = aX + bY + c$*

$$Var(Z \mid H) = a^2 Var(X \mid H) + b^2 Var(Y \mid H) + 2abCov(X, Y \mid H) \text{ a.s. } P$$

Proof. *From theorem 2.3.11. Equalities almost surely P*

$$\begin{aligned}
& Var\left[aX+bY+c \mid H\right] \\
= \; & Var\left[aX+bY \mid H\right] \\
= \; & E\left[(aX+bY)^2 \mid H\right] - \left(E\left[aX+bY \mid H\right]\right)^2 \\
= \; & a^2E\left[X^2 \mid H\right] + 2abE\left[XY \mid H\right] + b^2E\left[Y^2 \mid H\right] \\
& - \left(aE\left[X \mid H\right] + bE\left[Y \mid H\right]\right)^2 \\
= \; & a^2E\left[X^2 \mid H\right] + 2abE\left[XY \mid H\right] + b^2E\left[Y^2 \mid H\right] - a^2\left(E\left[X \mid H\right]\right)^2 \\
& -2abE\left[X \mid H\right]E\left[Y \mid H\right] - b^2\left(E\left[Y \mid H\right]\right)^2 \\
= \; & a^2Var\left(X \mid H\right) + b^2Var\left(Y \mid H\right) \\
& +2ab\left(E\left[XY \mid H\right] - E\left[X \mid H\right]E\left[Y \mid H\right]\right)
\end{aligned}$$

In the next section covariances will be properly defined, for now assume

$$Cov\left(X,Y \mid H\right) = E\left[XY \mid H\right] - E\left[X \mid H\right]E\left[Y \mid H\right].$$

Thus

$$\begin{aligned}
& Var\left[aX+bY+c \mid H\right] \\
= \; & a^2Var\left(X \mid H\right) + b^2Var\left(Y \mid H\right) + 2abCov\left(X,Y \mid H\right)
\end{aligned}$$

■

2.3.4 Covariance and Autocovariance

Definition 2.3.13 $(\Omega, \mathcal{F}, P)$. *For two square integrable random variables X and Y in our probability space, the covariance between X and Y is*

$$\begin{aligned}
cov\left[X,Y\right] &= E\left[\left(X - E\left[X\right]\right)\left(Y - E\left[Y\right]\right)\right] \\
&= E\left[XY\right] - E\left[X\right]E\left[Y\right]
\end{aligned}$$

Definition 2.3.14 $(\Omega, \mathcal{F}, P)$. *For a square integrable stochastic process $(X_t)_{t\in N}$, adapted to $\mathcal{F}$, the covariance between X_t and X_{t-k} for any $t,k \in N$ is*

$$\begin{aligned}
cov\left[X_t, X_{t-k}\right] &= E\left[\left(X_t - E\left[X_t\right]\right)\left(X_{t-k} - E\left[X_{t-k}\right]\right)\right] \\
&= E\left[X_tX_{t-k}\right] - E\left[X_t\right]E\left[X_{t-k}\right]
\end{aligned}$$

The covariance between elements of the same stochastic process is called the autocovariance.

The conditional covariance and autocovariance can be defined in a similar fashion as the conditional variance, bearing in mind that conditional covariances are random variables.

2.3.5 Correlation and Autocorrelation

Definition 2.3.15 $(\Omega, \mathcal{F}, P)$. *For two square integrable random variables* X *and* Y *the correlation between* X *and* Y *is*

$$cor[X,Y] = \frac{cov[X,Y]}{\sqrt{var[X]\,var[Y]}}$$

Definition 2.3.16 $(\Omega, \mathcal{F}, P)$. *For a square integrable stochastic process* $(X_t)_{t \in N}$, *adapted to* $\mathcal{F}$, *the correlation between* X_t *and* X_{t-k} *for any* $t, k \in N$ *is*

$$cor[X_t, X_{t-k}] = \frac{cov[X_t, X_{t-k}]}{\sqrt{var[X_t]\,var[X_{t-k}]}}$$

The correlation between elements of the same stochastic process is called the autocorrelation.

2.3.6 Lag

Definition 2.3.17 *Consider a stochastic process, say* $(X_t)_{t \in N}$. *At any time step* t *a lag of size* k *is an integer that represents the process at time* $t-k$, X_{t-k}.

2.3.7 Higher Moments

Definition 2.3.18 $(\Omega, \mathcal{F}, P)$. *The* r^{th} *moment of a random variable* X *(about its mean) is*

$$E[(X - E[X])^r]$$

The first moment of a random variable is defined as its mean. The second moment of a random variable is its variance. The second moments of a stochastic process also include the autocovariances. The third moment of a random variable is skewness and the fourth is kurtotsis. For a stochastic process $(X_t)_{t \in N}$ the set of r^{th} moments can be defined as

$$\left\{ E\left[\prod_{i=1}^{r} (X_{k_i} - E[X_{k_i}])\right] \mid \text{for all } k_i \in N \right\}$$

2.3.8 Stationarity

Definition 2.3.19 *A stochastic process is called stationary if all of its moments are constants.*

Definition 2.3.20 *A stochastic process is called weakly stationary if its first and second moments are constant. This means that its mean is constant and for every lag* k *and time* t *the* $cov[X_t, X_{t-k}]$ *is a constant.*

2.4 Cumulative Distribution Function and Probability Density Function

Definition 2.4.1 *The (cumulative) distribution function (cdf) of a random variable $X : \Omega \to \mathbb{R}$ is defined as*

$$F(x) = P\{X \leq x\}$$

Theorem 2.4.2 *The cdf F of a random variable $X : \Omega \to \mathbb{R}$ has the following properties*

1. $0 \leq F \leq 1$
2. $\lim_{x \to -\infty} F(x) = 0$ *and* $\lim_{x \to \infty} F(x) = 1$
3. F *is right-continuous,* $F(x) = \lim_{x_n \to x} F(x_n)$ *for a decreasing sequence* x_n
4. F *is increasing.*

Proof. *See Brzezniak et al. [7].* ■

Theorem 2.4.3 *If $g : \mathbb{R} \to \mathbb{R}$ is integrable then*

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)\,dF(x)$$

Proof. *A result of exercise 1.7 in Brzezniak et al. [7].* ■

Theorem 2.4.4 *A measurable function $f(x) \geq 0$ on $\mathbb{R}$ such that for any Borel measurable set B*

$$P\{X \in B\} = \int_B f(y)\,dy$$

is called the (probability) density function (pdf) of X. The pdf can in particular also be written in terms of the cdf of X, F

$$F(x) = \int_{-\infty}^{y} f(y)\,dy$$

Proof. *See Brzezniak et al. [7].* ■

Theorem 2.4.5 *If X has a continuous pdf f then*

$$\frac{d}{dx}F(x) = f(x)$$

Proof. *A result from the fundamental theorem of calculus.* ■

2.4.1 Joint Continuous Distributions

The joint distribution of a k-dimensional random variable,

$$\mathbf{X} = (X_1, X_2, ..., X_k),$$

is a measure $P_{\mathbf{X}}$ on $\mathbb{R}^n$ such that for any Borel set, $\mathbf{B} \in \mathbb{R}^n$

$$P_{\mathbf{X}}(\mathbf{B}) = P\{X \in \mathbf{B}\}$$

If the random variables of X are independently distributed then

$$P_{\mathbf{X}}(\mathbf{B}) = \prod_{i=1}^{k} P\{X_i \in B_i\}$$

where

$$\mathbf{B} = \begin{bmatrix} B_1 \\ ... \\ B_k \end{bmatrix}.$$

Definition 2.4.6 *The joint probability density function (joint pdf) of a k-dimensional random variable,*

$$\mathbf{X} = (X_1, X_2, ..., X_k),$$

is a Borel function

$$f(x_1, x_2, ..., x_k) : \mathbb{R}^n \to \mathbb{R}$$

such that

$$P_{\mathbf{X}}(\mathbf{B}) = \int_{\mathbf{B}} f(t_1, t_2, ..., t_k)\, dt_1 ... dt_k \tag{2.1}$$

Definition 2.4.7 *The joint cumulative distribution function (joint cdf) of a k-dimensional random variable,*

$$\mathbf{X} = (X_1, X_2, ..., X_k),$$

is

$$F(x_1, ..., x_k) = P[X_1 \le x_1, ..., X_k \le x_k]$$

If the random variables of $\mathbf{X}$ are independently distributed then

$$f(x_1, ..., x_k) = f(x_1) ... f(x_k)$$

and

$$F(x_1, ..., x_k) = F(x_1) ... F(x_k)$$

which follows directly from the case of independence of $P_{\mathbf{X}}$.

Theorem 2.4.8 *The joint cdf of a k-dimensional random variable* **X** *can be written in terms of the joint pdf of as follows*

$$F(x_1, x_2, ..., x_k) = \int_{-\infty}^{x_k} ... \int_{-\infty}^{x_1} f(l_1, l_2, ..., l_k)\, dl_1 ... dl_k \tag{2.2}$$

Proof. From definition 2.4.7 and the fact that $(-\infty, x_i]$ is a Borel set for every applicable i it is clear that the joint cdf of X is a special case of the joint probability of X. Equation 2.2 follows directly from 2.1. ■

Theorem 2.4.9 *If* X, $Y \in \mathbb{R}$ *are independent random variables and* $g(x)$ *and* $h(y)$ *are functions then*

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

and

$$cov[g(X), h(Y)] = 0$$

Proof. With a joint pdf $f(x, y)$

$$\begin{aligned} E[g(X)h(Y)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f(x,y)\,dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_1(x)f_2(y)\,dxdy \end{aligned}$$

due to independence. The cdfs of X and Y are f_1 and f_2 respectively, then

$$\begin{aligned} E[g(X)h(Y)] &= \int_{-\infty}^{\infty} h(y)f_2(y)\int_{-\infty}^{\infty} g(x)f_1(x)\,dxdy \\ &= \int_{-\infty}^{\infty} g(x)f_1(x)\,dx\int_{-\infty}^{\infty} h(y)f_2(y)\,dy \\ &= E[g(X)]E[h(Y)] \end{aligned}$$

The covariance can be expressed as

$$\begin{aligned} & cov[g(X), h(Y)] \\ =\ & E[g(X)h(Y)] - E[g(X)]E[h(Y)] \\ =\ & 0 \end{aligned}$$

■

2.5 The Normal Distribution and its Moment Generating Function

2.5.1 The Normal Distribution

The normal distribution, the most frequently used statistical distribution, was first published by Abraham de Moivre (1733).

A normal random variable $X \in \mathbb{R}$, with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}^+$ is denoted by

$$X \sim N\left(\mu, \sigma^2\right)$$

The probability density function (pdf) of the normal distribution is

$$f\left(x; \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Another way to define the pdf in terms of the probability space $(\Omega, \mathcal{F}, P)$ is as follows

$$P(A) = \int_A dP = \int_A f\left(x; \mu, \sigma^2\right) dx$$

The cumulative distribution function (cdf) of the normal distribution is given by

$$\begin{aligned} F\left(z; \mu, \sigma^2\right) &= P\{X \leq z\} \\ &= \int_{\{X \leq z\}} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx \\ &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx \end{aligned}$$

The standard normal distribution, frequently used in this dissertation is defined as the normal distribution with zero mean and a variance of one, $N(0,1)$. The standard normal distribution's pdf is

$$f(x; 0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

and cdf is

$$F(x; 0, 1) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx.$$

2.5.2 Moments of the Normal Distribution

Consider a normally distributed random variable $X \sim N\left(\mu, \sigma^2\right)$ with probability density function

$$f\left(x; \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

then the random variable

$$Y = \frac{X-\mu}{\sigma} \sim N(0,1)$$

Definition 2.5.1 *The moment generating function of X is defined as*

$$M_X(t) = E\left(e^{tX}\right)$$

Theorem 2.5.2 *The moment generating function of $X \sim N\left(\mu, \sigma^2\right)$ is*

$$M_X(t) = \exp\left(\mu t + \frac{(\sigma t)^2}{2}\right)$$

Proof.

$$\begin{aligned}
M_X(t) &= E\left(e^{tX}\right) \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{tx} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[tx - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{x^2 - 2\left(\mu + \sigma^2 t\right)x + \mu^2}{-2\sigma^2}\right) dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{\left(x - \mu + \sigma^2 t\right)^2 - 2\mu\sigma^2 t - \left(\sigma^2 t\right)^2}{-2\sigma^2}\right) dx \\
&= \exp\left(\mu t + \frac{(\sigma t)^2}{2}\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{\left(x - \mu + \sigma^2 t\right)^2}{-2\sigma^2}\right) dx \\
&= \exp\left(\mu t + \frac{(\sigma t)^2}{2}\right)
\end{aligned}$$

■

Theorem 2.5.3 *Moments about the mean of $X \sim N\left(\mu, \sigma^2\right)$. If r is even then*

$$E\left[(X-\mu)^r\right] = \frac{(2r)!\sigma^{2r}}{r!2^r},$$

if r is odd then

$$E\left[(X-\mu)^r\right] = 0$$

Proof. The

$$
\begin{aligned}
M_{X-\mu} &= \exp\left(\frac{(\sigma t)^2}{2}\right) \\
&= \sum_{n=0}^{\infty} \frac{\left(\frac{(\sigma t)^2}{2}\right)^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{\sigma^{2n} t^{2n}}{2^n n!} \\
&= \sum_{n=0}^{\infty} \frac{\sigma^{2n} t^{2n} (2n!)}{2^n n! \, (2n!)} \\
&= \sum_{n=0}^{\infty} \frac{\sigma^{2n} (2n!)}{2^n n!} \frac{t^{2n}}{(2n!)}
\end{aligned}
$$

The second line is due to the Maclaurin series expansion for e. Note that only positive integers are contained in the last line Theorem 2.5.1 in [2] states that

$$M_V(t) = 1 + \sum_{n=1}^{\infty} E[V^n] \frac{t^n}{n!}$$

for a random variable V. Thus

$$E[(X-\mu)^r] = 0$$

if r is odd and

$$E[(X-\mu)^r] = \frac{\sigma^r r!}{2^{r/2} (r/2)!}$$

if r is even. ■

The following characteristics of random variable $X \sim N(\mu, \sigma^2)$ follows from theorem 2.5.3:

1. The skewness of X is

$$E\left[(X-\mu)^3\right] = 0$$

2. The kurtosis of X is

$$
\begin{aligned}
E\left[(X-\mu)^4\right] &= \frac{\sigma^4 4!}{2^2 (2)!} \\
&= 3\sigma^4
\end{aligned}
$$

and thus if $\sigma^2 = 1$

$$E\left[(X-\mu)^4\right] = 3$$

2.5.3 Chi-square Distribution

Definition 2.5.4 *If random variable Y is chi-square distributed with v degrees of freedom then*

$$Y \sim \chi^2(v)$$

where the chi-square distributed is a special case of the gamma distribution

$$\chi^2(v) \sim GAM\left(2, \frac{v}{2}\right)$$

Theorem 2.5.5 *A random variable $Y \sim \chi^2(v)$ has the following characteristics*

1. *Probability density function*

$$f(y) = \frac{1}{2^{v/2}\Gamma(v/2)} y^{v/2-1} e^{-x/2},$$

where Γ is the gamma function

$$\Gamma(\kappa) = \int_0^\infty t^{k-1} e^{-t} dt$$

for all $\kappa > 0$.

2. *Moment generating function*

$$M_Y(t) = (1-2t)^{-v/2}$$

3. *Moments about the mean*

$$E[Y^r] = 2^r \frac{\Gamma(v/2+r)}{\Gamma(v/2)}$$

4. *Expected value*

$$E[Y] = v$$

5. *Variance*

$$Var[Y] = 2v$$

Proof. *Results follow from the gamma distribution. See Bain [2]* ■

Theorem 2.5.6 *If*

$$X \sim N\left(\mu, \sigma^2\right)$$

then

$$Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi^2(1)$$

Proof. *The moment generating function of* Z^2

$$\begin{aligned} M_{Z^2} &= E\left[e^{tZ^2}\right] \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(tz^2\right) \exp\left(-\frac{1}{2}z^2\right) dz \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(tz^2 - \frac{1}{2}z^2\right) dz \\ &= \frac{1}{\sqrt{1-2t}} \int_{-\infty}^{\infty} \frac{\sqrt{1-2t}}{\sqrt{2\pi}} \exp\left(\frac{z^2(1-2t)}{2}\right) dz \\ &= \frac{1}{\sqrt{1-2t}} \end{aligned}$$

which is the moment generating function of the chi-square distribution. ■

Definition 2.5.7 *If* $Z^2 \sim \chi^2(1)$ *then*

$$(Z-\lambda)^2$$

is noncentral chi-square distributed with 1 *degree of freedom and non-centrality parameter* λ.

Theorem 2.5.8 *The expected value of a noncentral chi-square distributed random variable is*

$$E\left[(Z-\lambda)^2\right] = 1 + \lambda^2$$

where $Z^2 \sim \chi^2(1)$.

Proof.

$$\begin{aligned} E\left[(Z-\lambda)^2\right] &= E\left[Z^2 - 2\lambda Z + \lambda^2\right] \\ &= E\left[Z^2\right] - 2\lambda E[Z] + \lambda^2 \\ &= 1 + \lambda^2 \end{aligned}$$

since $E[Z] = 0$ because $Z \sim N(0,1)$. ■

i16542502
b15954341

Theorem 2.5.9 *The expected value of a noncentral chi-square distributed random variable is*

$$E\left[(Z-\lambda)^4\right] = 3 + 6\lambda^2 + \lambda^4$$

where $Z^2 \sim \chi^2(1)$.

Proof.

$$\begin{aligned} E\left[(Z-\lambda)^4\right] &= E\left[\left(Z^2 - 2\lambda Z + \lambda^2\right)^2\right] \\ &= E\left[Z^4 - 4Z^3\lambda + 6Z^2\lambda^2 - 4\lambda^3 Z + \lambda^4\right] \\ &= E\left[Z^4\right] - 4\lambda E\left[Z^3\right] + 6\lambda^2 E\left[Z^2\right] - 4\lambda^3 E[Z] + \lambda^4 \\ &= 3 + 6\lambda^2 + \lambda^4 \end{aligned}$$

This is done by remembering that

$$Z \sim N(0,1)$$

thus the expected value of Z is

$$E[Z] = 0$$

the skewness is

$$E\left[Z^3\right] = 0$$

and the kurtosis is

$$E\left[Z^4\right] = 3$$

■

2.6 The Return Series and Lognormal Distribution

2.6.1 Returns Series

The financial value of a company or fund is represented by its (stock) price. The stock price has a clear, time dependent trend. It is hard to model series with trends, at least in an objective, scientific sense. To remove this trend, the financial time series is transformed into a series with "manageable" mean, a returns series. This is done with difference equations.

It will be proved that the returns series still has the same variance as the original series. The returns series is of great importance in risk management and derivatives pricing.

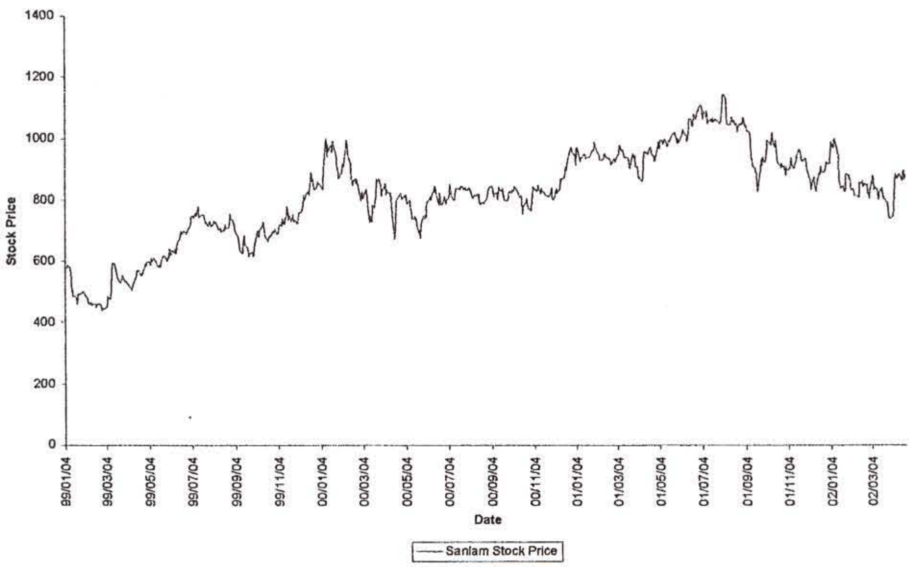

Figure 2.1: The stock price of Sanlam from 1999/01/05 to 2002/04/19.

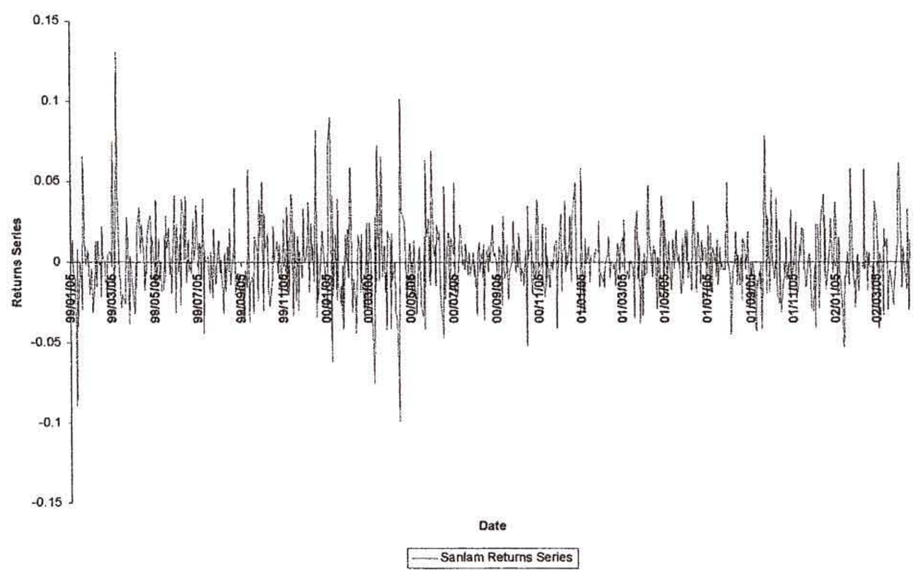

Figure 2.2: The returns series of Sanlam from 1999/01/05 to 2002/04/19.

2.6.2 The Arithmetic Returns Series

Definition 2.6.1 *The arithmetic returns series, for process* (S_t) *is defined as*

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}}$$

2.6.3 The Geometric Returns Series

The geometric returns series, for process (S_t) is defined as

$$\begin{aligned} r_t &= lnS_t - lnS_{t-1} \\ &= ln\left(\frac{S_t}{S_{t-1}}\right) \end{aligned}$$

The relationship between the geometric and arithmetic series, by the Taylor series expansion, are as follows:

$$\begin{aligned} ln\left(\frac{S_t}{S_{t-1}}\right) &= ln\left(\frac{S_t}{S_{t-1}} - 1 + 1\right) \\ &= ln\left(\frac{S_t - S_{t-1}}{S_{t-1}} + 1\right) \\ &\approx \frac{S_t - S_{t-1}}{S_{t-1}} \end{aligned}$$

if $\left|\frac{S_t - S_{t-1}}{S_{t-1}}\right| < 1$.

The geometric returns series will be considered in this dissertation.

Theorem 2.6.2 *If we assume that a returns series is normally distributed then the log of the stock process is also normal, and vice versa.*

Proof.

$$\begin{aligned} ln\left(\frac{S_1}{S_0}\right) + ln\left(\frac{S_2}{S_1}\right) + ln\left(\frac{S_2}{S_1}\right) + ... + ln\left(\frac{S_t}{S_{t-1}}\right) &= ln\left(\frac{S_t}{S_0}\right) \\ &= lnS_t - lnS_0 \end{aligned}$$

The sum of normally distributed random variables are also normal and we assume that S_0 is known. ∎

2.6.4 Lognormal Distribution

$(\Omega, \mathcal{F}, P)$. A random variable $X \in \mathbb{R}$, with mean μ and variance σ^2 is said to be lognormally distributed if $\ln(X)$ is normally distributed.

It's often observed that stock prices are lognormally distributed. In chapter 5.4.1 we deduce, given the assumed process 5.10, that a stock price

S_t can be defined in terms of an initial stock price S_0 and Brownian motion W_t ,

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right) \tag{2.3}$$

Taking the logarithm on both sides of equation 2.3 yields

$$\ln S_t \sim N\left(\ln S_0 + \left(\mu - \frac{1}{2}\sigma^2\right)t, \sigma^2 t\right)$$

The return on stock S_t is defined as $\ln\left(\frac{S_t}{S_{t-1}}\right)$ which clearly has the distribution

$$\ln \frac{S_t}{S_{t-1}} \sim N\left(\mu - \frac{1}{2}\sigma^2, \sigma^2\right)$$

It is possible to test with the Jarque-Bera test for normality whether the return is in reality normally distributed.

2.7 Hypothesis Testing[2]

Hypothesis tests are done to verify whether the properties of an observed series, say $\{\hat{\varepsilon}_t\}_{t\in N}$, are consistent with assumed properties under a model The properties that need to be tested include tests for normality, autocorrelation and heteroscedasticity.

The formal procedure for conducting a hypothesis test involves a statement of the null hypothesis and an alternative hypothesis. The sample estimate on which the decision to reject or not reject the null hypothesis comes from the sample space. The Neyman-Pearson methodology [20] involves partitioning the sample space into two regions. If the sample estimate falls in the critical region, the null hypothesis is rejected. If it falls in the acceptance region, it's not.

2.7.1 Jarque-Bera Test for Normality

The Jarque-Bera tests whether observations are not likely to have come from the normal distribution.

Define for n observations the following

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_t^2, \tag{2.4}$$

$$\hat{\mu}_3 = \frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_t^3, \tag{2.5}$$

[2]Suggested reading: [1], [2], [18] and [24].

$$\hat{\mu}_4 = \frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_t^4 \tag{2.6}$$

In equations 2.4 to 2.6 are the second, third and fourth moments of $\hat{\varepsilon}_t$ respectively.

The skewness is defined as

$$skewness = s = \frac{\hat{\mu}_3}{\hat{\sigma}^2}$$

and the kurtosis as

$$kurtosis = k = \frac{\hat{\mu}_4}{\hat{\sigma}^2}$$

The Jarque-Bera statistic is defined as

$$\begin{aligned} JB &= n\left(\frac{\hat{\mu}_3^2}{6\hat{\sigma}^6} + \frac{\left(\hat{\mu}_4 - 3\hat{\sigma}^4\right)^2}{24\hat{\sigma}^8}\right) \\ &= \frac{n}{6}\left(s^2 + 2\left(k-3\right)^2\right) \end{aligned}$$

$$JB \sim \chi^2\left(2\right)$$

The null hypothesis is

$$H_0 : s = 0 \; and \; k = 3$$

against the alternative

$$H_1 : reject\, H_0$$

Remark 2.7.1 *Many text books and computer packages calculates the adjusted kurtosis, that is the $k - 3$.*

2.7.2 Autocorrelation

Durbin-Watson

The most famous test for autocorrelation is the Durbin-Watson test

$$\frac{\sum_{r=2}^{t}\left(\hat{\varepsilon}_r - \hat{\varepsilon}_{r-1}\right)^2}{\sum_{r=1}^{t}\hat{\varepsilon}_r^2}$$

No exact distribution for this test is available.

Ljung-Box

For a series with m observations the Ljung-Box statistic over K lags is

$$m(m+2)\sum_{k=1}^{K}\frac{\tilde{\rho}_k^2}{m-k} \sim \chi^2(K)$$

where $\tilde{\rho}_k^2$ is the observed autocorrelation at lag k given by

$$\tilde{\rho}_k = \frac{\sum_{\tau=k+1}^{t}\left(\hat{\varepsilon}_t^2 - \tilde{\sigma}^2\right)\left(\hat{\varepsilon}_{t-k}^2 - \tilde{\sigma}^2\right)}{\sum_{\tau=1}^{t}\left(\hat{\varepsilon}_t^2 - \tilde{\sigma}^2\right)^2}$$

where $\hat{\varepsilon}_t$ is the observed return at time t and $\tilde{\sigma}^2$ is the sample variance.

2.7.3 Volatility Clustering

Many financial time series and also the Black-Scholes option pricing model make the assumption of constant volatility. Empirical evidence indicates that volatility of financial instruments tends to be dynamic. Volatility levels tend to alternate between periods of higher volatility and more tranquil periods. This clustering together of volatility levels for a period of time is called volatility clustering. Volatility clustering is due to the strong autocorrelation of squared returns or absolute returns. The Box-Pierce Lagrange multiplier test for the significance of first-order autocorrelation in squared returns, $\hat{\varepsilon}_t^2$, is

$$\frac{\sum_{t=2}^{T}\hat{\varepsilon}_t^2\hat{\varepsilon}_{t-1}^2}{\sum_{t=2}^{T}\hat{\varepsilon}_t^4}$$

The Lagrange multiplier tests are chi-squared distributed with T degrees of freedom.

2.7.4 The Leverage Effect

Volatility tends to be higher in a falling market, than in a rising market. Similarly volatility tends to be higher after a large negative return than after a large positive return, for an individual stock. The reason for this is that when a stock price falls, the leverage or debt/equity ratio increases. In laymen's terms, the part of the company's assets "owned" by the creditors increases, leaving less for the shareholders. This causes more uncertainty in the stock price.

An asymmetric version of the Lagrange multiplier test is used to investigate the influence of the leverage effect, and asymmetric returns levels in general

$$\frac{\sum_{t=2}^{T}\hat{\varepsilon}_t^2\hat{\varepsilon}_{t-1}}{\sum_{t=2}^{T}\hat{\varepsilon}_t^2\hat{\varepsilon}_{t-1}^2}$$

where $\hat{\varepsilon}_t$ is the observed return at time t.

Chapter 3

An Introduction to Time Series Models

3.1 Objectives[1]

The purpose of this introduction to Autoregressive Moving Averages (ARMA) time series is to provide enough background to the reader to understand and appreciate the more advanced models in later chapters. For a more complete discussion on ARMA time series see Ferreira [16].

3.2 Preliminaries

3.2.1 White Noise

A white noise series is often part of a time series in the form of an "error", an unpredictable randomness.

Definition 3.2.1 *A white noise series* (ε_t) *has the following characteristics for every* $t, s \in \mathbb{R}$

1. $E[\varepsilon_t] = 0$
2. $E[\varepsilon_t^2] = \sigma^2$
3. $E[\varepsilon_t \varepsilon_s] = 0$ *for* $s \neq t$

The white noise process is thus stationary.

[1] Suggested reading: [1] and [18]

3.2.2 Linear Time Series

Definition 3.2.2 $(\Omega, \mathcal{F}, P)$. *A linear time series at time t consists of a $\mathcal{F}_{t-1}$ predictable part plus a random part, that is for a time series*

$$Z_t = E\left[Z_t \mid \mathcal{F}_{t-1}\right] + \nu_t$$

where the expected value of the white noise process, ν_t where

$$E\left[\nu_t \mid \mathcal{F}_{t-1}\right] = 0$$

3.2.3 Lag Operators and Difference Operators

Definition 3.2.3 *A lag operator L is defined by*

$$L^k Z_t = Z_{t-k}$$

for all $k \in \mathbb{R}^+$.

Definition 3.2.4 *A difference operator Δ is defined by*

$$\Delta_k Z_t = Z_t - Z_{t-k}$$

for all $k \in \mathbb{R}^+$.

Example 3.2.5 *The power of a difference operator Δ^k is different from a higher order difference operator Δ_k.*

$$\begin{aligned}
\Delta^2 Z_t &= \Delta\left(Z_t - Z_{t-1}\right) \\
&= \Delta Z_t - \Delta Z_{t-1} \\
&= Z_t - 2Z_{t-1} + Z_{t-2}
\end{aligned}$$

Definition 3.2.6 *Invertibility of a time series: A time series (Z_t) is invertible if it is possible to write it in terms of an infinite combination of lags.*

3.3 Autoregressive Process (AR)

Definition 3.3.1 *For a stochastic process (Z_t) and white noise process (ε_t), the $AR(p)$ process is defined by*

$$\Phi_p(L) Z_t = \varepsilon_t$$

with

$$\Phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p \tag{3.1}$$

L is a lag operator and p the order of the autoregression polynomial 3.1.

The $AR(p)$ process (Z_t) can thus be written as

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t$$

3.4 Moving Averages Process (MA)

Definition 3.4.1 *For a stochastic process* (Z_t) *and white noise process* (ε_t)*, the* $MA(q)$ *process is defined by*

$$Z_t = \Theta_q(L)\,\varepsilon_t$$

with

$$\Theta_q(L) = 1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q \tag{3.2}$$

where L *is a lag operator and* q *the order of the moving averages polynomial 3.2.*

The $MA(q)$ process (Z_t) can thus be written as the sum of past errors

$$Z_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + .. + \theta_q\varepsilon_{t-q} \tag{3.3}$$

The lag operator thus acts on the white noise process not on Z_t.

3.5 Autoregressive Moving Averages (ARMA)

Definition 3.5.1 *For a stochastic process* (Z_t) *and white noise process* (ε_t)*, the* $ARMA(p,q)$ *process is defined by*

$$\Phi_p(L)\,Z_t = \Theta_q(L)\,\varepsilon_t$$

with

$$\begin{aligned}\Theta_q(L) &= 1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q \\ \Phi_p(L) &= 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p\end{aligned}$$

where L *is a lag operator,* p *the order of the autoregression polynomial and* q *the order of the moving averages polynomial*

The $ARMA(p,q)$ process (Z_t) is

$$\begin{aligned}Z_t &= \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \ldots + \theta_q\varepsilon_{t-q} \\ &= \sum_{i=1}^{p}\phi_i Z_{t-i} + \sum_{i=1}^{q}\theta_i\varepsilon_{t-i}\end{aligned}$$

where $\theta_0 = 1$. It is clear that the $ARMA(p,q)$ process, is a combination of an $AR(p)$ and an $MA(q)$ process.

3.6 Stationarity of ARMA Processes

The results in this section was proved in Ferreira [16].

An $MA(\infty)$ process

$$Z_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots$$

is stable if and only if its weights are square summable

$$\sum_{i=0}^{\infty}\theta_i^2 < \infty$$

The $AR(p)$ process

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t \tag{3.4}$$

can be rewritten in terms of the *Vector Autoregressive process* denoted by $VAR(1)$

$$\begin{bmatrix} Z_t \\ Z_{t-1} \\ Z_{t-2} \\ \vdots \\ Z_{t-(p-1)} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-2} \\ Z_{t-3} \\ \vdots \\ Z_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or

$$\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t$$

From this equation we can obtain

$$\boldsymbol{\xi}_t = \mathbf{F}^t\boldsymbol{\xi}_0 + \mathbf{F}^{\mathbf{t-1}}\boldsymbol{\varepsilon}_1 + \ldots + \mathbf{F}\boldsymbol{\varepsilon}_{t-1} + \mathbf{v}_t$$

Theorem 3.6.1 *If all eigenvalues of the matrix* $\mathbf{F}$ *lie within the unit circle,* $|\lambda| < 1$, *then*

$$\sum_{j=0}^{\infty}\mathbf{F}^j = (\mathbf{I} - \mathbf{F})^{-1} \tag{3.5}$$

where $\mathbf{I}$ *is the applicable identity matrix and the right-hand side of equation 3.5 is the inverse of* $\mathbf{I} - \mathbf{F}$.

Proof. Ferreira [16]. ■

Theorem 3.6.2 *If all the eigenvalues of the $p \times p$ matrix $\mathbf{F}$ lie within the unit circle, then*

$$(\mathbf{I}_p - \mathbf{F})^{-1}$$

exists and its element $(1,1)$ is

$$\frac{1}{1 - \phi_1 - \phi_2 - \ldots - \phi_p}$$

Proof. Ferreira [16]. ■

Corollary 3.6.3 *If all the eigenvalues of $\mathbf{F}$ are less than 1 in magnitude then $\mathbf{F}^j$ decays to zero as j increases to infinite. A time series with such a property is said to be stable.*

Process 3.4 can be rewritten as

$$\Phi(L)\, Z_t = \varepsilon_t$$

where

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p \tag{3.6}$$

Definition 3.6.4 *The characteristic function of the process 3.6 is defined by*

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p \tag{3.7}$$

We can then combine the ideas of the root of polynomial 3.7 and the eigenvalues of $\mathbf{F}$.

Theorem 3.6.5 *Factoring the characteristic function is equivalent to finding the eigenvalues of the matrix $\mathbf{F}$*

$$1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p = (1 - \lambda_1 L)(1 - \lambda_2 L) \ldots (1 - \lambda_p L)$$

Proof. Ferreira [16]. ■

Corollary 3.6.6 *The process 3.4 is stable if all the eigenvalues of $\mathbf{F}$ all lie inside of the unit circle.*

Theorem 3.6.7 *The characteristic function $\Phi(L)$ of an $AR(p)$ process can be written in terms of a characteristic function of a $MA(\infty)$ process, say $\pi(L)$*

$$\Phi(L) = \pi(L)^{-1}$$

Remark 3.6.8 *Note that only $\Phi(L)$, the characteristic function of the autoregressive terms influence stability.*

The results of this section is summarized as follows:

Summary 3.6.9 *An $AR(p)$ process is stationary if and only if the eigenvalues of the characteristic function of that process lie inside the unit circle.*

3.7 Estimation of ARMA Parameters

This section focusses on the *maximum likelihood estimation* (MLE) of the ARMA regression model. If we assume that the error process

$$\varepsilon_t = Z_t - \left(\phi_1 Z_{t-1} + \ldots + \phi_p Z_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q}\right)$$

is normally distributed. Then the likelihood function of the ARMA process is

$$\begin{aligned} f^*(\boldsymbol{\theta}) &= \prod_{i=p+1}^{n} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(-\left[\varepsilon_i/\sigma_\varepsilon\right]^2/2\right) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}}\right)^n \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\sum_{i=1}^{n}\varepsilon_i^2\right) \end{aligned}$$

where σ_ε^2 is the unconditional (stationary) variance of the error process (ε_t). The product is from the $(p+1)^{th}$ observation to the n^{th} since there are p parameters. Define $n' = n - p$.

Define the parameters matrix by

$$\boldsymbol{\theta} = \left(\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \ldots, \theta_q\right)'$$

The loglikelihood function (the ln of $f^*(\boldsymbol{\theta})$) is

$$f(\boldsymbol{\theta}) = -\frac{1}{2\sigma_\varepsilon^2}\sum_{i=p+1}^{n}\varepsilon_i^2.$$

The MLE parameters are those that maximizes $f^*(\boldsymbol{\theta})$ or $f(\boldsymbol{\theta})$ over a number of observations of (ε_t). Since only the error process is variate in terms of the parameters $\boldsymbol{\theta}$, maximizing $f(\boldsymbol{\theta})$ is equivalent to minimizing $\sum_{i=p+1}^{n}\varepsilon_i^2$.

To comment on the *significance* of the MLE parameter fit, define the information matrix

$$\mathbf{I} = -\lim_{n\to\infty} E\left[\frac{1}{n'}\frac{\partial^2 f(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]$$

The asymptotic distribution of MLE estimators is

$$\boldsymbol{\theta} \sim N\left(\boldsymbol{\theta}_0, \frac{1}{n'}\mathbf{I}^{-1}\right)$$

with $\mathbf{I}$ positive definite in the region of the optimal $\boldsymbol{\theta}_0$.

For the second derivative of $f(\boldsymbol{\theta})$ define

$$\mathbf{S} = \frac{\partial^2 f(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} = -\frac{1}{2\sigma_\varepsilon^2}\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\sum_{i=p+1}^{n}\varepsilon_i^2$$

thus we can approximate the covariance matrix of $\boldsymbol{\theta}$,

$$\begin{aligned} var\left(\boldsymbol{\theta}\right) &= \frac{1}{n'}\mathbf{I}^{-1} \\ &\approx 2\sigma_{\varepsilon}^{2}\mathbf{S} \end{aligned}$$

Chapter 4

Univariate Volatility Processes

4.1 Objectives[1]

A univariate model assumes only one source of randomness, in the case of volatility models the source of randomness is the conditional returns. Define, under measure P, the conditional returns as

$$\varepsilon_t = ln\frac{S_t}{S_{t-1}}$$

In this chapter two of the main univariate volatility processes are discussed. The Exponentially Weighted Moving Averages (EWMA) process is discussed in section 4.2 and the various GARCH processes is discussed in section 4.3 and further. This chapter includes a discussion on Asymmetric GARCH in section 4.7.

4.2 Exponentially Weighted Moving Averages

Weighing the $MA(q)$ process in equation 3.3, by the sum of its parameters yields

$$Z_t = \frac{\varepsilon_t + \lambda\varepsilon_{t-1} + \lambda^2\varepsilon_{t-2} + .. + \lambda^q\varepsilon_{t-q}}{1 + \lambda + \lambda^2 + .. + \lambda^q} \tag{4.1}$$

where $\theta_i = \lambda^i$ and $\lambda \in (0,1)$.

Taking the limit of 4.1 to infinite

$$\begin{aligned} \lim_{q\to\infty} Z_t &= \lim_{q\to\infty} \frac{\varepsilon_t + \lambda\varepsilon_{t-1} + \lambda^2\varepsilon_{t-2} + .. + \lambda^q\varepsilon_{t-q}}{1 + \lambda + \lambda^2 + .. + \lambda^q} \\ &= (1-\lambda)\sum_{i=1}^{\infty}\lambda^i\varepsilon_{t-i} \end{aligned} \tag{4.2}$$

[1]Suggested reading: [1], [18] and [23].

since $\lambda \in (0, 1)$.

Equation 4.2 is the basis of the EWMA conditional variance process,

$$\begin{aligned}\hat{\sigma}_t^2 &= (1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}\varepsilon_{t-i}^2 \\ &= (1-\lambda)\sum_{i=2}^{\infty}\lambda^{i-1}\varepsilon_{t-i}^2 + (1-\lambda)\,\varepsilon_{t-1}^2 \\ &= \lambda(1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-2}\varepsilon_{(t-1)-i}^2 + (1-\lambda)\,\varepsilon_{t-1}^2 \\ &= \lambda\sigma_{t-1}^2 + (1-\lambda)\,\varepsilon_{t-1}^2 \qquad (4.3)\end{aligned}$$

with $\lambda \in (0, 1)$.

Alexander [1] interprets the smoothing constant λ in the following two ways:

1. The term, $(1-\lambda)\,\varepsilon_{t-1}^2$ determines the *intensity of reaction* of volatility to market events. A low value of λ will give a process highly reactive to shocks. The effect of these shocks will quickly die away. Lower values of λ is mostly used for short term forecasts.

2. Term $\lambda\sigma_{t-1}^2$ determines the *persistence in volatility.* A high λ will give a process that persists at a certain level of volatility, despite recent shocks.

Parameters of the EWMA process can be estimated by minimizing the root mean square error or similar method. The accuracy of forecasts are however difficult to assess.

4.2.1 RiskMetrics

The EWMA model is also the basis of volatility forecasts in the RiskMetrics system by J.P. Morgan. The RiskMetrics model has the following to distinctive features:

1. The parameter λ is fixed, $\lambda = 0.94$.

2. The definition of volatility is different than the standard definition of volatility. Under the assumption of normality, the RiskMetrics volatility is the 95^{th} percentile or 1.65 times the standard deviation.

4.3 Generalized Conditional Autoregressive Conditional Heteroscedasticity

The Autoregressive Conditional Heteroscedastic (ARCH) process was introduced by Engle (1982) [14]. This process allows for the change of conditional volatility over time as a function of past errors.

The Generalized Autoregressive Conditional Heteroscedastic (GARCH) process by Bollerslev (1986) [6] is the most popular and widely used stochastic volatility measure and forecasting method.

The $GARCH(p,q)$ process is discussed in section 4.4 below. It will be shown that this discussion encompasses the ARCH process in a simple way. The GARCH process is also the basis for many subsequent models.

4.4 GARCH(p,q)

The GARCH(p,q) process under conditionally normal, discrete time errors, is defined by

$$\begin{aligned} \varepsilon_t \quad | \quad & \mathcal{F}_{t-1} \sim N\left(0, \sigma_t^2\right) \\ \sigma_t^2 \quad = \quad & \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \end{aligned} \tag{4.4}$$

where p, q are nonnegative integers, α_i, β_j are nonnegative real numbers for every applicable i, j and α_0 is a positive real.

For $p, q = 0$, the GARCH process is simple white noise. For $p = 0$, $q \neq 0$ the process is an ARCH process. Thus, the GARCH process is to volatility what the ARMA process is to the AR process, for means.

Any GARCH(p,q) process can be defined as a GARCH$(1,1)$ process. Define

$$\sigma_t^2 = \alpha_0 + A(L)\varepsilon_t^2 + B(L)\sigma_t^2$$

where for lag operator L,

$$\begin{aligned} A(L) &= \sum_{i=1}^{q} \alpha_i L^i \\ B(L) &= \sum_{i=1}^{p} \beta_i L^i \end{aligned}$$

4.4.1 Stationarity

Theorem 4.4.1 *A $GARCH(p,q)$ process is stationary, with (long-term) variance*

$$E\left[\sigma_t^2\right] = \frac{\alpha_0}{1 - A(1) - B(1)}$$

for any t if and only if $A(1)+B(1)<1$.

Proof. For any t

$$\begin{aligned} E\left[\sigma_t^2\right] &= E\left[var\left[\varepsilon_t \mid \mathcal{F}_{t-1}\right]\right] \\ &= E\left[E\left[\varepsilon_t^2 \mid \mathcal{F}_{t-1}\right]\right] \end{aligned}$$

since we assume that $E\left[\varepsilon_t \mid \mathcal{F}_{t-1}\right]=0$. It follows that

$$E\left[\sigma_t^2\right]=E\left[\varepsilon_t^2\right]$$

by the tower property of conditional expectation. Since ε_t is white noise, we have that

$$var\left[\varepsilon_t\right]=E\left[\varepsilon_t^2\right]=\sigma^2$$

for all t, where σ^2 is the long-term variance of ε_t. It follows directly then that

$$E\left[\varepsilon_t^2\right]=E\left[\varepsilon_{t-1}^2\right]$$

and

$$E\left[\sigma_t^2\right]=E\left[\sigma_{t-1}^2\right]$$

The expected value of the $GARCH\,(p,q)$ process

$$\sigma_t^2=\alpha_0+\sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2+\sum_{i=1}^{p}\beta_i\sigma_{t-i}^2$$

is

$$\begin{aligned} \sigma^2 &= E\left[\sigma_t^2\right] \\ &= \alpha_0+\sum_{i=1}^{q}\alpha_i E\left[\varepsilon_{t-i}^2\right]+\sum_{i=1}^{p}\beta_i E\left[\sigma_{t-i}^2\right] \\ &= \alpha_0+\sum_{i=1}^{q}\alpha_i\sigma^2+\sum_{i=1}^{p}\beta_i\sigma^2 \end{aligned}$$

It follows that

$$\sigma^2\left(1-\sum_{i=1}^{q}\alpha_i-\sum_{i=1}^{p}\beta_i\right)=\alpha_0$$

or

$$\sigma^2=\frac{\alpha_0}{\left(1-\sum_{i=1}^{q}\alpha_i-\sum_{i=1}^{p}\beta_i\right)}$$

For σ^2 to be finite it's required that

$$\sum_{i=1}^{q}\alpha_i+\sum_{i=1}^{p}\beta_i<1$$

■

4.4.2 Stylized Facts

In the financial literature four properties of returns series have been coined, stylized facts. These stylized facts are volatility clustering, mean reversion, excess kurtosis and the leverage effect. The leverage effect is discussed in section 4.7.

A stationary GARCH process captures these stylized facts in the following ways:

1. Volatility clustering is described in section 2.7.3 as strong autocorrelation of squared returns. Thus if σ_{t-1}^2 is high (low), then σ_t^2 will probably also be high (low). The long-term variance of a $GARCH\,(p,q)$ process was provided in theorem 4.4.1. The long-term variance of a $GARCH\,(1,1)$ process is

$$E\left[\sigma_t^2\right] = \frac{\alpha_0}{1-\alpha-\beta} \equiv V \tag{4.5}$$

thus

$$\alpha_0 \equiv V\left(1-\alpha-\beta\right)$$

and

$$\sigma_t^2 = V\left(1-\alpha-\beta\right) + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

equivalently,

$$\sigma_t^2 - V = \alpha\left(\varepsilon_{t-1}^2 - V\right) + \beta\left(\sigma_{t-1}^2 - V\right)$$

Taking expected value yields

$$\begin{aligned} E\left[\sigma_t^2 - V \mid \mathcal{F}_{t-2}\right] &= E\left[\alpha\left(\varepsilon_{t-1}^2 - V\right) + \beta\left(\sigma_{t-1}^2 - V\right) \mid \mathcal{F}_{t-2}\right] \\ &= \alpha E\left[\varepsilon_{t-1}^2 - V \mid \mathcal{F}_{t-2}\right] + \beta\left(\sigma_{t-1}^2 - V\right) \\ &= \left(\alpha+\beta\right)\left(\sigma_{t-1}^2 - V\right) \end{aligned} \tag{4.6}$$

since $E\left[\varepsilon_{t-1} \mid \mathcal{F}_{t-2}\right] = 0$ and $Var\left[\varepsilon_{t-1} \mid \mathcal{F}_{t-2}\right] = \sigma_{t-1}^2$. This equation can be rewritten as

$$E\left[\sigma_t^2 \mid \mathcal{F}_{t-2}\right] = V + \left(\alpha+\beta\right)\left(\sigma_{t-1}^2 - V\right)$$

thus if σ_{t-1}^2 is large (small) then it's expected for σ_t^2 also to be large (small).

2. Mean reversion is the gradual return of variance levels, after a shock, to a long-term variance level. Equation 4.6 can be rewritten as

$$E\left[\sigma_{t+k}^2 - V \mid \mathcal{F}_t\right] = \left(\alpha+\beta\right) E\left[\sigma_{t+k-1}^2 - V \mid \mathcal{F}_t\right]$$

By repeating this relationship yields

$$E\left[\sigma_{t+k}^2 - V \mid \mathcal{F}_t\right] = (\alpha+\beta)^k \left(\sigma_t^2 - V\right)$$

or

$$E\left[\sigma_{t+k}^2 \mid \mathcal{F}_t\right] = V + (\alpha+\beta)^k \left(\sigma_t^2 - V\right) \tag{4.7}$$

Since the GARCH process is stationary, $\alpha + \beta < 1$. This means that the second term of equation 4.7 tends to zero, as k tends to infinity. Thus the expected value of the conditional variance tends to the long-term variance level, V.

3. Excess kurtosis in returns series can be described as kurtosis, see section 2.5.2, larger than that of the normal distribution. In theorem 4.4.1 above, we proved that for the $GARCH\,(1,1)$ process

$$\begin{aligned} E\left[\varepsilon_t^2\right] &= E\left[\sigma_t^2\right] \\ &= \frac{\alpha_0}{1-\alpha-\beta} \end{aligned}$$

Bollerslev, see [6], proved that if $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$ the stationary fourth moment of ε exists,

$$E\left[\varepsilon_t^4\right] = \frac{3\alpha_0^2\,(1+\alpha+\beta)}{(1-\alpha-\beta)\left(1-\beta^2-2\alpha\beta-3\alpha^2\right)}$$

The stationary kurtosis is

$$K = \frac{E\left[\varepsilon_t^4\right]}{E\left[\varepsilon_t^2\right]^2} = \frac{3\left(1-(\alpha+\beta)^2\right)}{1-\beta^2-2\alpha\beta-3\alpha^2} > 3$$

thus the GARCH process is heavy-tailed (leptokurtic).

4.4.3 Estimation of GARCH Regression Model

This section focusses on the maximum likelihood estimation (MLE) of the GARCH regression model. The GARCH model in equation 4.4 may be written in terms of the following nonlinear regression model

$$\varepsilon_t = y_t - \mathbf{x}_t\mathbf{b}$$

which is the means process of the error ε_t, which is conditionally normal

$$\varepsilon_t \mid \mathcal{F}_{t-1} \sim N\left(0, \sigma_t^2\right)$$

where

$$\sigma_t^2 = \mathbf{z}_t'\boldsymbol{\omega}$$

is the GARCH(p,q) process. The vector

$$\mathbf{z}_t' = \left(1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, \sigma_{t-1}^2, \ldots, \sigma_{t-p}^2\right)$$

and parameter vector

$$\boldsymbol{\omega}' = \left(\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p\right)$$

Define Θ as a compact subspace of a Euclidean space, with $\boldsymbol{\theta} = (\mathbf{b}', \boldsymbol{\omega}') \in \Theta$. Denote the true parameter values of by θ_0, where $\theta_0 \in int\,\Theta$.

The likelihood function of ε_t is the pdf of the error process ε_t, written in terms of its parameters

$$f^*\left(0, \sigma_t^2\right) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{-[\varepsilon_t/\sigma_t]^2/2} \tag{4.8}$$

since the conditional mean is zero and the process follows GARCH variance. There are T observations.

It's computationally easier to take the ln of equation 4.8. The loglikelihood function is

$$f\left(0, \sigma_t^2\right) = \sum_{t=1}^{T} -\frac{1}{2}\ln\sigma_t^2 - \frac{1}{2}\frac{\varepsilon_t^2}{\sigma_t^2} + \text{constants} \tag{4.9}$$

The constants will have no effect on later results, thus redefine

$$\begin{aligned} f\left(0, \sigma_t^2\right) &= \sum_{t=1}^{T} -\frac{1}{2}\ln\sigma_t^2 - \frac{1}{2}\frac{\varepsilon_t^2}{\sigma_t^2} \\ &= \sum_{t=1}^{T} l_t(\boldsymbol{\theta}) \end{aligned} \tag{4.10}$$

where $l_t(\boldsymbol{\theta})$ is the likelihood function of observation t.

Differentiating $l_t(\boldsymbol{\theta})$ with respect to the variance parameters yields

$$\begin{aligned} \frac{\partial l_t}{\partial \omega} &= -\frac{1}{2}\sigma_t^{-2}\frac{\partial\sigma_t^2}{\partial\omega} + \frac{1}{2}\varepsilon_t^2\left(\sigma_t^2\right)^{-2}\frac{\partial\sigma_t^2}{\partial\omega} \\ &= \frac{1}{2\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\omega}\left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right) \end{aligned}$$

the second derivative

$$\frac{\partial l_t}{\partial\omega\partial\omega'} = \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)\frac{\partial\sigma_t^2}{\partial\omega}\left[\frac{1}{2\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\omega}\right] - \frac{1}{2\left(\sigma_t^2\right)^2}\frac{\partial\sigma_t^2}{\partial\omega}\frac{\partial\sigma_t^2}{\partial\omega'}\frac{\varepsilon_t^2}{\sigma_t^2}$$

where

$$\frac{\partial\sigma_t^2}{\partial\omega} = z_t + \sum_{i=1}^{p}\beta_i\frac{\partial h_{t-i}}{\partial\omega}$$

Differentiating $l_t(\boldsymbol{\theta})$ with respect to the mean parameters yields

$$\frac{\partial l_t}{\partial b} = \frac{\varepsilon_t x_t}{\sigma_t^2} + \frac{1}{2\left(\sigma_t^2\right)^2}\frac{\partial\sigma_t^2}{\partial b}\left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)$$

the second derivative

$$\begin{aligned}\frac{\partial l_t}{\partial b\partial b'} &= -\frac{1}{\sigma_t^2}x_t x_t' - \frac{1}{2}\frac{1}{\left(\sigma_t^2\right)^2}\frac{\partial\sigma_t^2}{\partial b}\frac{\partial\sigma_t^2}{\partial b'}\left(\frac{\varepsilon_t^2}{\sigma_t^2}\right)\\ &\quad -2\frac{1}{\left(\sigma_t^2\right)^2}\varepsilon_t x_t\frac{\partial\sigma_t^2}{\partial b} + \left(\frac{\varepsilon_t^2}{\sigma_t^2} - 1\right)\frac{\partial}{\partial b'}\left[\frac{1}{2\sigma_t^2}\frac{\partial\sigma_t^2}{\partial b}\right]\end{aligned}$$

where

$$\frac{\partial\sigma_t^2}{\partial b} = -2\sum_{j=1}^{q}\alpha_j x_{t-j}\varepsilon_{t-j} + \sum_{j=1}^{q}\beta_j\frac{\partial\sigma_{t-j}^2}{\partial b}$$

4.5 Integrated GARCH

The Integrated GARCH or I-GARCH process is defined as the standard $GARCH(p,q)$ process defined in equation 4.4 where $\alpha_1 + \beta_1 = 1$, thus if we put $\beta_1 = \lambda$ then

$$\sigma_t^2 = \alpha_0 + (1-\lambda)\varepsilon_{t-1}^2 + \lambda\sigma_{t-1}^2$$

where $\varepsilon_t \mid \mathcal{F}_{t-1} \sim N\left(0,\sigma_t^2\right)$ and clearly $\lambda \in [0,1]$.

From the stationary variance of the $GARCH(1,1)$ process defined in equation 4.5, it's clear that the stationary variance of the I-GARCH process doesn't exists. I-GARCH processes are often encountered in foreign exchange and commodity markets.

When the constant term $\alpha_0 = 0$ then the I-GARCH process is an EWMA process.

The I-GARCH process can however by strictly stationary, this result follows from Nelson (see [18]). For the $GARCH(1,1)$ process

$$\begin{aligned}\sigma_t^2 &= \alpha_0 + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2\\ &= \alpha_0 + \alpha\varepsilon_{t-1}^*\sigma_{t-1}^2 + \beta\sigma_{t-1}^2\end{aligned}$$

where $\varepsilon_t^* \mid \mathcal{F}_{t-1} \sim N(0,1)$. Further

$$\begin{aligned}
\sigma_t^2 &= \alpha_0 + \left(\alpha\varepsilon_{t-1}^* + \beta\right)\sigma_{t-1}^2 \\
&= \alpha_0 + \left(\alpha\varepsilon_{t-1}^* + \beta\right)\left(\alpha_0 + \left(\alpha\varepsilon_{t-2}^* + \beta\right)\sigma_{t-2}^2\right) \\
&= \alpha_0\left(1 + \left(\alpha\varepsilon_{t-1}^* + \beta\right)\right) + \left(\alpha\varepsilon_{t-1}^* + \beta\right)\left(\alpha\varepsilon_{t-2}^* + \beta\right)\sigma_{t-2}^2 \\
&\vdots \\
&= \alpha_0 + \left(1 + \sum_{i=1}^{t-1}\prod_{j=1}^{i}\left(\alpha\varepsilon_{t-j}^* + \beta\right)\right) + \prod_{j=1}^{i}\left(\alpha\varepsilon_{t-i}^* + \beta\right)\sigma_0^2
\end{aligned}$$

where σ_0^2 is the first conditional variance. Nelson proved that the process is strictly stationary if

$$E\left[\ln\left(\alpha\varepsilon_{t-i}^* + \beta\right)\right] < 1$$

for every applicable i.

4.6 GARCH-in-Mean

The ARCH-in-Mean (GARCH-M) process was introduced by Engle, Lilien & Robins in 1987. In this process the connection between returns and risk, represented by AR and GARCH processes respectively, is set. Risk averse investors are expected to demand higher returns on risky assets than on less risky ones. The GARCH process in this model is therefore fixed to a risk premium. This risk premium can be seen as the positive correlation between current return and conditional covariance.

An example of an GARCH-M process is

$$y_t = \phi_0 + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + g\left(\sigma_t, \lambda\right) + \varepsilon_t \tag{4.11}$$

where the ϕ-parameters are AR parameters and g is a function of a GARCH process, σ_t and the risk premia, λ. The function is mostly taken as the identity or square root function of σ_t multiplied with λ.

The GARCH-M process by Duan, discussed in chapter 6, is

$$S_t = S_{t-1}\exp\left(r\Delta t - \frac{1}{2}\sigma_t^2 + \lambda\sigma_t + \varepsilon_t\right) \tag{4.12}$$

or

$$\ln\frac{S_t}{S_{t-1}} = r\Delta t - \frac{1}{2}\sigma_t^2 + \lambda\sigma_t + \varepsilon_t$$

where, for an annual risk-free rate r and daily volatility measurements t, $\Delta t = 1/252$, since we assume 252 trading days in a year.

GARCH-M process can be extended by any other GARCH process.

4.7 Asymmetric GARCH and the Leverage Effect

The leverage effect was reviewed in section 2.7.4. The jest of the leverage effect is: markets tend to react more volatile to negative information than to positive information. Symmetric GARCH processes react equally to positive and negative news.

Asymmetric GARCH processes have an extra parameter, denoted by γ in this dissertation, that skew returns information to market reaction. Here follow a few Asymmetric GARCH processes:

4.7.1 Exponential GARCH

The Exponential GARCH (EGARCH) was introduced by Nelson (1991). The EGARCH process is given by

$$\ln \sigma_t^2 = \alpha_0 + \beta_1 \ln \sigma_{t-1}^2 + \beta_2 \left(|\varepsilon_{t-1}| - \gamma \varepsilon_{t-1} \right)$$

where $\beta_2, \gamma > 0$.

The upside of EGARCH is that it generally fits empirical financial data well, but the downside is that EGARCH has no analytic form for its term structure.

4.7.2 Asymmetric GARCH

The Asymmetric GARCH (AGARCH) process is by Engle and Ng (1993). The AGARCH process is as follows

$$\sigma_t^2 = \alpha_0 + \alpha \left(\varepsilon_{t-1} - \gamma \right)^2 + \beta \sigma_{t-1}^2$$

where $\alpha_0 > 0$ and $\alpha, \beta, \gamma \geq 0$.

The parameters of the AGARCH process is easier to estimate than that of the EGARCH process, and it possesses an analytical term structure.

4.7.3 Glosten, Jagannathan and Runkle GARCH

The Glosten, Jagannathan and Runkle GARCH (GJR) process (1993), is named after its founders. The process is

$$\sigma_t^2 = \alpha_0 + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2 + \gamma \max \left(-\varepsilon_t, 0 \right)^2$$

where $\gamma > 0$.

4.8 Limitations of the GARCH Process

The GARCH processes have the following limitations:

1. The GARCH processes perform best under stable market conditions. This process often fails to capture highly unexpected shocks, like market crashes. Except for the direct effect of a sudden shock, it may also cause structural changes in the market.

2. It's often hard to decide which GARCH process fits empirical data the best. There is no single GARCH process that can adequately model all conditional volatility processes. The conditional volatility structure of underlying assets also occasionally changes, which necessitates the using a different process.

3. The GARCH processes presented here depends on normal innovations. These processes often fail to fully capture the heavy tails observed in return series. Student's t-distribution and distributions like the Normal Inverse Gaussian distribution are often used as sources of innovation.

4. Investment decisions mustn't be solely based on the results of the GARCH processes. Other sources of information and models must also be used to make such decisions.

Part II

Risk-Neutral Valuation

Chapter 5

Risk-Neutral Valuation

5.1 Objectives[1]

The aim of this chapter is to provide essential background to continuous-time finance concepts and the standard risk-neutral valuation framework, which is the cornerstone of the Black-Scholes option pricing framework. The Black-Scholes framework is the benchmark pricing method for options. In this framework we assume constant volatility of stock returns which leads to the helpful property of a complete market model.

Empirical evidence shows that the constant volatility assumption is generally incorrect. The GARCH option pricing model discussed in chapters 6 and 7 is an attempt to include stochastic volatility into the option pricing framework, the price is that the market model is no longer complete. Although volatility is generally stochastic, it is important to know the risk-neutral valuation framework, since it is so widely used and because many of the concepts are used in incomplete market models.

In this chapter only the bare skeleton of the risk-neutral valuation framework is given. For more complete discussions see [25], [4], [32] or any of the many other similar books.

An introduction to continuous time stochastic calculus is given in section 5.2. The essential definitions of Brownian motion, martingales and Ito processes are given. The proofs of the Ito formula, absolute continuous measures and equivalent measures, the Radon-Nikodym theorem and Girsanov's theorem are excluded.

Continuous-time finance concepts are briefly discussed in section 5.3.

Section 5.4 is the core section of this chapter. The risk-neutral valuation framework is discussed under the assumption of constant volatility. Only the proofs vital for a better understanding of the model investigated in chapters 6 and 7 are proved. Special attention is paid to the concept of the market price of risk.

[1]Suggested reading: [4], [13], [17], [26] and [32].

5.2 Essentials of Continuous-time Stochastic Calculus

5.2.1 Brownian Motion

Definition 5.2.1 *Brownian motion, W_t, is a real-valued stochastic process satisfying the following conditions:*

1. *Continuous sample paths:* $t \to W_t \quad P$ *a.s..*
2. *Stationary increments:* $W_{t+s} - W_t$ *has the same probability law for any* $t \in \mathbb{R}^+$ *varying and* $s \in \mathbb{R}^+$ *fixed.*
3. *Independent increments:* $W_{t+s} - W_t$ *is independent of*

$$\mathcal{F}_t = \sigma(W_u, u \leq t)$$

4. $W_0 = 0 \quad P$ *a.s.*

The probability law mentioned in point 2, will throughout this dissertation be the Normal distribution with mean zero and variance s.

5.2.2 Martingales

Definition 5.2.2 *In discrete time: An adapted process, $(M_t)_{t \in I}$, where I is a countable index and $E|M_t| < \infty$, is called:*

1. *A martingale if*

$$E(M_t \,|\, \mathcal{F}_s) = M_s \; Pa.s.$$

for all $s, t \in I$, $s \leq t$.

2. *A super-martingale if*

$$E(M_t \,|\, \mathcal{F}_s) \leq M_s \quad P \text{ a.s.}$$

for all $s, t \in I$, $s \leq t$.

Definition 5.2.3 *In continuous time: An adapted process, $(M_t)_{t \in \mathbf{R}^+}$, where $\mathbb{R}^+$ is the positive real numbers and, $E|M_t| < \infty$ is called:*

1. *A martingale if*

$$E[M_t \,|\, \mathcal{F}_s] = M_s \quad P \text{ a.s.}$$

for all $s, t \in I$, $s \leq t$.

2. *A super-martingale if*

$$E[M_t \,|\, \mathcal{F}_s] \leq M_s \quad P \text{ a.s.}$$

for all $s, t \in I$, $s \leq t$.

5.2.3 Ito Process

Definition 5.2.4 *A stochastic process, X_t, is called an Ito process if it has a.s. continuous paths and*

$$X_t = X_0 + \int_0^T A(t,\omega)dt + \int_0^T B(t,\omega)dW_t \tag{5.1}$$

where $A(t,\omega)$ and $B(t,\omega)$ are $\mathcal{F}_t$ measurable,

$$\int_0^T |A(t,\omega)|\, dt < \infty \quad P \text{ a.s.}$$

and

$$E\left[\int_0^T B(t,\omega)^2 dt\right] < \infty \quad P \text{ a.s.}$$

X_t is also called the stock price process. In short hand notation

$$dX_t = A(t,\omega)dt + B(t,\omega)dW_t$$

Definition 5.2.5 *A stochastic process, S_t, follows a geometric Brownian motion if*

$$dS_t = S_t\mu(t,\omega)dt + S_t\sigma(t,\omega)dW_t$$

5.2.4 Ito Formula (in 1-Dimension)

Definition 5.2.6 *Let X_t be an Ito process as defined in equation (5.1). For the function*

$$f(t,x) \in C^2([0,\infty) \times \mathbb{R})$$

the Ito formula is given by

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 \tag{5.2}$$

$$= (\frac{\partial f}{\partial t} + A\frac{\partial f}{\partial x} + \frac{1}{2}B^2\frac{\partial^2 f}{\partial x^2})dt + B\frac{\partial f}{\partial x}dW_t \tag{5.3}$$

In integral notation this is:

$$f_t = f_0 + \int_0^t (\frac{\partial f}{\partial t} + A\frac{\partial f}{\partial x} + \frac{1}{2}B^2\frac{\partial^2 f}{\partial x^2})dt + \int_0^t B\frac{\partial f}{\partial x}dW_t \tag{5.4}$$

5.2.5 Absolute Continuous

Definition 5.2.7 *In our probability space $(\Omega, \mathcal{F}, P)$, probability measure P_1 is said to be absolutely continuous with respect to P if*

$$P(A) = 0 \Rightarrow P_1(A) = 0$$

for all $A \in \mathcal{F}$. This is sometimes denoted by

$$P_1 << P$$

Theorem 5.2.8 *Probability measure P_1 is absolutely continuous with respect to P if and only if there exists an adapted random variable K such that*

$$P_1(A) = \int_A K(\omega)\, dP \tag{5.5}$$

Proof. *See Lamberton and Lapeyre [26].* ∎

Definition 5.2.9 *The state price density is defined as*

$$\frac{dP_1}{dP}$$

thus from integral (5.5)

$$\frac{dP_1}{dP} = K$$

Definition 5.2.10 *In the probability space $(\Omega, \mathcal{F})$ two probability measures P_1 and P_2 are equivalent if*

$$P_1(A) = 0 \Leftrightarrow P_2(A) = 0$$

for all $A \in \mathcal{F}$.(See Lamberton and Lapeyre [26])

5.2.6 Radon-Nikodym

Theorem 5.2.11 *Let measure Q be absolutely continuous with respect to measure P. There then exists a random variable $\Lambda \geq 0$, such that*

$$E^P[\Lambda] = 1$$

and

$$Q(A) = \int_A dQ = \int_A \Lambda dP \tag{5.6}$$

for all $A \in \mathcal{F}$. Λ is $P-a.s.$ unique. Conversely, if there exists a random variable, Λ with the mentioned properties and Q is defined by equation 5.6, then Q is a probability measure and Q is absolutely continuous with respect to P.

Proof. See [25]. ∎

5.2.7 Risk-neutral Probability Measure

Definition 5.2.12 *A probability measure, Q, is called a risk-neutral probability measure if*

1. *Q is equivalent to the "real world" measure P.*
2. *$\frac{S_t}{B_t} = E^Q\left(\frac{S_{t+\tau}}{B_{t+\tau}}|\mathcal{F}_t\right)$ for all $t, \tau \in \mathbb{R}^+$.*

In this definition, B_t is the deterministic price process of a risk-free asset, where

$$B_t = B_0 \exp\left(\int_0^t r(s)ds\right)$$

The variable $r(t)$ is the short rate.

5.2.8 Girsanov's Theorem in One Dimension

Girsanov's theorem is used to transform stochastic processes in terms of their drift parameters. In option pricing, Girsanov's theorem is used to find a probability measure under which the risk-free rate adjusted stock price process is a martingale.

Definition 5.2.13 *A function $f(s,t) \in v(s,t)$ if*

$$f(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}$$

and the following holds:

1. *$(t,\omega) \to f(t,\omega)$ is $B \times \mathcal{F}$-measurable, where B is the Borel sets on $[0,\infty)$*
2. *$f(t,\omega)$ is adapted*
3. *$E\left[\int_S^T f(t,\omega)^2\, dt\right] < \infty$*

Theorem 5.2.14 *Girsanov's theorem. Let $X_t \in \mathbb{R}$ be an Ito process, of the form*

$$dX_t = \beta(t,\omega) + \theta(t,\omega)dW_t$$

with $t \leq T < \infty$. Suppose that there exist a $v(t,\omega)$-process $u(t,\omega) \in \mathbb{R}$ and $\alpha(t,\omega) \in \mathbb{R}$ such that

$$\theta(t,\omega)u(t,\omega) = \beta(t,\omega) - \alpha(t,\omega)$$

Since we are only looking at the one dimensional case

$$u(t,\omega) = \frac{(\beta(t,\omega) - \alpha(t,\omega))}{\theta(t,\omega)}$$

We further assume that

$$E^P\left[\exp\left(\frac{1}{2}\int_0^T u^2(s,\omega)\,ds\right)\right] < \infty \tag{5.7}$$

Let

$$M_t = \exp\left(-\int_0^t u(s,\omega)\,dW_t - \int_0^t u^2(s,\omega)\,ds\right) \tag{5.8}$$

and

$$dQ = M_T dP \tag{5.9}$$

We then have that

$$\tilde{W}_t = W_t + \int_0^t u(s,\omega)\,ds$$

is a Brownian motion with respect to Q. X_t in terms of $\tilde{W}_t$ is

$$dX_t = \alpha(t,\omega) + \theta(t,\omega)d\tilde{W}_t$$

M_t is a martingale.

Proof. *See Girsanov theorem II, Oksendal [27].* ■

Remark 5.2.15 *Result 5.9 is equivalent to*

$$E^Q[B] = E^P[BM_t]$$

for all Borel measurable sets B on $C[0,T]$.

5.3 Continuous-time Finance Essentials

This section contains a short summary of vital continuous-time finance concepts. For complete discussions on continuous-time finance see Bjork [4], Lamberton and Lapeyre [26] and Steele [32].

5.3.1 Self-financing

Definition 5.3.1 *A trading strategy is called self-financing if the value of the portfolio is due to the initial investment and gains and losses realized on the subsequent investments. This means that no funds are added or withdrawn from the portfolio.*

Theorem 5.3.2 *Let $\phi = \left(H_t^0, H_t\right)_{0\leq t\leq T}$ be an adapted process of portfolio weights satisfying*

$$\int_0^T \left|H_t^0\right| dt + \int_0^T H_t^2 dt < \infty \quad P\ a.s.$$

Then the discounted value of portfolio $V_t(\phi) = H_t^0\beta_t + H_t S_t$ namely, $\tilde{V}_t(\phi) = V_t(\phi)/\beta$ can be expressed for all $t \in [0,T]$ as

$$\tilde{V}_t(\phi) = V_0(\phi) + \int_0^t H_u d\tilde{S}_u \quad Q\ a.s.$$

if and only if ϕ is a self-financing strategy.

Proof. The product of $V_t(\phi)$ and with the bond process β yields

$$\begin{aligned}\frac{1}{\beta_t}V_t(\phi) &= V_0(\phi) + \int_0^t \frac{1}{\beta_t} dV_t(\phi) + \int_0^t V_s(\phi)\, d\frac{1}{\beta_t} + \left\langle V_t(\phi), \frac{1}{\beta_t}\right\rangle \\ &= V_0(\phi) + \int_0^t \frac{1}{\beta_t} dV_t(\phi) + \int_0^t V_s(\phi)\, d\frac{1}{\beta_t}\end{aligned}$$

since the process $\frac{1}{\beta_t}$ doesn't have a stochastic term. Since we can express $V_t(\phi)$ as

$$V_t(\phi) = H_t^0\beta_t + H_t S_t$$

a change in $V_t(\phi)$ can be expressed by

$$dV_t(\phi) = H_t^0 d\beta_t + H_t dS_t$$

thus

$$\begin{aligned}&\frac{1}{\beta_t}V_t(\phi) \\ &= V_0(\phi) + \int_0^t \frac{1}{\beta_t}\left(H_t^0 d\beta_t + H_t dS_t\right) + \int_0^t \left(H_t^0\beta_t + H_t S_t\right) d\frac{1}{\beta_t} \\ &= V_0(\phi) + H_t^0\left(\int_0^t \frac{1}{\beta_t} d\beta_t + \beta_t d\frac{1}{\beta_t}\right) + H_t\left(\int_0^t \frac{1}{\beta_t} H_t dS_t + H_t S_t d\frac{1}{\beta_t}\right) \\ &= V_0(\phi) + H_t^0 d\frac{\beta_t}{\beta_t} + H_t d\frac{S_t}{\beta_t} \\ &= V_0(\phi) + H_t d\frac{S_t}{\beta_t}\end{aligned}$$

■

5.3.2 Admissible Trading Strategy

Definition 5.3.3 *A trading strategy is admissible if it is self-financing and if the corresponding discounted portfolio,* $\tilde{V}_t$ *is nonnegative and* $\sup_{t\in[0,T]} \tilde{V}_t$ *is square integrable under the risk-neutral probability measure* Q.

5.3.3 Attainable Claim

Definition 5.3.4 *A claim is attainable if there exists an admissible trading strategy replicating that claim.*

5.3.4 Arbitrage Opportunity

Definition 5.3.5 *An arbitrage opportunity is an admissible trading strategy, such that the value of the portfolio at initialization,* $V(0) = 0$ *and* $E[V(T)] > 0$.

5.3.5 Complete Market

The completeness of a market can be defined in terms of the risk-neutral probability measure or in terms of the attainability of a contingent claim.

Definition 5.3.6 *Under no arbitrage conditions, the market model is complete if and only if every contingent claim is attainable.*

Theorem 5.3.7 *The market model is complete if and only if there exists a unique risk-neutral probability measure.*

Proof. See Pliska [28]. ■

5.4 Risk-Neutral Valuation under Constant Volatility

The aim of this section is to introduce the notion of risk-neutral valuation. The process of risk-neutral valuation is as follows:

1. In section 5.4.1 a simple stock price process is evaluated. A solution to this process is found and its distribution is discussed. The solution is obtained by applying the Ito process.

2. The next step, in section 5.4.2, is to evaluate the discounted stock price process. We get the discounted stock price process by discounting the solution to the original process in step 1 and then utilizing the Ito formula in reverse order.

3. This new process still has a trend. The so-called risk-neutral measure and related Brownian process is derived with Girsanov's theorem in section 5.4.3.

4. A wide-class of options are priced under risk-neutral valuation in section 5.4.4.

5.4.1 The Stock Price Process

It is generally assumed that stock prices follow geometric Brownian motion, under the real world measure P,

$$dS_t = S_t \mu dt + S_t \sigma dW_t \tag{5.10}$$

where $\mu \in \mathbb{R}$ and $S_0, \sigma \in \mathbb{R}^+$, W_t is Brownian motion and the process is defined on $[0, T]$.

A solution, S_t, to this equation can be found with the help of Ito's formula. Let $f(t,x) = \ln(x)$. It follows from section 5.2.4 that $f(t,x) \in C^2([0,\infty) \times \mathbb{R})$. Fortunately, if we assume that $S_t \in \mathbb{R}^+$, we can define $f(t,x) \in C^2([0,\infty) \times \mathbb{R}^+)$. From (5.4) we have[2]

$$\begin{aligned}
d\ln(S_t) &= \frac{1}{S_t} dS_t - \frac{1}{2}\frac{1}{S_t^2} dS_t^2 \\
&= \frac{1}{S_t}\left(S_t \mu dt + S_t \sigma dW_t\right) \\
&\quad -\frac{1}{2}\frac{1}{S_t^2}\left(S_t \mu dt + S_t \sigma dW_t\right)^2 \\
&= \mu dt + \sigma dW_t - \frac{1}{2}\sigma^2 dt \\
&= \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t
\end{aligned}$$

which in integral notation is

$$\begin{aligned}
\ln(S_t) &= \ln(S_0) + \int_0^t \left(\mu - \frac{1}{2}\sigma^2\right) du + \int_0^t \sigma dW_u \\
&= \ln(S_0) + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t
\end{aligned} \tag{5.11}$$

The solution, S_t, is

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right) \tag{5.12}$$

[2]In this chapter the drift μ, the variance σ and the risk-free interest rate r are all defined in terms of the same time period for instance 1 year.

Thus by assuming that the stock price follows the geometric Brownian motion described in equation 5.10, we are also assuming that the stock price process is lognormally distributed. There are ample empirical evidence to support this assumption. This means that from equation 5.11

$$\ln(S_t) \sim N\left(\ln(S_0) + \left(\mu - \frac{1}{2}\sigma^2\right)t, \sigma^2 t\right)$$

5.4.2 The Discounted Stock Price Process

The next aim is to find a probability measure under which $\tilde{S}_t = S_t/B_t$ is a martingale, called the risk-neutral probability measure. The discounted process

$$\tilde{S}_t = S_0 \exp\left(\left(\mu - r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right) \tag{5.13}$$

where $B_t = e^{rt}$ and r is the constant risk-free rate of interest.

To get the stochastic process driving $\tilde{S}_t = S_t e^{-rt}$, we again use Ito's formula

$$\begin{aligned}
df(t, S_t) &\equiv d\tilde{S}_t \\
&\equiv d\left(S_t e^{-rt}\right) \\
&= -rS_t e^{-rt}dt + e^{-rt}dS_t \\
&= -rS_t e^{-rt}dt + e^{-rt}\left(S_t\mu dt + S_t\sigma dW_t\right) \\
&= (\mu - r) S_t e^{-rt}dt + e^{-rt}S_t\sigma dW_t \\
&= (\mu - r)\tilde{S}_t dt + \tilde{S}_t\sigma dW_t
\end{aligned}$$

thus

$$d\tilde{S}_t = (\mu - r)\tilde{S}_t dt + \sigma\tilde{S}_t dW_t \tag{5.14}$$

In integral form this is

$$\tilde{S}_t = S_0 + \int_0^t \left(\mu - r - \frac{1}{2}\sigma^2\right)du + \int_0^t \sigma dW_u$$

5.4.3 Girsanov's Theorem Applied

It's clear that the process $\tilde{S}_t$ has a trend, $(\mu - r)\tilde{S}_t$. This trend causes $\tilde{S}_t$ not to be a P-martingale (a martingale under probability measure P).

The risk-neutral probability measure is found by employing Girsanov's theorem. By using the notation of the Girsanov theorem in section 5.2.8, we can define, for the process $\tilde{S}_t$,

$$\begin{aligned} u(t,\omega) &= \frac{(\mu - r)\tilde{S}_t}{\sigma\tilde{S}_t} \\ &= \frac{(\mu - r)}{\sigma}. \end{aligned}$$

Note that $\alpha(t,\omega) \equiv 0$ (in the sense of theorem 5.2.14) and $u(t,\omega) = u$ is a finite scalar since we assumed that σ is strictly positive. The result of this is that condition 5.7 is met and $u \in v(t,\omega)$.

M_t was defined in equation 5.8, as follows

$$M_t = \exp\left(-\int_0^t u(s,\omega)\,dW_t - \int_0^t u^2(s,\omega)\,ds\right)$$

In this case, for $u(t,\omega) = u$

$$M_t = \exp\left(-uW_t - u^2 t\right)$$

The new measure, the risk-neutral probability measure can be defined as

$$dQ = M_T dP$$

We can define a new process

$$\tilde{W}_t = ut + W_t$$

which is a Q- Brownian motion. The original process, $\tilde{S}_t$, in terms of $\tilde{W}_t$ is

$$d\tilde{S}_t = \sigma\tilde{S}_t d\tilde{W}_t \tag{5.15}$$

Remark 5.4.1 *The scalar $u(t,s) = \frac{(\mu - r)}{\sigma}$ is also known as the market price of risk. If $\mu = r$ then the investor is called risk-neutral and $dP = dQ$. Under the measure Q we price instruments as if they are risk-neutral.*

5.4.4 Pricing Options under Constant Volatility

Theorem 5.4.2 *The option price at time t defined by a nonnegative, $\mathcal{F}_t$-measurable random variable h such that*

$$E^Q\left[h^2\right] < \infty$$

is replicable and its value at time t is given by

$$V_t = e^{-r(T-t)} E^Q\left[h \mid \mathcal{F}_t\right] \tag{5.16}$$

Proof. Lets *assume* there exists an admissible trading strategy $\phi = (H_t^0, H_t)_{t\in[0,T]}$ *replicating* the option. The value of the replicating portfolio at time t is

$$V_t = H_t^0\beta_t + H_tS_t$$

The discounted value of the process at time t is

$$\begin{aligned}\tilde{V}_t &= e^{-rt}V_t \\ &= H_t^0 + H_t\tilde{S}_t\end{aligned}$$

Since no new funds are added or removed from the replicating portfolio, the portfolio is self-financing, by theorem 5.3.2 we can write the portfolio as

$$\tilde{V}_t(\phi) = V_0(\phi) + \int_0^t H_u d\tilde{S}_u$$

by equation 5.15 we can write

$$\tilde{V}_t(\phi) = V_0(\phi) + \int_0^t H_u\sigma\tilde{S}_u d\tilde{W}_u$$

By the assumption of an admissible trading strategy we have by theorem 5.3.3 proved that $\sup_{t\in[0,T]}\tilde{V}_t^2$ is square integrable. It can then be proven (see Lamberton and Lapeyre [26]) that if

$$E^Q\left[\sup_{t\in[0,T]}\tilde{V}_t^2\right] < \infty$$

then

$$E^Q\left[\int_0^t \left(H_u\sigma\tilde{S}_u\right)^2 du\right] < \infty \tag{5.17}$$

Further, there exists a unique continuous mapping from the class of adapted processes with property 5.17 to the space of continuous $\mathcal{F}_t$ martingales on $[0,T]$. We thus have that

$$\tilde{V}_t = E^Q\left[\tilde{V}_T \mid \mathcal{F}_t\right]$$

and hence

$$V_t = E^Q\left[e^{-r(T-t)}h \mid \mathcal{F}_t\right] \tag{5.18}$$

which is a square-integrable martingale.

We have assumed that there exists a portfolio replicating the option, an admissible trading strategy can easily be found by the use of the martingale representation theorem (see Lamberton and Lapeyre [26]). By the martingale representation theorem there exists a square integrable martingale under Q with respect to $\mathcal{F}_t$ such that for every $0 \leq t \leq T$,

$$M_t = E^Q\left[e^{-rT}h \mid \mathcal{F}_t\right]$$

and that any such martingale is a stochastic integral with respect to $\tilde{W}$, such that

$$E^Q\left[e^{-rT}h \mid \mathcal{F}_t\right] = M_0 + \int_0^t \eta_u d\tilde{W}_u$$

where η_t is adapted to $\mathcal{F}_t$ and

$$E^Q\left[\int_0^T (\eta_s)^2 \, ds\right] < \infty.$$

By letting $H_0 = M_t - H_t\tilde{S}_t$ and $H_t = \eta_t/\left(\sigma\tilde{S}_t\right)$ we have found a self-financing trading strategy. ■

5.4.5 The Black-Scholes Formula and Implied Volatility

The Black-Scholes formula for a European put option is a solution to equation 5.16 when

$$h = (X - S_T)_+$$

Black and Scholes (1973) and Merton (1973) proved that this as a solution to the Black-Scholes partial differential equation (pde). A martingale proof was later discovered. For the derivation of the pde proof for this formula see Black and Scholes [5], for a martingale proofs see Lamberton and Lapeyre [26] and Steele [32]. The Black-Scholes formula for a European put option at time t is

$$P^{BS} = e^{-r(T-t)}KN(-d_2) - S_tN(-d_1)$$

where

$$d_1 = \frac{\ln(S_0/X) + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T}$$

In this formula K is the strike price of the option and $N(\cdot)$ is the cumulative normal distribution. The risk-free interest rate r and the variance σ^2 are both annualized.

Volatility is the only parameter of the Black-Scholes formula that isn't directly observable. *Implied volatility*, σ, is the solution to the following problem

$$\min_{\sigma} \left| P^{BS}(\sigma) - P \right|$$

where $P^{BS}(\sigma)$ is the estimate of the put option as a function of implied volatility and P is the market value of the put option at time t.

Part III

Option pricing under the Local Risk-Neutral Valuation Relationship

Chapter 6

Local Risk-Neutral Valuation

6.1 Introduction

One of the properties of Brownian motion is that equally spaced increments are stationary, that is, it can be assumed that they are independently and identically distributed. The vast majority of empirical studies show that this is generally not the case.

Stochastic volatility in stock prices complicates the pricing of derivative instruments. The assumption of a complete market model and therefore the risk-neutral probability measure derived in chapter 5 no longer holds. This is because we cannot completely hedge away the risk posed by stochastic volatility.

Jin-Chuan Duan (1995) [10] defined a new measure, the *local risk-neutral probability measure*. He showed that an economic agent maximizes its expected utility by using this measure. In this incomplete market, extra assumptions are made about the consumer (its utility function) and the risk premium demanded by the market for taking additional risk. Duan named the properties of the measure, the *local risk-neutral valuation relationship* (LRNVR).

In this chapter the GARCH, EGARCH and GJR-GARCH processes are considered in the GARCH-M framework. The GARCH processes are in discrete time, thus unlike the risk-neutral pricing framework which forms the basis for Black-Scholes framework, the LRNVR is in discrete time.

In section 6.2, the continuous-time option pricing model discussed in chapter 5 is converted into a discrete time model. The goal of this section is to translate and compare some of the well-known continuous time finance concepts into discrete time statistical concepts. For example the continuous time concept of Brownian motion is converted in discrete time to that of expected returns.

The GARCH-in-Mean model for the volatility of a discrete time stock price process used by Duan for option pricing, is introduced in section 6.3.

Utility functions and the risk aversion of economic agents are discussed in section 6.4. A general consumption-investment strategy is maximized in section 6.5.

The LRNVR is defined in section 6.6 after which the local risk-neutral measure is derived in section 6.7. The stock price process under the new measure is discussed in section 6.8.

6.2 The Stock Price Process in Discrete Time

Recall the stock price process of section 5.13 with solution

$$\tilde{S}_t = S_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right]$$

In discrete time, with equally spaced observations,

$$\tilde{S}_{t-1} = S_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right) \times (t-1) + \sigma W_{t-1}\right]$$

which gives

$$S_0 = \tilde{S}_{t-1} \exp\left[-\left(\mu - \frac{1}{2}\sigma^2\right) \times (t-1) - \sigma W_{t-1}\right]$$

At time t, the value of

$$\begin{aligned} \tilde{S}_t &= S_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right] \\ &= \tilde{S}_{t-1} \exp\left[-\left(\mu - \frac{1}{2}\sigma^2\right)(t-1) - \sigma W_{t-1}\right] \\ &\quad \times \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right] \end{aligned}$$

Since

$$W_t - W_{t-1} \sim W_{t-t+1} \sim N(0,1)$$

let

$$\varepsilon_t | \mathcal{F}_{t-1} \overset{P}{\sim} N(0,1)$$

where ε_t is $\mathcal{F}_{t-1}$ measurable.

The one period ahead stock price is defined by

$$\tilde{S}_t = \tilde{S}_{t-1} \exp\left(\mu - \frac{1}{2}\sigma^2 + \sigma\varepsilon_t\right) \tag{6.1}$$

where

$$\varepsilon_t | \mathcal{F}_{t-1} \overset{P}{\sim} N(0,1)$$

If we let

$$\lambda = \frac{\mu - r}{\sigma}$$

where r is the risk-free rate of interest, equation 6.1 would become

$$\tilde{S}_t = \tilde{S}_{t-1} \exp\left(r - \frac{1}{2}\sigma^2 + \lambda\sigma + \sigma\varepsilon_t\right)$$

In the discrete case where the information on time $t-1$ is known, we could just as well have considered a volatility process which is constant between time $t-1$ and t.

6.3 The Stock Price Model under certain GARCH Volatility

Jin-Chuan Duan proposed the following conditional, lognormally distributed stock price process, with stochastic volatility, under the P measure

$$S_t = S_{t-1} \exp\left(r^*\Delta t - \frac{1}{2}\sigma_t^2 + \lambda\sigma_t + \sigma_t\varepsilon_t\right) \tag{6.2}$$

where

$$\varepsilon_t | \mathcal{F}_{t-1} \overset{P}{\sim} N(0,1)$$

is the conditional error process, σ_t^2 is the conditional variance (GARCH process) and λ[1] the unit risk premium. $\mathcal{F}_{t-1}$ is the σ-algebra of information up to time t. The yearly risk-free rate of return is r^*. Henceforth define r over period Δt, the same time period over which the conditional variance is taken. From this point on the period is daily.

[1] A possible interpretation of the unit risk premium follows from section 6.2 which deals with the market price of risk. Define the risk premium as

$$\lambda^* = \frac{\mu - r}{\sigma}$$

where σ is the long term or unconditional standard deviation of the series $\{X_t\}$.We can simplify the term with the risk premium from equation 6.3 to get

$$\lambda^*\sigma_t = (\mu - r)\frac{\sigma_t}{\sigma}.$$

$(\mu - r)$ can be seen as a fixed (positive) premium. $\lambda^*\sigma_t$ increases as the predicted conditional volatility σ_t increases over the long term volatility σ. The economic interpretation is that the market agent demands a higher premium as the expected volatility increases.

The conditional expected rate of return is defined as

$$\begin{aligned} \ln \frac{S_t}{S_{t-1}} &= r - \frac{1}{2}\sigma_t^2 + \lambda\sigma_t + \sigma_t\varepsilon_t \\ &\sim N\left(r - \frac{1}{2}\sigma_t^2 + \lambda\sigma_t, \sigma_t^2\right) \end{aligned} \tag{6.3}$$

This is derived by transforming equation 6.2.

The GARCH option pricing model prices options under conditional heteroscedasticity. This means that conditional variance is allowed to change over time while keeping unconditional variance constant. In this dissertation, options whose variance follows (vanilla) GARCH, GJR-GARCH and EGARCH process will be investigated. The main focus will be on the GARCH(p, q) process and specifically GARCH($1, 1$) process.

The GARCH(p, q) conditional variance process is

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i\sigma_{t-i}^2$$

where $\alpha_0 > 0$ and $\alpha_i, \beta_i \geq 0$ for all applicable numbers i. Notice that σ_t^2 is predictable at time $t-1$.

The GJR-GARCH variance process is

$$\sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + \alpha\varepsilon_{t-1}^2 + \gamma \max\left(-\varepsilon_t, 0\right)^2$$

where $\gamma > 0$.

The EGARCH variance process is

$$\ln\sigma_t^2 = \alpha_0 + \beta_1 \ln\sigma_{t-1}^2 + \beta_2\left(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1}\right)$$

where $\beta_2, \gamma > 0$.

6.4 Consumer Utility Essentials

6.4.1 Utility Functions

The satisfaction (utility) an economic agent gets from consumption can often not be described on a monetary scale. A utility function represents an economic agent's welfare from consumption.

In this dissertation we assume that utility is measurable and possible to represent in a function. This function is called a (cardinal) utility function. Define the utility function by

$$u(x) : \mathbb{R}^+ \to \mathbb{R}^+$$

that is

$$u\,(\text{monetary cost of consumption}) = \text{“welfare” from that consumption}$$

It is generally assumed that a utility function has the following three properties:

1. $u(x)$ is twice differentiable
2. $u(x)$ is an increasing function of x, $u'(x) > 0$
3. $u(x)$ is a concave function of x, $u''(x) < 0$.

Property 2 is due to the fact that an economic agent prefers to consume more than less. Property 3 can be interpreted in terms of additional consumption. The utility that an economic agent gains from additional consumption g, in excess of an existing consumption x,

$$u(x+g) - u(x) < u(g)$$

Each economic agent has its own unique preferences and thus utility function.

6.4.2 Risk Aversion

Absolute Risk Aversion

For a given utility function $u(x)$, in continuous time, we can associate an absolute risk aversion function defined by

$$R(x) = \frac{-u''(x)}{u'(x)} = -\frac{d}{dx}\ln u'(x) \tag{6.4a}$$

Properties 2 and 3 of section 6.4.1 insures that $R(x) > 0$ for all x. The bigger $R(x)$ is, the less risk the economic agent is willing to take for additional consumption. The discrete time version of equation 6.4a

$$\begin{aligned}\widetilde{R}(x) &= -\frac{\ln u'(x_t) - \ln u'(x_{t-1})}{x_t - x_{t-1}} \\ &= -\frac{\ln \frac{u'(x_t)}{u'(x_{t-1})}}{x_t - x_{t-1}}\end{aligned}$$

Relative Risk Aversion

The relative risk aversion for a utility function $u(x)$ is defined by

$$\begin{aligned}r(x) &= x\widetilde{R}(x) = -\frac{u''(x)}{u'(x)}x \\ &= -\frac{\frac{d}{dx}\ln u'(x)}{\frac{d}{dx}\ln x}\end{aligned}$$

The discrete time risk aversion function is

$$\begin{aligned}\widetilde{r}(x) &= -\frac{\ln u'(x_t) - \ln u'(x_{t-1})}{x_t - x_{t-1}} \div \frac{\ln x_t - \ln x_{t-1}}{x_t - x_{t-1}} \\ &= -\frac{\ln u'(x_t) - \ln u'(x_{t-1})}{\ln x_t - \ln x_{t-1}} \\ &= -\frac{\ln \frac{u'(x_t)}{u'(x_{t-1})}}{\ln \frac{x_t}{x_{t-1}}}\end{aligned}$$

6.5 A General Consumption-Investment Strategy

Consider an investor with the following discrete time consumption-investment plan: The investor maximizes its differentiable utility function, $u(x) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, at any point in time $t-1$ by either consuming, $C_{t-1} \in \mathbb{R}^+$ or by investing, $H_{t-1} \in \mathbb{R}$, in a portfolio with random payoff $S_t \in \mathbb{R}^+$ at time t. At time t the investor again has the same choice between consumption and investment from the payoff from time $t-1$. Like most investors, this investor gets more satisfaction out of consuming immediately than waiting for the next period, hence define the impatience factor $\rho \in \mathbb{R}$. At time $t-1$ this plan can be written as

$$\max E^P \left[u(C_{t-1}) + \exp(-\rho) u(C_t) \mid \mathcal{F}_{t-1}\right] \tag{6.5}$$

subject to

$$v = C_{t-1} + H_{t-1}S_{t-1} \tag{6.6}$$

$$H_{t-1}S_t = C_t + H_t S_t \tag{6.7}$$

v is the payoff of the investment made at time $t-2$. Take note that at time $t-1$ the only choices this investor make is to consume now or invest for one period ahead, thus the expected utility of consumption of periods after time t isn't of concern. Since C_{t-1} is predictable at time $t-1$ problem (6.5) can be rewritten as

$$\max u(C_{t-1}) + \exp(-\rho) E^P \left[u(C_t) \mid \mathcal{F}_{t-1}\right] \tag{6.8}$$

The aim here is to maximize utility in terms of consumption and investment. From equations (6.6) and (6.7) consumption in subject (6.8) can be rewritten in terms of investment as

$$\max\ u(v - H_{t-1}S_{t-1}) + \exp(-\rho) E^P \left[u(H_{t-1}S_t - H_t S_t) \mid \mathcal{F}_{t-1}\right] \tag{6.9}$$

If we then maximize problem (6.9) in terms of H_{t-1} we get

$$\begin{aligned} 0 &= \frac{\partial}{\partial H_{t-1}} u\left(v - H_{t-1}S_{t-1}\right) \\ &\quad + \frac{\partial}{\partial H_{t-1}} \exp\left(-\rho\right) E^P\left[u\left(H_{t-1}S_t - H_tS_t\right) \mid \mathcal{F}_{t-1}\right] \\ &= -S_{t-1}u'\left(v - H_{t-1}S_{t-1}\right) \\ &\quad + \exp\left(-\rho\right) E^P\left[S_t u'\left(H_{t-1}S_t - H_tS_t\right) \mid \mathcal{F}_{t-1}\right] \end{aligned}$$

which by further simplification and equations (6.6) and (6.7) yield

$$S_{t-1} = E^P\left[\exp\left(-\rho\right) \frac{u'\left(C_t\right)}{u'\left(C_{t-1}\right)} S_t \mid \mathcal{F}_{t-1}\right] \tag{6.10}$$

Thus the price of the portfolio at time $t-1$ is written in terms of the expected value of the economic agent's utility, its impatience and the expected future portfolio value.

6.6 The Local Risk-Neutral Valuation Relationship

The conventional risk-neutral valuation doesn't accommodate heteroscedasticity of stock returns. The Local Risk-neutral Valuation Relationship (LRNVR) is a way to generalize risk-neutral valuation to accommodate heteroscedasticity.

Definition 6.6.1 $(\Omega, \mathcal{F}, P)$. *A probability measure Q is said to be a local risk-neutral probability measure if*

1. *Q is equivalent to measure P*
2. $E^Q\left[\ln\left(\frac{S_t}{S_{t-1}}\right) | \mathcal{F}_{t-1}\right] = r$ *for all* $t \in \mathbb{R}^+$
3. $Var^Q\left[\ln\left(\frac{S_t}{S_{t-1}}\right) | \mathcal{F}_{t-1}\right] = Var^P\left[\ln\left(\frac{S_t}{S_{t-1}}\right) | \mathcal{F}_{t-1}\right] \qquad P$ *a.s.*

Remark 6.6.2 *Condition 1 is the same as in 5.2.7. Condition 2 is also similar but only defined over one period. The expected return doesn't locally depend on preferences. The one period conditional variance of the returns are invariant almost surely under the equivalent measures.*

The rest of this chapter focuses on the possible characteristics of an economic agent and the distribution of returns for which the LRNVR will hold.

6.7 The Local Risk-Neutral Probability Measure

Theorem 6.7.1 *Let a process Y_t be such that $Y_t \mid \mathcal{F}_{t-1} \overset{P}{\sim}$ is normally distributed with constant mean and variance under the P-measure. Define Q as*

$$dQ = \exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s\right) dP$$

then Q is a measure and is equivalent to P.

Proof. Measure. *Q is a measure by Corollary 4.9 of Bartle [3], since*

$$\exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s\right)$$

is a nonnegative $\mathcal{F}_{t-1}$ measurable function from Ω to $\mathbb{R}$.

Equivalence. *Let $A \in \mathcal{F}$ be a set such that*

$$P(A) = 0$$

Then

$$\begin{aligned} P(A) &= 0 \\ &\Leftrightarrow \int_A dP = \int_\Omega I_A dP = 0 \end{aligned}$$

where I_A is a characteristic function for A, I_A is a measurable and nonnegative function.

$$\int_\Omega I_A dP = 0 \Leftrightarrow I_A = 0 \quad P-a.s.$$

This means that

$$I(\omega) = 0$$

for all $\omega \in A \equiv \Omega \backslash M$ where $P(M) = 0$. This holds if and only if

$$I(\omega) f(\omega) = 0$$

for all $\omega \in A \equiv \Omega \backslash M$ where $P(M) = 0$. $f(\omega)$ is a measurable positive continuous function from Ω to $\mathbb{R}^+$. The product of real measurable functions $f_A I_A$ is also measurable. This is the same as

$$f I_A = 0 \quad P-a.s.$$

It is also equivalent to [3]

$$\int_\Omega f_A I_A dP = \int_A f dP$$

It is clear that

$$\tilde{f}(\omega) = \exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s(\omega)\right)$$

is a nonnegative, measurable continuous function from Ω *to* $\mathbb{R}^+$ *(* $Y_s : \Omega \to \mathbb{R}$ *for every* s*).*

$$\tilde{f}(\omega) = 0$$

if and only if $Y_s(\omega) = -\infty$ *for any* s*. Fortunately* $P\{Y_s = -\infty\} = 0$ *since* Y_s *is normal.*

Thus

$$\begin{aligned} P(A) &= 0 \\ &\Leftrightarrow \int_{\Omega} I_A dP = 0 \\ &\Leftrightarrow \int_{\Omega} \exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s(\omega)\right) I_A dP = 0 \\ &\Leftrightarrow Q(A) = 0 \end{aligned}$$

Thus the measure Q *is equivalent to measure* P*.* ■

The measure Q isn't, in general, a probability measure. In the next theorem conditions under which Q is a probability measure will be defined and a desirable property of Q will be derived.

Theorem 6.7.2 *If*

$$S_{t-1} = E^P\left[S_t \exp(-\rho + Y_t) \mid \mathcal{F}_{t-1}\right] \tag{6.11}$$

then

1. *Q is a probability measure*
2. *If W_t is $\mathcal{F}_t$ measurable then*

$$E^Q\left[W_t \mid \mathcal{F}_{t-1}\right] = E^P\left[W_t \exp((r-\rho) + Y_t) \mid \mathcal{F}_{t-1}\right]$$

Proof. *From the definition of* Q

$$dQ = \exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s\right) dP$$

In integral notation

$$\begin{aligned}
\int_{\Omega} dQ &= \int_{\Omega} \exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s\right) dP \\
&= E^P\left[\exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s\right)\right] \\
&= E^P\left[\exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s\right) \mid \mathcal{F}_0\right] \\
&= E^P\left[\exp\left((r-\rho)(T-1) + \sum_{s=1}^{T-1} Y_s\right) \exp\left(r-\rho+Y_T\right) \mid \mathcal{F}_0\right] \\
&= E^P\left[e^{\left((r-\rho)(T-1)+\sum_{s=1}^{T-1} Y_s\right)} e^r E^P\left[e^{\rho+Y_T} \mid \mathcal{F}_{T-1}\right] \mid \mathcal{F}_0\right]
\end{aligned}$$

This last step is due to the tower property of conditional expectation. Assumption 6.11 states that

$$E^P\left[\exp\left(-\rho+Y_t\right) S_t \mid \mathcal{F}_{t-1}\right] = S_{t-1}$$

thus

$$E^P\left[\exp\left(-\rho+Y_T\right) \mid \mathcal{F}_{T-1}\right] = \exp\left(-r\right) \tag{6.12}$$

for a risk-free asset. The result is that

$$\begin{aligned}
\int dQ &= E^P\left[\exp\left((r-\rho)(T-1) + \sum_{s=1}^{T-1} Y_s\right) \exp\left(r\right)\exp\left(-r\right) \mid \mathcal{F}_0\right] \\
&= E^P\left[\exp\left((r-\rho)(T-1) + \sum_{s=1}^{T-1} Y_s\right) \mid \mathcal{F}_0\right]
\end{aligned}$$

The tower property can again be invoked and an argument similar to 6.12 can be derived.

$$\begin{aligned}
\int_{\Omega} dQ &= E^P\left[e^{\left((r-\rho)(T-2)+\sum_{s=1}^{T-2} Y_s\right)} e^r E^P\left[e^{(\rho+Y_{T-1})} \mid \mathcal{F}_{T-2}\right] \mid \mathcal{F}_0\right] \\
&= E^P\left[\exp\left((r-\rho)(T-2) + \sum_{s=1}^{T-2} Y_s\right) \mid \mathcal{F}_0\right]
\end{aligned}$$

This can be repeated until we have, at filtration $\mathcal{F}_0$,

$$\int_{\Omega} dQ = E^P\left[\exp\left((r-\rho)+Y_1\right) \mid \mathcal{F}_0\right] = 1$$

Thus

$$Q\left(\Omega\right) = 1$$

with this property, the measure Q is a probability measure. We also have that

$$Q(\Omega) = E^P\left[\exp\left((r-\rho)+Y_1\right)\right] = 1$$

it is clear that

$$\exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s\right) \geq 0$$

and we proved in theorem 6.7.1 that Q is equivalent to P. The Radon-Nikodym theorem can be invoked thus

$$\exp\left((r-\rho)T + \sum_{s=1}^{T} Y_s\right)$$

is $P-a.s.$ unique and for any $\mathcal{F}_t$ measurable set W_t,

$$E^Q\left[W_t \mid \mathcal{F}_{t-1}\right] = E^P\left[W_t \exp\left((r-\rho)+Y_t\right) \mid \mathcal{F}_{t-1}\right]$$

■

Theorem 6.7.3 *If*

$$S_{t-1} = E^P\left[\exp\left(-\rho + Y_t\right) S_t \mid \mathcal{F}_{t-1}\right]$$

then

1. *$ln\left(\frac{S_t}{S_{t-1}}\right) \mid \mathcal{F}_{t-1} \overset{Q}{\sim}$ normal*
2. *$E^Q\left[\frac{S_t}{S_{t-1}} \mid \mathcal{F}_{t-1}\right] = e^r$ for all $t \in \mathbb{R}^+$*
3. *$Var^Q\left[\ln\left(\frac{S_t}{S_{t-1}}\right) | \mathcal{F}_{t-1}\right] = Var^P\left[\ln\left(\frac{S_t}{S_{t-1}}\right) | \mathcal{F}_{t-1}\right] \qquad P$ a.s.*

Proof. Lemma 2.
From theorem 6.7.2 we have

$$\begin{aligned} & E^Q\left[\frac{S_t}{S_{t-1}} \mid \mathcal{F}_{t-1}\right] \\ = \; & E^P\left[\frac{S_t}{S_{t-1}} \exp\left((r-\rho)+Y_t\right) \mid \mathcal{F}_{t-1}\right] \\ = \; & \frac{e^r}{S_{t-1}} E^P\left[S_t \exp\left(-\rho+Y_t\right) \mid \mathcal{F}_{t-1}\right] \\ = \; & e^r \end{aligned}$$

Proof of lemmas 1 and 3.

In theorem 6.7.2 we proved that

$$E^Q[W_t \mid \mathcal{F}_{t-1}] = E^P[W_t \exp((r-\rho)+Y_t) \mid \mathcal{F}_{t-1}]$$

for all $\mathcal{F}_t$ measurable sets W_t. If W_t is $\mathcal{F}_t-$measurable, so is W_t^c for all $c \in \mathbb{R}$. From theorem 6.7.2 we have that

$$E^Q[S_t^c \mid \mathcal{F}_{t-1}] = E^P\left[S_t^c e^{((r-\rho)+Y_t)} \mid \mathcal{F}_{t-1}\right]$$

then

$$\begin{aligned} E^Q\left[\frac{S_t^c}{S_{t-1}^c} \mid \mathcal{F}_{t-1}\right] &= E^P\left[\frac{S_t^c}{S_{t-1}^c} e^{((r-\rho)+Y_t)} \mid \mathcal{F}_{t-1}\right] \\ E^Q\left[e^{c\ln\frac{S_t}{S_{t-1}}} \mid \mathcal{F}_{t-1}\right] &= E^P\left[e^{c\ln\frac{S_t}{S_{t-1}}} e^{((r-\rho)+Y_t)} \mid \mathcal{F}_{t-1}\right] \\ E^Q\left[e^{cX_t} \mid \mathcal{F}_{t-1}\right] &= E^P\left[e^{cX_t} e^{((r-\rho)+Y_t)} \mid \mathcal{F}_{t-1}\right] \end{aligned}$$

if we define

$$X_t = \ln\frac{S_t}{S_{t-1}}.$$

Throughout this chapter there's been assumed that $X_t \mid \mathcal{F}_{t-1}$ is normally distributed under P, say

$$X_t \mid \mathcal{F}_{t-1} \overset{P}{\sim} N\left(\mu_t, v_t^2\right)$$

In theorem 6.7.1 we assumed that Y_t is also conditionally normal. Y_t can thus be written in terms of X_t, a constant α and another random variable with zero mean U_t, which is independent of X_t. Then

$$Y_t = \alpha + \beta X_t + U_t$$

with $\beta \in \mathbb{R}$. Thus

$$\begin{aligned} & E^Q\left[e^{cX_t} \mid \mathcal{F}_{t-1}\right] \\ = \ & E^P\left[e^{cX_t} e^{((r-\rho)+Y_t)} \mid \mathcal{F}_{t-1}\right] \\ = \ & E^P\left[e^{cX_t+\beta X_t+\alpha+U_t+(r-\rho)} \mid \mathcal{F}_{t-1}\right] \\ = \ & e^{\alpha+r-\rho} E^P\left[e^{(c+\beta)X_t+U_t} \mid \mathcal{F}_{t-1}\right] \end{aligned} \tag{6.13}$$

The joint variance of $(c+\beta)X_t$ and U_t under P is

$$var\left((c+\beta)X_t + U_t\right) = (c+\beta)^2 v_t^2 + E^P\left[U_t^2\right]$$

since U_t is of zero mean. By the moment generating function

$$\begin{aligned}
& E^P\left[e^{(c+\beta)X_t+U_t} \mid \mathcal{F}_{t-1}\right] \\
= \; & e^{\mu_t(c+\beta)+\frac{1}{2}\left((c+\beta)^2 v_t^2+E^P\left[U_t^2\right]\right)} \\
= \; & e^{\mu_t(c+\beta)+\frac{1}{2}\left(c^2+2c\beta+\beta^2\right)v_t^2+\frac{1}{2}E^P\left[U_t^2\right]} \\
= \; & e^{\frac{1}{2}\beta^2 v_t^2+\mu_t\beta+\frac{1}{2}E^P\left[U_t^2\right]+\frac{1}{2}c^2 v_t^2+c\left(\mu_t+\beta v_t^2\right)}
\end{aligned}$$

Then equation 6.13 becomes

$$\begin{aligned}
E^Q\left[e^{cX_t} \mid \mathcal{F}_{t-1}\right] &= e^{\left[(r-\rho)+\frac{1}{2}E^P\left[U_t^2 \mid \mathcal{F}_{t-1}\right]+\mu_t\beta+\beta^2 v_t^2\right]} \times \\
& \quad e^{\left[\frac{1}{2}c^2 v_t^2+c\left(\mu_t+\beta v_t^2\right)\right]}
\end{aligned}$$

This equation holds for all $c \in \mathbb{R}$. If we let $c = 0$ then

$$\begin{aligned}
1 &= E^Q\left[1 \mid \mathcal{F}_{t-1}\right] \\
&= e^{(r-\rho)+\frac{1}{2}E^P\left[U_t^2 \mid \mathcal{F}_{t-1}\right]+\mu_t\beta+\beta^2 v_t^2}
\end{aligned}$$

so we are left with

$$E^Q\left[e^{cX_t} \mid \mathcal{F}_{t-1}\right] = e^{\frac{1}{2}c^2 v_t^2+c\left(\mu_t+\beta v_t^2\right)}$$

If we let $c = 1$, then by the form of the answer of a moment generating function,

$$X_t = \ln\left(\frac{S_t}{S_{t-1}}\right) \mid \mathcal{F}_{t-1} \overset{Q}{\sim} N\left(\mu_t+\beta v_t^2, v_t^2\right)$$

Which proves 1. The conditional variance under P of X_t is also σ^2 thus lemma 3 is also proved. ■

Theorem 6.7.4 *An economic agent who's an expected utility maximizer and whose utility function is separable and additive is a LRNVR investor under the following conditions:*

1. *The utility function is of constant relative risk aversion and the changes in the logarithm of the aggregate consumption are conditionally normally distributed with constant mean and variance under the P measure*
2. *The utility function is of constant absolute risk aversion and the changes in the logarithm of the aggregate consumption are conditionally normally distributed with constant mean and variance under the P measure*
3. *The utility function is linear.*

The local risk-neutral measure is

$$dQ = e^{-(r-\rho)T} \frac{U'(C_t)}{U'(C_{t-1})} dP$$

The implied interest rate is assumed constant.

Proof.

1. *From the discussion on the utility function and risk aversion it is possible to define conditions 1 to 3:*

   (a) ***Condition 1:*** *A utility function of constant relative risk aversion is defined by*

$$\begin{aligned} \lambda_1 &= -\frac{d \ln U'(C)}{dC} \div \frac{d \ln C}{dC} \\ &= -\frac{\ln U'(C_t) - \ln U'(C_{t-1})}{\ln C_t - \ln C_{t-1}} \end{aligned}$$

$$\begin{aligned} \ln U'(C_t) - \ln U'(C_{t-1}) &= (-\lambda_1)(\ln C_t - \ln C_{t-1}) \\ \ln \left(\frac{U'(C_t)}{U'(C_{t-1})} \right) &= (-\lambda_1) \ln \left(\frac{C_t}{C_{t-1}} \right) \end{aligned} \tag{6.14}$$

   Since we assume that $\ln(C_t/C_{t-1})$ is normally distributed with constant mean and variance under P, $\ln(U'(C_t)/U'(C_{t-1}))$ is also normal with constant mean and variance.

   (b) ***Condition 2:*** *A utility function of constant absolute risk aversion is defined by*

$$\begin{aligned} \lambda_2 &= -\frac{d \ln U'(C)}{dC} \\ &= -\frac{\ln U'(C_t) - \ln U'(C_{t-1})}{C_t - C_{t-1}} \end{aligned}$$

   thus

$$\begin{aligned} \ln U'(C_t) - \ln U'(C_{t-1}) &= (-\lambda_2)(C_t - C_{t-1}) \\ \ln \left(\frac{U'(C_t)}{U'(C_{t-1})} \right) &= (-\lambda_2)(C_t - C_{t-1}) \end{aligned}$$

   By the assumption that $C_t - C_{t-1}$ is normally distributed with constant mean and variance under P, $\ln(U'(C_t)/U'(C_{t-1}))$ is also normal with constant mean and variance.

   (c) ***Condition 3:*** *A linear utility function is defined by*

$$U(C_t) = aC_t + c$$

   thus

$$U'(C_t) = a$$

and

$$\frac{U'(C_t)}{U'(C_{t-1})} = 1$$

The ratio of marginal utilities

$$\ln\left(\frac{U'(C_t)}{U'(C_{t-1})}\right) = 0 \overset{P}{\sim} N(0,0)$$

From all three conditions it is clear that $\ln\left(\frac{U'(C_t)}{U'(C_{t-1})}\right)$ *is normal with constant mean and variance.*

2. *In section* 6.5 *we saw that under the* $P-$*measure*

$$\begin{aligned} S_{t-1} &= E^P\left[e^{-\rho}\frac{u'(C_t)}{u'(C_{t-1})}S_t \mid \mathcal{F}_{t-1}\right] \\ &= E^P\left[e^{-\rho+\ln\left(\frac{u'(C_t)}{u'(C_{t-1})}\right)}S_t \mid \mathcal{F}_{t-1}\right] \\ &\equiv E^P\left[e^{-\rho+Y_t}S_t \mid \mathcal{F}_{t-1}\right] \end{aligned} \tag{6.15}$$

where $Y_t = \ln\left(\frac{u'(C_t)}{u'(C_{t-1})}\right)$. Y_t, *as mentioned, is normally distributed under conditions 1 to 3. If we define* Q *as*

$$dQ = e^{(r-\rho)T+\sum_{s=1}^{T} Y_s}dP$$

then from theorem 6.7.1, Q *is a measure which is equivalent to* P. *From theorem 6.7.2 we see that* Q *is a probability measure and*

$$E^Q[W_t \mid \mathcal{F}_{t-1}] = E^P\left[W_t e^{(r-\rho)T+Y_y} \mid \mathcal{F}_{t-1}\right]$$

for any W_t *which is* $\mathcal{F}_t$ *measurable. Another result from equation 6.15 stated in theorem 6.7.3 is that*

(a) $ln\left(\frac{S_t}{S_{t-1}}\right) \mid \mathcal{F}_{t-1} \overset{Q}{\sim}$ *normal*

(b) $E^Q\left[\frac{S_t}{S_{t-1}} \mid \mathcal{F}_{t-1}\right] = e^r$ *for all* $t \in \mathbb{R}^+$

(c) $Var^Q\left[\ln\left(\frac{S_t}{S_{t-1}}\right)|\mathcal{F}_{t-1}\right] = Var^P\left[\ln\left(\frac{S_t}{S_{t-1}}\right)|\mathcal{F}_{t-1}\right] \quad P$ *a.s.*

3. *Thus for an economic agent who's an expected utility maximizer, whose utility function is separable, additive and fulfills one of the three stated conditions, the Local risk-neutral Valuation Relationship also holds.*

■

6.8 The Stock Price Process under LRNVR

In this section the stock price process under the LRNVR is derived.

Theorem 6.8.1 *Under the Q – measure, implied by the LRNVR,*

$$\ln \frac{S_t}{S_{t-1}} = r - \frac{1}{2}\sigma_t^2 + \xi_t \sigma_t$$

where

$$\xi_t | \mathcal{F}_{t-1} \overset{Q}{\sim} N(0,1)$$

and

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \left(\xi_{t-i} - \lambda \sigma_{t-i}\right)^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{6.16}$$

Note that the parameters T and t in this context are in terms of time i.e. fractions with the days of the year as the denominator, not as the discrete index. That is, for 63 days in a 252 day year $t = 0.25$.

Proof. *As proved in theorem 6.7.3, $\ln \frac{S_t}{S_{t-1}} | \mathcal{F}_{t-1}$ is normally distributed under measure Q. It can thus be written in terms of a deterministic and random variable*

$$\ln \frac{S_t}{S_{t-1}} = v_t + \xi_t \tag{6.17}$$

under Q.The random variable is obviously normal with mean zero and variance the same of that of $\ln \frac{S_t}{S_{t-1}}$ under Q. It will be proved that

1. $v_t = r - \frac{1}{2}\sigma_t^2$
2. $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \left(\xi_{t-i} - \lambda \sigma_{t-i}\right)^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2$

Proof of 1:
From equation 6.17

$$\begin{aligned} \frac{S_t}{S_{t-1}} &= e^{v_t + \sigma_t \xi_t} \\ E^Q\left[\frac{S_t}{S_{t-1}} | \mathcal{F}_{t-1}\right] &= E^Q\left[e^{v_t + \sigma_t \xi_t} | \mathcal{F}_{t-1}\right] \\ &= e^{v_t} E^Q\left[e^{\sigma_t \xi_t} | \mathcal{F}_{t-1}\right] \end{aligned}$$

then by the moment generating function for an normally distributed random variable we have

$$E^Q\left[\frac{S_t}{S_{t-1}} | \mathcal{F}_{t-1}\right] = e^{v_t + \frac{1}{2} Var^Q\left[\ln \frac{S_t}{S_{t-1}} | \mathcal{F}_{t-1}\right]} E^Q\left[1 | \mathcal{F}_{t-1}\right]$$

Since

$$\begin{aligned} Var^Q\left[\ln\left(\frac{S_t}{S_{t-1}}\right)|\mathcal{F}_{t-1}\right] &= Var^P\left[\ln\left(\frac{S_t}{S_{t-1}}\right)|\mathcal{F}_{t-1}\right] \\ &= \sigma_t^2 \ P-a.s. \end{aligned}$$

from theorem 6.7.3 we can write

$$E^Q\left[\frac{S_t}{S_{t-1}}|\mathcal{F}_{t-1}\right] = e^{\upsilon_t+\frac{1}{2}\sigma_t^2}$$

It was also proved in theorem 6.7.3 that

$$E^Q\left[\frac{S_t}{S_{t-1}} \mid \mathcal{F}_{t-1}\right] = e^r$$

thus

$$\upsilon_t + \frac{1}{2}\sigma_t^2 = r$$

$$\upsilon_t = r - \frac{1}{2}\sigma_t^2$$

Proof of 2.

Recall the original stock price process with GARCH volatility under the P measure, equation 6.2,

$$\ln\frac{S_t}{S_{t-1}} = r + \lambda\sigma_t - \frac{1}{2}\sigma_t^2 + \varepsilon_t$$

and the process implied by proof 1 above under measure $-Q$

$$\ln\frac{S_t}{S_{t-1}} = r - \frac{1}{2}\sigma_t^2 + \xi_t$$

Again using the result

$$Var^Q\left[\ln\left(\frac{S_t}{S_{t-1}}\right)|\mathcal{F}_{t-1}\right] = Var^P\left[\ln\left(\frac{S_t}{S_{t-1}}\right)|\mathcal{F}_{t-1}\right] \ P-a.s$$

from theorem 6.7.3 we can write

$$r + \lambda\sigma_t - \frac{1}{2}\sigma_t^2 + \varepsilon_t = r - \frac{1}{2}\sigma_t^2 + \xi_t$$

thus

$$\varepsilon_t = \xi_t - \lambda\sigma_t$$

Substituting this result into

$$\ln\frac{S_t}{S_{t-1}} = r + \lambda\sigma_t - \frac{1}{2}\sigma_t^2 + \varepsilon_t$$

yields

$$\ln \frac{S_t}{S_{t-1}} = r - \frac{1}{2}\sigma_t^2 + \xi_t$$

and into the GARCH process, yields

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \left(\xi_{t-i} - \lambda\sigma_{t-i}\right)^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{6.18}$$

under the Q measure. ■

The equivalent GJR-GARCH process under the Q measure is

$$\sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + \alpha\left(\xi_{t-1} - \lambda\sigma_{t-1}\right)^2 + \gamma \max\left(-\xi_{t-1} + \lambda\sigma_{t-1}, 0\right)^2$$

where $\gamma > 0$.

The EGARCH variance process under the Q measure is

$$\ln \sigma_t^2 = \alpha_0 + \beta_1 \ln \sigma_{t-1}^2 + \beta_2 \left(\left|\xi_{t-1} - \lambda\sigma_{t-1}\right| - \gamma\left(\xi_{t-1} - \lambda\sigma_{t-1}\right)\right)$$

where $\beta_2, \gamma > 0$.

Theorem 6.8.1 can easily be proved for the above two GARCH processes by substituting equation 6.18 with the respective process and replacing the P variable ε_t with the Q variable $\xi_{t-1} - \lambda\sigma_{t-1}$.

Corollary 6.8.2 *Theorem 6.8.1 implies that under the Q measure*

$$S_T = S_t \exp\left((T-t) \times r - \frac{1}{2}\sum_{s=t+1}^{T} \sigma_s + \sum_{s=t+1}^{T} \xi_s\right) \tag{6.19}$$

Proof. *From theorem 6.8.1 we have that*

$$\ln \frac{S_t}{S_{t-1}} = r - \frac{1}{2}\sigma_t^2 + \xi_t$$

for every $t \in \mathbb{R}$ under Q. Thus

$$\begin{aligned} \ln \frac{S_T}{S_t} &= \sum_{s=t+1}^{T} \ln \frac{S_s}{S_{s-1}} \\ &= \sum_{s=t+1}^{T} r - \frac{1}{2}\sigma_s^2 + \xi_s \\ &= r(T-t) - \frac{1}{2}\sum_{s=t+1}^{T} \sigma_s^2 + \sum_{s=t+1}^{T} \xi_s \end{aligned} \tag{6.20}$$

which means that

$$S_T = S_t \exp\left((T-t)\,r - \frac{1}{2}\sum_{s=t+1}^{T}\sigma_s + \sum_{s=t+1}^{T}\xi_s\right)$$

by taking exponents on both sides of equation 6.20. ■

Corollary 6.8.3 *The discounted process $e^{-rt}S_t$ is a martingale under the Q measure.*

Proof. *Corollary 6.8.2 is equivalent to*

$$S_t = S_{t-1}\exp\left(r - \frac{1}{2}\sigma_t^2 + \xi_t\right)$$

thus the conditional expected value of $e^{-rt}S_t$ is

$$\begin{aligned}
& E^Q\left[\exp(-rt)\,S_t \mid \mathcal{F}_{t-1}\right] \\
= \;& E^Q\left[S_{t-1}\exp(-rt)\exp\left(r - \frac{1}{2}\sigma_t^2 + \xi_t\right) \mid \mathcal{F}_{t-1}\right] \\
= \;& S_{t-1}\exp(-r(t-1))\,E^Q\left[\exp\left(-\frac{1}{2}\sigma_t^2 + \xi_t\right) \mid \mathcal{F}_{t-1}\right] \\
= \;& S_{t-1}\exp(-r(t-1))
\end{aligned}$$

because $\xi_t|\mathcal{F}_{t-1} \overset{Q}{\sim} N\left(0,\sigma_t^2\right)$ and by the moment generating function

$$E^Q\left[\exp(\xi_t) \mid \mathcal{F}_{t-1}\right] = \exp\frac{1}{2}\sigma_t^2$$

which completes the proof. ■

Chapter 7

GARCH Option Pricing and Hedging

7.1 Introduction

This chapter builds on the results of chapter 6. European option on stocks with GARCH volatility is priced under the LRNVR. The delta hedge for such options is also derived.

Delta hedging is defined in Hull [23] as a hedging scheme that is designed to make the price of a portfolio of derivatives insensitive to small changes in the price of the underlying.

In the last section some of the properties of the most widely used GARCH process, the GARCH$(1,1)$ process is discussed.

7.2 Option Pricing under the LRNVR

The stock price process under LRNVR was discussed in the previous chapter. The machinery to model stocks with GARCH volatility can also be adapted to price European options.

Theorem 7.2.1 *GARCH option price. The price of a European call option on a non-dividend paying stock, S_t, expiring at T under LRNVR at time t is*

$$C_t^G = e^{-(T-t)r} E^Q \left[\max\left(S_T - K\right)_+ \mid \mathcal{F}_{t-1} \right]$$

where $\max(x)_+$ is the maximum between x and 0. Note that the parameters T and t in this context is in terms of time i.e. fractions of with the days of the year as the denominator, not as the discrete position of a variable or element of a process.

Proof. See theorem 5.4.2, since $e^{-rt}S_t$ is a martingale under the Q measure. ■

Theorem 7.2.2 *Delta hedge under LRNVR. The delta hedge for a stock with a European call option is*

$$\Delta_t^G = e^{-(T-t)r} E^Q \left[\frac{S_T}{S_t} 1_{[S_T \geq K]} \mid \mathcal{F}_{t-1} \right]$$

where $1_{[S_T \geq K]}$ *is an indicator function and* K *the exercise price of the option.*

Proof. *From corollary 6.8.2*

$$S_T = S_t \exp\left((T-t)\, r - \frac{1}{2} \sum_{s=t+1}^{T} \sigma_s + \sum_{s=t+1}^{T} \xi_s \right)$$

define

$$Y_{t,T} \equiv (T-t)\, r - \frac{1}{2} \sum_{s=t+1}^{T} \sigma_s + \sum_{s=t+1}^{T} \xi_s$$

then

$$E^Q [S_T] = E^Q [S_t \exp(Y_{t,T})]$$

The GARCH option price for a European option proved in theorem 7.2.1 now is

$$C_t^G (S_t) = e^{-(T-t)r} E^Q \left[\max\left(S_t e^{Y_{t,T}} - K\right)_+ \mid \mathcal{F}_{t-1} \right]$$

The delta hedge is the first partial derivative of the option price with respect to the underlying asset price. The strategy is to approximate this derivative with the function C_t^G. *For an arbitrary* $h > 0$

$$\begin{aligned} & C_t^G (S_t + h) - C_t^G (S_t) \\ = \; & e^{-(T-t)r} E^Q \Big[\max\left((S_t + h)\, e^{Y_{t,T}} - K\right)_+ \\ & - \max\left(S_t e^{Y_{t,T}} - K\right)_+ \mid \mathcal{F}_{t-1} \Big] \\ = \; & e^{-(T-t)r} \int_{-\infty}^{\infty} \max\left((S_t + h)\, e^{y} - K\right)_+ \\ & - \max\left(S_t e^{y} - K\right)_+ dF(y \mid \mathcal{F}_t) \end{aligned} \tag{7.1}$$

where $F(y \mid \mathcal{F}_t)$ *is the cdf of* $Y_{t,T}$ *under* Q. *With an indicator function we can express the max function*

$$\max\left((S_t + h)\, e^{Y_{t,T}} - K\right)_+ = \left((S_t + h)\, e^{Y_{t,T}} - K\right) I_{[(S_t+h)e^{Y_{t,T}} - K > 0]}$$

Consider that $h > 0$ *then*

$$(S_t + h)\, e^{Y_{t,T}} - K > 0$$

can be rewritten as

$$\begin{aligned} e^{Y_{t,T}} &> \frac{K}{(S_t+h)} \\ Y_{t,T} &= \ln e^{Y_{t,T}} > \ln \frac{K}{(S_t+h)} \end{aligned}$$

similarly

$$S_t e^{Y_{t,T}} - K > 0$$

can be rewritten as

$$Y_{t,T} = \ln e^{Y_{t,T}} > \ln \frac{K}{S_t}.$$

Equation 7.1 then becomes

$$\begin{aligned} & e^{-(T-t)r} \int_{\ln \frac{K}{(S_t+h)}}^{\infty} (S_t+h)\, e^{y} - K\, dF\,(y \mid \mathcal{F}_t) \\ & -e^{-(T-t)r} \int_{\ln \frac{K}{S_t}}^{\infty} S_t e^{y} - K\, dF\,(y \mid \mathcal{F}_t) \\ = \; & e^{-(T-t)r} \int_{\ln \frac{K}{(S_t+h)}}^{\ln \frac{K}{S_t}} S_t e^{y} - K\, dF\,(y \mid \mathcal{F}_t) \\ & +e^{-(T-t)r} \int_{\ln \frac{K}{(S_t+h)}}^{\infty} h e^{y}\, dF\,(y \mid \mathcal{F}_t)\,. \end{aligned}$$

Since

$$\lim_{h \to 0} e^{-(T-t)r} \int_{\ln \frac{K}{(S_t+h)}}^{\ln \frac{K}{S_t}} S_t e^{y} - K\, dF\,(y \mid \mathcal{F}_t) = 0$$

the

$$\begin{aligned} & \lim_{h \to 0} \frac{C_t^G\,(S_t+h) - C_t^G\,(S_t)}{h} \\ = \; & e^{-(T-t)r} \int_{\ln \frac{K}{S_t}}^{\infty} e^{y}\, dF\,(y \mid \mathcal{F}_t) \\ = \; & e^{-(T-t)r} E^{Q}\left[e^{y} 1_{[S_T > K]}\right] \end{aligned}$$

This argument could similarly have been proven from the left for $h < 0$. *Thus*

$$\begin{aligned} \frac{dC_t}{dS_t} &= \Delta_t \\ &= e^{-(T-t)r} E^{Q}\left[e^{y} 1_{[S_T > K]}\right]^{\cdot} \end{aligned}$$

This completes the proof. ■

The delta hedge of a stock and a European put option can be derived similarly. The delta hedge is

$$\Delta_t^G = e^{-(T-t)r} E^Q \left[\frac{S_T}{S_t} 1_{[K \geq S_T]} \mid \mathcal{F}_{t-1} \right]$$

7.3 Some Properties of the GARCH$(1,1)$ Process under LRNVR

Theorem 7.3.1 *Under measure Q innovations of the GARCH process is $\chi^2(1)$ distributed with non-centrality parameter λ.*

Proof. *From theorem 6.8.1 we have*

$$\sigma_t^2 = \alpha_0 + \alpha\left(\xi_{t-1} - \lambda\sigma_{t-1}\right)^2 + \beta\sigma_{t-1}^2$$

where

$$\xi_t | \mathcal{F}_{t-1} \sim N\left(0, \sigma_t^2\right)$$

thus

$$\frac{\xi_t}{\sigma_t} | \mathcal{F}_{t-1} \sim N(0,1)$$

The innovations of the GARCH process under LRNVR is

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha\left(\xi_{t-1} - \lambda\sigma_{t-1}\right)^2 + \beta\sigma_{t-1}^2 \\ &= \alpha_0 + \alpha\sigma_{t-1}^2\left(\frac{\xi_{t-1}}{\sigma_{t-1}} - \lambda\right)^2 + \beta\sigma_{t-1}^2 \end{aligned}$$

then

$$\frac{1}{\alpha}\frac{\sigma_t^2 - \alpha_0}{\sigma_{t-1}^2} - \frac{\beta}{\alpha} = \left(\frac{\xi_{t-1}}{\sigma_{t-1}} - \lambda\right)^2$$

where

$$\frac{\xi_{t-1}}{\sigma_{t-1}} | \mathcal{F}_{t-2} \sim N(0,1)$$

which completes the proof. ■

Theorem 7.3.2 *Stationary (unconditional) variance of a GARCH process. If*

$$|\lambda| < \sqrt{\frac{(1-\alpha-\beta)}{\alpha}}$$

under probability measure Q then

1. *The stationary variance of* ξ_t,

$$var(\xi_t) = \frac{\alpha_0}{1-(1+\lambda^2)\alpha-\beta}$$

2. ξ_t *is leptokurtic*

3. *The*

$$cov^Q\left(\frac{\xi_t}{\sigma_t}, \sigma^2_{t+1}\right) = \frac{-2\lambda\alpha_0\alpha}{1-(1+\lambda^2)\alpha-\beta}$$

Proof. Proof of part 1.
Under the Q probability measure

$$\begin{aligned}\sigma_t^2 &= \alpha_0 + \alpha\left(\xi_{t-1} - \lambda\sigma_{t-1}\right)^2 + \beta\sigma^2_{t-1} \\ &= \alpha_0 + \alpha\sigma^2_{t-1}\left(\frac{\xi_{t-1}}{\sigma_{t-1}} - \lambda\right)^2 + \beta\sigma^2_{t-1}\end{aligned}$$

let

$$z_t \equiv \frac{\xi_t}{\sigma_t} - \lambda$$

then

$$\begin{aligned}\sigma_t^2 &= \alpha_0 + \alpha\sigma^2_{t-1}z^2_{t-1} + \beta\sigma^2_{t-1} \\ &= \alpha_0 + \sigma^2_{t-1}\left(\alpha z^2_{t-1} + \beta\right)\end{aligned}$$

Using this relationship

$$\sigma^2_{t-1} = \alpha_0 + \sigma^2_{t-2}\left(\alpha z^2_{t-2} + \beta\right)$$

thus

$$\begin{aligned}\sigma_t^2 &= \alpha_0 + \left(\alpha_0 + \sigma^2_{t-2}\left(\alpha z^2_{t-2} + \beta\right)\right)\left(\alpha z^2_{t-1} + \beta\right) \\ &= \alpha_0 + \alpha_0\left(\alpha z^2_{t-1} + \beta\right) + \sigma^2_{t-2}\left(\alpha z^2_{t-2} + \beta\right)\left(\alpha z^2_{t-1} + \beta\right)\end{aligned}$$

and further

$$\begin{aligned}\sigma_t^2 &= \alpha_0 + \alpha_0\left(\alpha z^2_{t-1} + \beta\right) \\ &\quad + \left(\alpha_0 + \sigma^2_{t-3}\left(\alpha z^2_{t-3} + \beta\right)\right)\left(\alpha z^2_{t-2} + \beta\right)\left(\alpha z^2_{t-1} + \beta\right) \\ &= \alpha_0 + \alpha_0\left(\alpha z^2_{t-1} + \beta\right) + \alpha_0\left(\alpha z^2_{t-2} + \beta\right)\left(\alpha z^2_{t-1} + \beta\right) \\ &\quad + \sigma^2_{t-3}\left(\alpha z^2_{t-3} + \beta\right)\left(\alpha z^2_{t-2} + \beta\right)\left(\alpha z^2_{t-1} + \beta\right) \\ &= \alpha_0\left[1 + \sum_{k=1}^{2}\prod_{i=1}^{k}\left(\alpha z^2_{t-i} + \beta\right)\right] + \sigma^2_{t-3}\prod_{k=1}^{3}\left(\alpha z^2_{t-i} + \beta\right)\end{aligned}$$

Then by substituting previous equations for the variance from time $t-1$ to time 0 we obtain

$$\sigma_t^2 = \alpha_0 \left[1 + \sum_{k=1}^{t-1} \prod_{i=1}^{k} \left(\alpha z_{t-i}^2 + \beta\right)\right] + \sigma_0^2 \prod_{k=1}^{t} \left(\alpha z_{t-i}^2 + \beta\right) \tag{7.2}$$

$$\equiv \alpha_0 \sum_{k=0}^{t-1} G_k + \sigma_0^2 G_t \tag{7.3}$$

where

$$\begin{aligned} G_k &\equiv \prod_{i=1}^{k} \left(\alpha z_{t-i}^2 + \beta\right) \qquad (7.4) \\ &= G_{k-1} \left(\alpha z_{t-k}^2 + \beta\right) \\ G_0 &\equiv 1 \end{aligned}$$

From theorem 7.3.1 and the discussion on the chi-square distribution in section 2.5.3

$$z_t^2 = \left(\frac{\xi_t}{\sigma_t} - \lambda\right)^2$$

is chi-square distributed with one degree of freedom and non-centrality parameter λ, since

$$z_t \mid \mathcal{F}_{t-1} \overset{Q}{\sim} N\left(0, \sigma_t^2\right)$$

Thus from the tower property of conditional expectation and theorem 2.5.8

$$\begin{aligned} E^Q\left[E^Q\left[z_t^2 \mid \mathcal{F}_{t-1}\right] \mid \mathcal{F}_0\right] &= E^Q\left[z_t^2 \mid \mathcal{F}_0\right] \\ &= 1 + \lambda^2 \end{aligned}$$

Now from equation 7.4 for $t > k$

$$G_k = \prod_{i=1}^{k} \left(\alpha z_{t-i}^2 + \beta\right)$$

and the conditional expected value of G_k

$$E^Q\left[G_k \mid \mathcal{F}_0\right] = E^Q\left[\prod_{i=1}^{k} \left(\alpha z_{t-i}^2 + \beta\right) \mid \mathcal{F}_0\right]$$

Since z_r and z_s are independently distributed for all applicable r, s. z_r^2 and z_s^2 are also independent (see theorem 2.4.9). This allows us to write

$$cov\left(\alpha z_{t-i}^2 + \beta, \alpha z_{t-j}^2 + \beta\right) = 0$$

which follows from theorem 2.4.9 such that

$$\begin{aligned} E^Q\left[G_k \mid \mathcal{F}_0\right] &= \prod_{i=1}^{k} E^Q\left[\left(\alpha z_{t-i}^2+\beta\right) \mid \mathcal{F}_0\right] \\ &= \prod_{i=1}^{k} a\left(1+\lambda^2\right)+\beta \\ &= \left[a\left(1+\lambda^2\right)+\beta\right]^k \end{aligned}$$

Using this result we can write the conditional expectation of equation 7.2

$$\begin{aligned} & E^Q\left[\sigma_t^2 \mid \mathcal{F}_0\right] \\ = \ & E^Q\left[\alpha_0\left[1+\sum_{k=1}^{t-1}\prod_{i=1}^{k}\left(\alpha z_{t-i}^2+\beta\right)\right]+\sigma_0^2\prod_{k=1}^{t}\left(\alpha z_{t-i}^2+\beta\right) \mid \mathcal{F}_0\right] \\ = \ & \alpha_0\sum_{k=0}^{t-1}\left[a\left(1+\lambda^2\right)+\beta\right]^k+\sigma_0^2\left[a\left(1+\lambda^2\right)+\beta\right]^t \end{aligned}$$

Using the condition that

$$|\lambda|<\sqrt{\frac{(1-\alpha-\beta)}{\alpha}}$$

the term

$$a\left(1+\lambda^2\right)+\beta<a\left(1+\frac{(1-\alpha-\beta)}{\alpha}\right)+\beta=1$$

The stationary variance is the limit of t to infinite of $E\left[\sigma_t^2\right]$. By again using the tower property of conditional expectation

$$E^Q\left[E^Q\left[\sigma_t^2 \mid \mathcal{F}_0\right]\right]=E^Q\left[\sigma_t^2\right]$$

$$\begin{aligned} \lim_{t\to\infty} E^Q\left[\sigma_t^2\right] &= \alpha_0\sum_{k=0}^{\infty}\left[a\left(1+\lambda^2\right)+\beta\right]^k \\ &= \frac{\alpha_0}{a\left(1+\lambda^2\right)+\beta} \end{aligned}$$

Proof of part 2.
We need to prove that $E^Q\left[\xi_t^4\right]>3\left(E^Q\left[\xi_t^2\right]\right)^2$ since

$$\xi_t \mid \mathcal{F}_0 \overset{Q}{\sim} N\left(0,\sigma_t^2\right)$$

In theorem 2.5.9 it was proved that

$$E^Q\left[z_t^4 \mid \mathcal{F}_0\right]=3+6\lambda^2+\lambda^4$$

thus for $t > k$

$$\begin{aligned}
& E^Q\left[G_k^2 \mid \mathcal{F}_0\right] \\
= \;& E^Q\left[\left(\prod_{i=1}^{k} \alpha z_{t-i}^2 + \beta\right)^2 \mid \mathcal{F}_0\right] \\
= \;& E^Q\left[\prod_{i=1}^{k} \left(\alpha z_{t-i}^2 + \beta\right)^2 \mid \mathcal{F}_0\right]
\end{aligned}$$

Since

$$z_t^2 = \left(\frac{\xi_t}{\sigma_t} - \lambda\right)^2$$

is noncentral chi-square distributed with 1 degree of freedom and non-centrality parameter λ. Again as in part 1, it follows from theorem 2.4.9 that

$$cov\left(\left(\alpha z_{t-i}^2 + \beta\right)^2, \left(\alpha z_{t-j}^2 + \beta\right)^2\right) = 0$$

for all $i, j \in \{0, 1, ..., k\}$ and $i \neq j$. Then

$$\begin{aligned}
& E^Q\left[\prod_{i=1}^{k} \left(\alpha z_{t-i}^2 + \beta\right)^2 \mid \mathcal{F}_0\right] \\
= \;& \prod_{i=1}^{k} E^Q\left[\left(\alpha z_{t-i}^2 + \beta\right)^2 \mid \mathcal{F}_0\right] \\
= \;& \prod_{i=1}^{k} E^Q\left[\alpha^2 z_{t-i}^4 + 2\alpha\beta z_{t-i}^2 + \beta^2 \mid \mathcal{F}_0\right] \\
= \;& \prod_{i=1}^{k} \left[\alpha^2\left(3 + 6\lambda^2 + \lambda^4\right) + 2\alpha\beta\left(1 + \lambda^2\right) + \beta^2\right] \\
= \;& \left[\alpha^2\left(3 + 6\lambda^2 + \lambda^4\right) + 2\alpha\beta\left(1 + \lambda^2\right) + \beta^2\right]^k
\end{aligned}$$

For notational purposes define

$$\begin{aligned}
u &\equiv \alpha^2\left(3 + 6\lambda^2 + \lambda^4\right) + 2\alpha\beta\left(1 + \lambda^2\right) + \beta^2 \\
v &= \alpha\left(1 + \lambda^2\right) + \beta
\end{aligned}$$

then

$$u \equiv v^2 + 2\alpha^2\left(1 + 2\lambda^2\right)\alpha^2 \tag{7.5}$$

and

$$u > v$$

since all terms of equation 7.5 are positive.

For $k > j$

$$
\begin{aligned}
& E^Q\left[G_k G_j \mid \mathcal{F}_0\right] \\
= & E^Q\left[\prod_{i=1}^{k}\left(\alpha z_{t-i}^2+\beta\right)\prod_{i=1}^{j}\left(\alpha z_{t-i}^2+\beta\right) \mid \mathcal{F}_0\right] \\
= & E^Q\left[\prod_{i=1}^{j}\left(\alpha z_{t-i}^2+\beta\right)^2\prod_{i=j+1}^{k}\left(\alpha z_{t-i}^2+\beta\right) \mid \mathcal{F}_0\right]
\end{aligned}
$$

By theorem 2.4.9

$$
cov\left[\left(\alpha z_{t-i}^2+\beta\right)^2,\left(\alpha z_{t-i}^2+\beta\right)\right]=0
$$

for all $i, j \in \{0,1,\ldots,k\}$, $k > j$. *Thus*

$$
\begin{aligned}
& E^Q\left[\prod_{i=1}^{j}\left(\alpha z_{t-i}^2+\beta\right)^2\prod_{i=j+1}^{k}\left(\alpha z_{t-i}^2+\beta\right) \mid \mathcal{F}_0\right] \\
= & \prod_{i=1}^{j}E^Q\left[\left(\alpha z_{t-i}^2+\beta\right)^2 \mid \mathcal{F}_0\right]\prod_{i=j+1}^{k}E^Q\left[\left(\alpha z_{t-i}^2+\beta\right) \mid \mathcal{F}_0\right] \\
= & u^j v^{k-j}
\end{aligned}
$$

Then the conditional expected value of σ_t^4, *the square of the GARCH process under LRNVR at time t follows from equation 7.3*

$$
\begin{aligned}
& E^Q\left[\sigma_t^4 \mid \mathcal{F}_0\right] \\
= & E^Q\left[\left(\alpha_0\sum_{k=0}^{t-1}G_k+\sigma_0^2G_t\right)^2 \mid \mathcal{F}_0\right] \\
= & E^Q\left[\left(\alpha_0\sum_{k=0}^{t-1}G_k\right)^2+2\sigma_0^2G_t\left(\alpha_0\sum_{k=0}^{t-1}G_k\right)+\left(\sigma_0^2G_t\right)^2 \mid \mathcal{F}_0\right] \\
= & \alpha_0^2E^Q\left[\left(\sum_{k=0}^{t-1}G_k\right)^2 \mid \mathcal{F}_0\right]+2\alpha_0\sigma_0^2\sum_{k=0}^{t-1}E^Q\left[G_tG_k \mid \mathcal{F}_0\right] \\
& +\sigma_0^4E^Q\left[G_t^2 \mid \mathcal{F}\right] \\
= & \sigma_0^4u^t+2\alpha_0\sigma_0^2\sum_{k=0}^{t-1}u^kv^{t-k}+\alpha_0^2\left[\sum_{k=0}^{t-1}u^k+2\sum_{k=0}^{t-1}\sum_{j=0}^{k-1}u^jv^{k-j}\right] \qquad (7.6)
\end{aligned}
$$

where the third term is a common mathematical expansion. The properties of geometric series (see Haggarty [21]) are used to simplify equation 7.6:

1. *Geometric series*

$$\sum_{k=0}^{t-1} u^k = \frac{1-u^t}{1-u}$$

2. *Geometric series where v^t is independent of the summation*

$$\begin{aligned}\sum_{k=0}^{t-1} u^k v^{t-k} &= v^t \sum_{k=0}^{t-1} \left(\frac{u}{v}\right)^k \\ &= v^t \frac{\left(1-\left(\frac{u}{v}\right)^t\right)}{1-\frac{u}{v}} \\ &= v\frac{u^t - v^t}{u-v}\end{aligned}$$

3. *Geometric series using point 2, where $\frac{v}{u-v}$ is independent of the summation*

$$\begin{aligned}\sum_{k=0}^{t-1}\sum_{j=0}^{k-1} u^j v^{k-j} &= \sum_{k=0}^{t-1} v\frac{u^k - v^k}{v-u} \\ &= \frac{v}{u-v}\sum_{k=0}^{t-1} u^k - \sum_{k=0}^{t-1} v^k \\ &= \frac{v}{u-v}\left(\frac{1-u^t}{1-u} - \frac{1-v^t}{1-v}\right)\end{aligned}$$

Equation 7.6 is simplified such that

$$\begin{aligned}& E^Q\left[\sigma_t^4 \mid \mathcal{F}_0\right] \\ = \; & \sigma_0^4 u^t + 2\alpha_0\sigma_0^2 v\frac{u^t - v^t}{u-v} \\ & +\alpha_0^2\left[\frac{1-u^t}{1-u} + 2\frac{v}{u-v}\left(\frac{1-u^t}{1-u} - \frac{1-v^t}{1-v}\right)\right]\end{aligned}$$

Now, to derive the value of the unconditional kurtosis of ξ_t, we take the limit of $E^Q\left[\sigma_t^4 \mid \mathcal{F}_0\right]$. If we remember that $u > v$ and assume that $u \geq 1$ then

$$E^Q\left[\sigma_t^4\right] = \lim_{t\to\infty} E^Q\left[\sigma_t^4 \mid \mathcal{F}_0\right] = \infty \tag{7.7}$$

and if $u < 1$

$$\lim_{t\to\infty} E^Q\left[\sigma_t^4 \mid \mathcal{F}_0\right] = \alpha_0^2 \frac{(1-v)}{(1-u)(1-v)}$$

since

$$\begin{aligned}
&\lim_{t\to\infty} \alpha_0^2 \left[\frac{1-u^t}{1-u} + 2\frac{v}{u-v}\left(\frac{1-u^t}{1-u} - \frac{1-v^t}{1-v}\right)\right] \\
&= \lim_{t\to\infty} 2\alpha_0^2 \left[\frac{v}{u-v}\left(\frac{1-u^t}{1-u} - \frac{1-v^t}{1-v}\right)\right] + \frac{\alpha_0^2}{1-u} \\
&= 2\alpha_0^2 \left[\frac{v}{u-v}\left(\frac{1}{1-u} - \frac{1}{1-v}\right)\right] + \frac{\alpha_0^2}{1-u} \\
&= \alpha_0^2 \left[\frac{2v}{u-v}\frac{(1-v)-(1-u)}{(1-u)(1-v)} + \frac{1}{1-u}\right] \\
&= \alpha_0^2 \left[\frac{2v}{u-v}\frac{u-v}{(1-u)(1-v)} + \frac{1}{1-u}\right] \\
&= \alpha_0^2 \left[\frac{2v+1-v}{(1-u)(1-v)}\right] = \frac{\alpha_0^2(1+v)}{(1-u)(1-v)} \\
&= E^Q\left[\sigma_t^4\right] \qquad (7.8)
\end{aligned}$$

Since σ_t^2 *is* $\mathcal{F}_{t-1}$ *measurable under* Q *and*

$$\frac{\xi_t}{\sigma_t} \mid \mathcal{F}_0 \overset{Q}{\sim} N(0,1)$$

the

$$\begin{aligned}
E^Q\left[\xi_t^4\right] &= E^Q\left[E^Q\left[\xi_t^4 \mid \mathcal{F}_{t-1}\right]\right] \\
&= E^Q\left[\sigma_t^4 E^Q\left[\left(\frac{\xi_t}{\sigma_t}\right)^4 \mid \mathcal{F}_{t-1}\right]\right] \\
&= 3E^Q\left[\sigma_t^4\right],
\end{aligned}$$

where

$$E^Q\left[\left(\frac{\xi_t}{\sigma_t}\right)^4 \mid \mathcal{F}_{t-1}\right]$$

is the kurtosis under Q.

Finally, from equation 7.7 it is clear that ξ_t *is leptokurtic if* $u \geq 1$. *If* $u < 1$ *then*

$$\begin{aligned}
E^Q\left[\xi_t^4\right] &= 3\frac{\alpha_0^2(1+v)}{(1-u)(1-v)} \\
&= 3\frac{1-v^2}{1-u}\left(E^Q\left[\xi_t^2\right]\right)^2
\end{aligned}$$

from the definition of v. *Since* $u > v > 0$

$$E^Q\left[\xi_t^4\right] > 3\left(E^Q\left[\xi_t^2\right]\right)^2$$

Proof of part 3.

From theorem 6.8.1 equation 6.16 we have

$$\sigma_t^2 = \alpha_0 + \alpha\left(\xi_{t-1} - \lambda\sigma_{t-1}\right)^2 + \beta\sigma_{t-1}^2$$

and

$$\xi_t|\mathcal{F}_{t-1} \sim N\left(0, \sigma_t^2\right)$$

Thus

$$\frac{\xi_t}{\sigma_t}\sigma_{t+1}^2 = \frac{\xi_t}{\sigma_t}\alpha_0 + \alpha\frac{\xi_t}{\sigma_t}\left(\xi_t - \lambda\sigma_t\right)^2 + \beta\frac{\xi_t}{\sigma_t}\sigma_t^2$$

and

$$\begin{aligned} E^Q\left[\frac{\xi_t}{\sigma_t}\sigma_{t+1}^2|\mathcal{F}_{t-1}\right] &= E^Q\left[\frac{\xi_t}{\sigma_t}\alpha_0 + \alpha\frac{\xi_t}{\sigma_t}\left(\xi_t - \lambda\sigma_t\right)^2 + \beta\frac{\xi_t}{\sigma_t}\sigma_t^2|\mathcal{F}_{t-1}\right] \\ &= E^Q\left[\alpha\frac{\xi_t}{\sigma_t}\left(\xi_t - \lambda\sigma_t\right)^2|\mathcal{F}_{t-1}\right] \end{aligned}$$

since

$$E^Q\left[\xi_t|\mathcal{F}_{t-1}\right] = 0$$

Then

$$\begin{aligned} & E^Q\left[\alpha\frac{\xi_t}{\sigma_t}\left(\xi_t - \lambda\sigma_t\right)^2|\mathcal{F}_{t-1}\right] \\ = \;& \alpha E^Q\left[\frac{\xi_t}{\sigma_t}\left(\xi_t^2 - 2\lambda\xi_t\sigma_t + (\lambda\sigma_t)^2\right)|\mathcal{F}_{t-1}\right] \\ = \;& \alpha E^Q\left[\frac{\xi_t^3}{\sigma_t} - 2\lambda\xi_t^2 + \lambda^2\xi_t\sigma_t|\mathcal{F}_{t-1}\right] \\ = \;& \frac{\alpha}{\sigma_t}E^Q\left[\xi_t^3|\mathcal{F}_{t-1}\right] - 2\alpha\lambda E^Q\left[\xi_t^2|\mathcal{F}_{t-1}\right] + \alpha\lambda^2\sigma_t E^Q\left[\xi_t|\mathcal{F}_{t-1}\right] \end{aligned}$$

since σ_t is $\mathcal{F}_{t-1}$ measurable, the

$$\begin{aligned} & E^Q\left[\alpha\frac{\xi_t}{\sigma_t}\left(\xi_t - \lambda\sigma_t\right)^2|\mathcal{F}_{t-1}\right] \\ = \;& -2\alpha\lambda E^Q\left[\xi_t^2|\mathcal{F}_{t-1}\right] \\ = \;& -2\alpha\lambda\sigma_t^2. \end{aligned}$$

Finally,

$$
\begin{aligned}
& cov^Q\left[\frac{\xi_t}{\sigma_t}, \sigma_{t+1}^2\right] \\
= \; & E^Q\left[\frac{\xi_t}{\sigma_t}\sigma_{t+1}^2\right] - E^Q\left[\frac{\xi_t}{\sigma_t}\right]E^Q\left[\sigma_{t+1}^2\right] \\
= \; & E^Q\left[\frac{\xi_t}{\sigma_t}\sigma_{t+1}^2\right] \\
= \; & E^Q\left[E^Q\left[\frac{\xi_t}{\sigma_t}\sigma_{t+1}^2 | \mathcal{F}_{t-1}\right]\right]
\end{aligned}
$$

the tower property of conditional expectation. The

$$
\begin{aligned}
& cov^Q\left[\frac{\xi_t}{\sigma_t}, \sigma_{t+1}^2\right] \\
= \; & E^Q\left[E^Q\left[\alpha\frac{\xi_t}{\sigma_t}\left(\xi_t - \lambda\sigma_t\right)^2 | \mathcal{F}_{t-1}\right]\right] \\
= \; & E^Q\left[-2\alpha\lambda\sigma_t^2\right] \\
= \; & -2\alpha\lambda E^Q\left[\sigma_t^2\right] \\
= \; & \frac{-2\lambda\alpha_0\alpha}{1-\left(1+\lambda^2\right)\alpha-\beta}
\end{aligned}
$$

by the proof of 2. ∎

Part IV

Implementation and Numerical Results

Chapter 8

Implementation of GARCH Option Pricing

8.1 Introduction

In this chapter, methods to implement the GARCH option pricing model is discussed. Two separate numerical procedures are required in the implementation of this model, the first is the calibration of the parameters to the stock or option data and the second is the forecast of the option price.

8.2 Calibrating the GARCH Process to Empirical Data

8.2.1 Historical Data

In the GARCH option pricing procedure proposed by Jin-Chuan Duan, the GARCH process is "fitted" to the process of the underlying stock or index. This means that the parameters of the GARCH-M process under the P measure is fitted to the returns series of the underlying by maximizing its loglikelihood function.

For the (vanilla) GARCH(1,1) $-M$ process under the P measure

$$S_t = S_{t-1} \exp\left(r\Delta t - \frac{1}{2}\sigma_t^2 + \lambda\sigma_t + \varepsilon_t\right)$$

where ε_t is the returns at time t, the rest of the parameters and variables are as in section 4.6. The vanilla GARCH(1,1) process is[1]

$$\hat{\sigma}_t^2 = \hat{\alpha}_0 + \hat{\alpha}\varepsilon_t^2 + \hat{\beta}\hat{\sigma}_{t-1}^2.$$

[1] Estimates are written with hats.

From section 4.4.3, equation 4.9 the likelihood function for the variance up to time t is

$$\begin{aligned} f\left(0, \hat{\sigma}_t^2\right) &= \sum_{i=1}^{t} -\frac{1}{2} \ln \hat{\sigma}_i^2 - \frac{1}{2} \frac{\varepsilon_i^2}{\hat{\sigma}_i^2} \\ &= \sum_{i=1}^{t} l_i \left(\alpha_0, \alpha, \beta, \sigma_0^2\right) \end{aligned}$$

The optimization problem for the variance is as follows

$$\max_{\hat{\alpha}_0, \hat{\alpha}, \hat{\beta}, \hat{\sigma}_0^2} f\left(0, \hat{\sigma}_t^2\right)$$

where the likelihood function $f\left(0, \hat{\sigma}_t^2\right)$ is maximized in terms of the parameters $\hat{\alpha}_0, \hat{\alpha}$ and $\hat{\beta}$. Since the value of $\hat{\sigma}_0^2$ isn't known, it forms part of the optimization problem.

The value of parameter $\hat{\lambda}$ is then estimated by minimizing the sum of squares between the actual and estimated stock prices up to time t

$$\begin{aligned} & \min_{\hat{\lambda}} \sum_{i=1}^{t} \left(S_i - \hat{S}_i\right)^2 \\ = & \min_{\hat{\lambda}} \sum_{i=1}^{t} \left(S_i - \hat{S}_{i-1} \exp\left(r\Delta t - \frac{1}{2}\hat{\sigma}_i^2 + \hat{\lambda}\hat{\sigma}_i + \varepsilon_i\right)\right)^2 \end{aligned}$$

$\hat{S}_i$ is an estimate of the of the stock price at time i and S_i is the actual stock price.

Both of the optimization problems are due to overdetermined systems. This means that there are more equations than variables. Tim Bollerslev (1986) suggests the use of the Berndt, Hall, Hall and Hausman algorithm for the estimation of the variance optimization problem. A similar algorithm can also be used for the mean optimization problem.

Many new statistical computer packages have built-in GARCH algorithms. Often, the problem with these built-in algorithms are that they are designed to solve only certain types of GARCH parameter estimation problems.

The GARCH toolbox available with Matlab R 12 is only limited to solving vanilla GARCH problems. Fortunately the optimization toolbox of Matlab is excellent. The procedure `fmincon` can be used to fit a tailor-made GARCH and means process.

8.2.2 Implied Volatility

As mentioned in section 1.3 of the introduction, the levels of implied volatility of warrants are substantially higher than that of the historical volatility of the underlying. This means that there is little use in pricing an option with a model based on the history of the underlying.

In this dissertation we investigate the calibration of a GARCH process to the implied volatility of warrant..

The approach is as follows:

1. In section 5.4.5, the discussion of the Black-Scholes formula and implied volatility, it was mentioned that implied volatility is annualized. Implied volatility at time t, say σ_t must thus by multiplied by the square root of the relevant time fraction, for instance if the available returns series is daily then the new adjusted series must be $\sigma_t\sqrt{1/252}$.

2. The GARCH process is a variance process, not a standard deviations process, thus the square of the new series in point 1 must be taken, which gives $\sigma_t^2/252 \equiv I_t$.

3. Implied volatility is used to price options, thus it is already under the Q measure. This means that the unit risk premium λ is already "absorbed" into the GARCH process.

4. The parameter estimation for the Asymmetric GARCH(p,q) process is as follows,

$$\begin{aligned}
& \min_{\hat{\alpha}_0,\hat{\alpha},\hat{\beta}} \sum_{i=1}^{t} \left(I_t - \sigma_t^2\right)^2 \\
= & \min_{\hat{\alpha}_0,\hat{\alpha},\hat{\beta},\lambda} \sum_{i=1}^{t} \\
& \left(I_t - \alpha_0 + \sum_{j=1}^{q} \alpha_j \left(\varepsilon_{i-j} - \lambda\sigma_{i-j}\right)^2 + \sum_{j=1}^{p} \beta_j \sigma_{i-j}^2\right)^2 \qquad (8.1)
\end{aligned}$$

   Unlike the parameter estimation in section 8.2.1 above, the value of σ_0^2 here isn't part of the minimization problem. That is because if we let σ_0^2 equal I_0, the value at time $i = 0$ in equation 8.1 is zero.

5. Optimization here is again done with the `fmincon` procedure of Matlab.

8.3 Monte Carlo Simulations

Monte Carlo simulations is a method to solve stochastic integrals numerically. This is done by simulating N sample paths of a stochastic processes, say f by the generation of random variables from the underlying probability distribution. All the versions of f are then added together and divided by the amount of simulations. By the law of large numbers we can write

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}f\left(X_n\right)=\int_R f\left(x\right)q\left(x\right)dx \tag{8.2}$$

where (X_n) are independently drawn from the distribution with pdf q.

In this dissertation q is the pdf of the normal distribution discussed in section 2.5.

8.3.1 European Option with Constant Volatility

The pricing theorem for a European option in the Black-Scholes sense, theorem 5.4.2, yields

$$\begin{aligned} V_t &= e^{-r(T-t)}E^Q\left[V_T \mid \mathcal{F}_t\right] \\ &= e^{-r(T-t)}E^Q\left[f_T \mid \mathcal{F}_t\right] \\ &= e^{-r(T-t)}\int_R f_T\left(x\right)q\left(x\right)dx \end{aligned}$$

where q is a pdf. For a put option, $f_T = (X - S_T)_+$, where X and S_T are the strike price and the stock price at time T respectively. S_T and thus f_T is a function of Brownian motion. By equation 5.12,

$$S_T = S_0\exp\left(\left(\mu-\frac{1}{2}\sigma^2\right)T+\sigma W_t\right)$$

In discrete time, this can be estimated by

$$\widehat{S_T} = \widehat{S_0}\exp\left(\left(\mu-\frac{1}{2}\sigma^2\right)T+\sigma\xi_t\sqrt{\Delta t}\right)$$

where

$$\xi_t|\mathcal{F}_{t-1}\overset{Q}{\sim}N\left(0,1\right)$$

8.3.2 European Options with GARCH Volatility

The aim is again to estimate the value of f_T at time T. This time it must be remembered the the GARCH-M process used in this dissertation is defined in discrete time, we are thus not solving an integral. The stock price process, under the LRNVR with GARCH conditional volatility, as defined in theorem 6.8.1 is

$$S_T = S_t\exp\left(\left(T-t\right)r-\sum_{i=t}^{T}\left(\frac{1}{2}\sigma_i^2+\xi_i\sigma_i\right)\right)$$

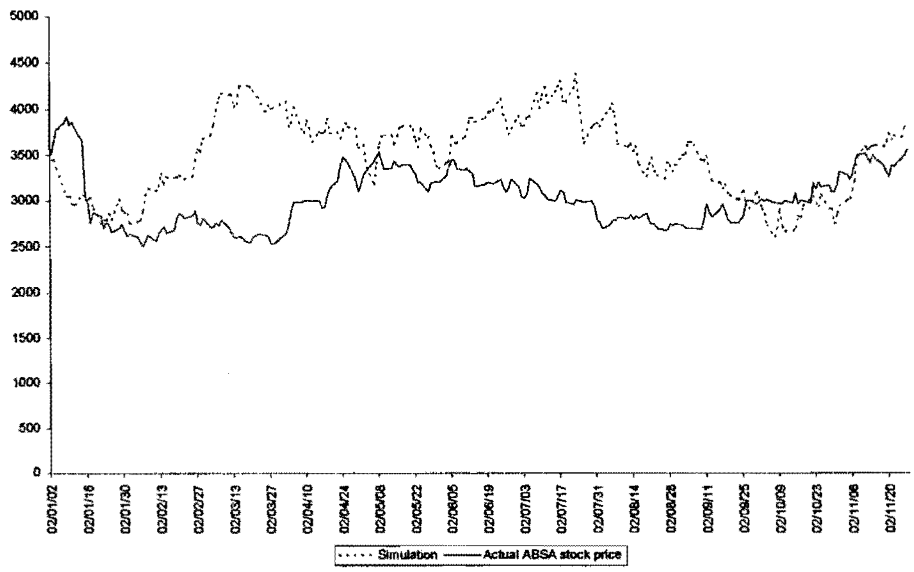

Figure 8.1: A sample path of a Monte Carlo simulation compared with the actual ABSA stock price process.

where

$$\xi_i|\mathcal{F}_{i-1} \overset{Q}{\sim} N(0,1)$$

and

$$\sigma_i^2 = \alpha_0 + \sum_{j=1}^{q} \alpha_j \left(\xi_{i-j} - \lambda\sigma_{i-j}\right)^2 + \sum_{i=1}^{p} \beta_j \sigma_{i-j}^2$$

8.3.3 Notes

1. To simulate possible sample paths of the stock price, a random number ξ_t is generated for each interval $t \in N \cap [1, T]$. The intervals are equally spaced, say of size Δt. If we use an annual risk-free interest rate and daily time intervals, Δt would be $1/252$, since we usually assume 252 trading days in a year.

2. A large number of future paths are simulated. The number depends on the accuracy required. This can vary between a 1000 and 50000 or even more simulations.

8.3.4 Generating Other Distributions from the Uniform Distribution

Many computer packages can only generate uniform random variables between 0 and 1. Most other packages, like Matlab and Excel, generate only random variables from certain famous distributions. The following famous technique is a way to generate random variables from uniform random variables:

To generate random variables from this cdf, use the following famous result: Say we are able to generate a uniform random variable, U, between 0 and 1. Define the inverse of $\widehat{F_Z(z)}$ as

$$\widehat{F_Z(y)}^{-1} = \inf\left\{z \mid \widehat{F_Z(z)} \leq y\right\}$$

where $0 \leq y \leq 1$, thus

$$\widehat{F_Z(U)}^{-1} = Z.$$

It then simply follows that

$$\begin{aligned}\widehat{F_Z(z)} &= P(Z \leq z) \\ &= P\left(\widehat{F_Z(U)}^{-1} \leq z\right) \\ &= P\left(U \leq \widehat{F_Z(z)}\right)\end{aligned}$$

Thus by generating a value u from U, calculate $\widehat{F_Z(u)}^{-1}$ which is set equal to z. This yields a $\widehat{F_Z(z)}$ distributed random variable.

8.4 Variance Reduction Techniques

Monte Carlo simulations are computationally expensive. It is practical to employ variance reduction procedures to decrease the number of simulations needed. Hull [23] gives a broad summary of variance reduction procedures.

The variance reduction procedures used in this dissertation are the *antithetic variable* and *moment matching* techniques. The control variate technique[2] can possibly also be used. To be certain of the soundness of the use of the control variate technique for the particular simulations done in this dissertation, further investigation is needed. This is beyond the scope of the dissertation.

[2] For the simulations in this dissertation, for the option price resultant from a forecasted GARCH-M process, we can do the Controle Variate Technique as follows:

Two simulations, the standard Black-Scholes option pricing integral and the Duan GARCH integral are done in parallel using the same random variables.

The Black-Scholes simulation is then subtracted from the GARCH one and the equivalent analytical Black-Scholes value is added.

8.4.1 The Antithetic Variable Technique

With this technique two values for the derivative is calculated. The first value, f^1 is calculated the normal way with random sample vector $[\xi_t]_{T\times 1}$ taken from the applicable distribution. For the second value f^2, $-1 \times [\xi_t]_{T\times 1}$ is used. The final answer is the average of the two values

$$\frac{f^1 + f^2}{2}$$

The advantage of this technique is that a value above the true value can be "canceled out" by one below and vice versa.

8.4.2 Moment Matching

In this dissertation the standard normal distribution is used. In the moment matching technique all of the sampled random variables for each sample path, say the vector $[\xi_t]_{T\times 1}$ is standardized. This is done by subtracting the mean of the sample m from each element of the sample and then dividing that by the standard deviation s of the sample,

$$y_t = \frac{\xi_t - m}{s}$$

yielding a standard normal random variable.

Chapter 9

Study and Results

9.1 Aim

It is generally assumed that the current implied volatility level is the best proxy for the future level of implied volatility and hence the future price of an option. In this chapter, the GARCH option pricing method is applied to the implied volatility history of a warrant. The method is as described in section 8.2.2, where I propose calibrating the GARCH process over the "historical" implied volatility of the underlying financial instrument, in this case stock.

In this study, the current implied volatility level is compared to the GARCH level or equivalently, the predicted future price of the option compared to the predicted price of under the GARCH option pricing method

The predicted future price of a European option, to avoid arbitrage, is its current value adjusted for the relevant risk-free interest rate.

9.2 Methodology and Data

JSE warrants are generally short dated, that is of maturity less than one year. An option pricing model must thus be able to price a warrant, with as little calibration to historical data as possible. Here, a 30-day period of calibration to implied volatility was decided on in each case.

Although the warrants market is more liquid than the options market of SAFEX, there are still days where no new trade takes place in a specific warrant. The result, is that after a sharp drop in the price in the underlying equity, the intrinsic value of the replicating portfolio may be greater than the market price of the an untraded put warrant. The implied volatility of that warrant is thus undefined at such a date.

In this study, the chosen warrants where priced in a rolling window of one day (with a thirty day history each), from approximately thirty days after they where first traded, up to a date where either the implied volatility is undefined or zero, the warrant reaches maturity or 2002/11/27[1].

The 11 to 20 day ahead values of both the forecasts due to the actual warrant prices and the GARCH option prices are compared to the actual warrant prices of 11 to 20 days ahead. The measurement over a 10 day period gives a better indication of the forecasting power of the two methods.

The following put warrants were selected:

Result no:	Warrant	Underlying	Issuer	Date From	Date To
1	*3ASAIB*	*ABSA*	*Investec Bank*	*2002/05/15*	*2002/11/06*
2	3ASAUB	ABSA	UBS	2002/01/02	2002/11/06
3	*5ASAIB*	*ABSA*	*Investec Bank*	*2002/01/08*	*2002/11/06*
4	2AGLUB	Anglo American	UBS	2001/06/08	2002/04/04
5	*3AGLIB*	*Anglo American*	*Investec Bank*	*2001/09/25*	*2002/02/27*
6	7AGLIB	Anglo American	Investec Bank	2002/02/22	2002/07/25
7	*BAGLIB*	*Anglo American*	*Investec Bank*	*2002/07/29*	*2002/09/17*
8	3NEDUB	Nedcor	UBS	2002/01/02	2002/08/19
9	*6NEDIB*	*Nedcor*	*Investec Bank*	*2002/07/08*	*2002/10/03*
10	6NEDSG	Nedcor	Societe General	2002/08/23	2002/09/20
11	*3OMLUB*	*Old Mutual*	*UBS*	*2002/01/02*	*2002/06/24*
12	4OMLSG	Old Mutual	Societe General	2002/08/16	2002/09/16
13	*5OMLIB*	*Old Mutual*	*Investec Bank*	*2002/05/15*	*2002/06/13*
14	3SAPIB	Sappi	Investec Bank	2002/04/16	2002/10/07
15	*3SAPUB*	*Sappi*	*UBS*	*2002/01/02*	*2002/11/06*

Each warrant can be categorized in terms of time to maturity and moneyness:

- Time to maturity of a warrant is the amount of days left in the life of the warrant. A warrant's implied volatility tends to increase dramatically 70 to 60 days and closer, to maturity. It seems sensible to categorize results in terms of the time to maturity of the warrant. The two categories are maturity of less than 70 days and maturity of 70 days and more.

- Moneyness is defined as the stock price divided by the exercise price of a warrant. A put warrant is defined to be "out of the money" when the moneyness ratio is more that 1.1, "at the money" if the ratio is between 0.9 and 1.1 and "in the money" if the ration is less than 0.9.

9.3 Measures of Results

The accuracy of the implied volatility method and the GARCH option pricing method is measured in the following way:

[1] The last date on which data was captured.

1. The current market implied volatility, $I(t)$ and the GARCH option price, $\sigma^2(t)$ at time t are adjusted to the (annual) risk-free rate r for each day of the 10 day period starting in 11 days,

$$\begin{aligned} I(t,i) &\equiv e^{r\times(10+i)/252} I(t) \\ \sigma^2(t,i) &\equiv e^{r\times(10+i)/252}\sigma^2(t) \end{aligned}$$

for $i = 1, \ldots, 10$.

2. The absolute percentage deviations between the two forecasts,

$$I(t,i) \text{ and } \sigma^2(t,i)$$

and the actual observed implied volatility in the market,

$$I(t+i+10,0)$$

is taken for each day of the 10 day period and weighed as follow

$$\begin{aligned} \Delta^i_{Actual} &= \frac{|I(t,i) - I(t+i+10,0)|}{I(t+i+10,0)} \\ \Delta^i_{GARCH} &= \frac{|\sigma^2(t,i) - \sigma^2(t+i+10,0)|}{I(t+i+10,0)} \end{aligned}$$

3. The following risk-measures are determined

$$\begin{aligned} \Delta^{below}_{Actual} &= \sum_{i=1}^{10} \Delta^i_{Actual} \mathbf{1}_{I(t,i)<I(t+i+10,0)} \\ \Delta^{above}_{Actual} &= \sum_{i=1}^{10} \Delta^i_{Actual} \mathbf{1}_{I(t,i)>I(t+i+10,0)} \\ \Delta^{below}_{GARCH} &= \sum_{i=1}^{10} \Delta^i_{GARCH} \mathbf{1}_{\sigma^2(t,i)<I(t+i+10,0)} \\ \Delta^{above}_{GARCH} &= \sum_{i=1}^{10} \Delta^i_{GARCH} \mathbf{1}_{\sigma^2(t,i)>I(t+i+10,0)} \end{aligned}$$

where and $\mathbf{1}$ is an indicator function.

The measure $\Delta^{below}_{\bullet}$ ($\Delta^{above}_{\bullet}$) is the sum of the absolute percentage deviations below (above) the actual implied volatilities. These measures don't only measure the absolute deviation, but also measures if the forecasts are above or below the actual implied volatilities. The sum of the measure $\Delta^{below}_{\bullet}$ and the measure $\Delta^{above}_{\bullet}$ give the absolute deviation.

9.4 Results

The results are given for the 15 mentioned warrants

- The columns denoted by time to maturity and moneyness are as explained in section 9.2.
- The column named "Observations" indicates the amount of separate tests done in each category of the specific warrant.
- The columns marked less and more are as explained in section 9.3.
- The following abbreviations are used:
  - ITM: In the money
  - ATM: At the money
  - OTM: Out of the money
  - CTM: Close to maturity
  - FFM: Far from maturity

9.4.1 The Results:

Warrant Name: 3ASAIB						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	36	In	2.47	0.05	0.19	0.20
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	83	In	0.59	1.00	0.86	0.85
	0	At	-	-	-	-
	0	Out	-	-	-	-
Description:	The GARCH model predicts ITM, CTM warrants worse than the Actual model does. The GARCH models underpredicts ITM, CTM and overpredicts ITM FFM warrants. The GARCH model predicts ITM, FFM warrants slightly better than the Actual model does.					

Warrant Name: 3ASAUB						
Time to maturity (days)	Observations	Moneyness	GARCH		Actual	
			Below	Above	Below	Above
Less than 70	49	In	0.17	1.86	0	1.61
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	160	In	0.89	0.98	0.92	0.93
	0	At	-	-	-	-
	0	Out	-	-	-	-
Description:	The GARCH model predicts ITM, CTM warrants worse than the Actual model does. Both the GARCH and Actual models overpredicts ATM and OTM CTM warrants. The GARCH model predicts ITM, FFM warrants slightly worse than the Actual model does.					

Warrant Name: 5ASAIB						
Time to maturity (days)	Observations	Moneyness	GARCH		Actual	
			Below	Above	Below	Above
Less than 70	37	In	0.03	5.53	0.01	5.54
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	45	In	0.63	0.70	0.79	0.62
	0	At	-	-	-	-
	0	Out	-	-	-	-
Description:	The GARCH model predicts ITM, CTM warrants slightly worse than the Actual model does. The GARCH model predicts ITM, FFM warrants slightly better than the Actual model does.					

Warrant Name: 2AGLUB						
Time to maturity (days)	Observations	Moneyness	GARCH		Actual	
			Below	Above	Below	Above
Less than 70	0	In	-	-	-	-
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	8	In	1.31	1.05	0	2.52
	90	At	1.26	8.71	0.01	10.36
	5	Out	1.49	7.03	0	6.92

Description:

The GARCH model predicts ITM, FFM warrants worse than the Actual model does.

The GARCH model predicts ATM, FFM warrants better than the Actual model does.

Both models overpredicts ATM and OTM FFM warrants.

The GARCH model predicts OTM, FFM warrants better than the Actual model does.

Warrant Name: 3AGLIB						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	0	In	-	-	-	-
	0	At	-	-	-	-
	50	Out	2.61	2.90	0.03	6.13
70 and Above	0	In	-	-	-	-
	10	At	0	5.28	0	5.68
	43	Out	0.02	14.63	0	14.51

Description:

The GARCH model predicts OTM, CTM warrants better than the Actual model does.

The GARCH model predicts ATM, FFM warrants slghtly better than the Actual model does.

Both models overpredicts ATM and OTM FFM warrants.

The GARCH model predicts OTM, FFM warrants slightly worse than the Actual model does.

Warrant Name: 7AGLIB						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	7	In	3.72	0	3.44	0
	24	At	2.76	0.03	3.04	0.01
	0	Out	-	-	-	-
70 and Above	0	In	-	-	-	-
	68	At	0.31	1.32	0.49	0.98
	6	Out	0.53	0.19	0.65	0.15

Description:

The GARCH model predicts ITM, CTM warrants better than the Actual model does.
The GARCH model predicts ATM, CTM warrants slightly better than the Actual model does.
Both the GARCH and Actual models underpredicts ATM and OTM CTM warrants.
The GARCH model predicts ATM, FFM warrants slightly worse than the Actual model does.
The GARCH model predicts OTM, FFM warrants slightly better than the Actual model does.

Warrant Name: BAGLSG						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	7	In	3.72	0	3.44	0
	24	At	2.76	0.03	3.04	0.01
	0	Out	-	-	-	-
70 and Above	0	In	-	-	-	-
	68	At	0.31	1.32	0.49	0.98
	6	Out	0.53	0.19	0.65	0.15

Description:

The GARCH model predicts ITM, CTM warrants worse than the Actual model does.
The GARCH model predicts ATM, CTM warrants better than the Actual model does.
Both the GARCH and Actual models underpredicts ATM and OTM CTM warrants.
The GARCH model predicts ATM, FFM warrants worse than the Actual model does.
The GARCH model predicts OTM, FFM warrants slightly better than the Actual model does.

Warrant Name: 3NEDUB						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	0	In	-	-	-	-
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	157	In	0.84	0.88	0.65	0.68
	0	At	-	-	-	-
	0	Out	-	-	-	-
Description:	The GARCH model predicts ITM, FFM warrants worse than the Actual model does.					

Warrant Name: 6NEDIB						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	17	In	0.35	0.33	0	0.83
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	45	In	1.50	0.09	1.45	0.13
	0	At	-	-	-	-
	0	Out	-	-	-	-

Description:

The GARCH model predicts ITM, CTM warrants better than the Actual model does.

The GARCH models underpredicts overpredicts ITM, FFM warrants.

The GARCH model predicts ITM, FFM warrants slightly worse than the Actual model does.

Warrant Name: 6NEDSG						
Time to maturity (days)	Observations	Moneyness	GARCH		Actual	
			Below	Above	Below	Above
Less than 70	7	In	0.73	0.11	0.76	0.13
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	14	In	0.47	0.46	0.66	0.31
	0	At	-	-	-	-
	0	Out	-	-	-	-
Description:	The GARCH model predicts ITM, CTM warrants worse than the Actual model does. The GARCH model predicts ITM, FFM warrants slightly better than the Actual model does.					

Warrant Name: 3OMLUB						
Time to maturity (days)	Observations	Moneyness	GARCH		Actual	
			Below	Above	Below	Above
Less than 70	0	In	-	-	-	-
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	118	In	0.18	1.99	0.42	1.28
	0	At	-	-	-	-
	0	Out	-	-	-	-

Description: The GARCH model predicts ITM, FFM warrants worse than the Actual model does.
Both the GARCH and Actual models overpredicts ITM, FFM warrants.

Warrant Name: 4OMLSG						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	3	In	0.73	0.07	1.09	0
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	19	In	2.56	0	2.43	0
	0	At	-	-	-	-
	0	Out	-	-	-	-
Description:	The GARCH model predicts ITM, CTM warrants better than the Actual model does. The GARCH model predicts ITM, FFM warrants worse than the Actual model does. Both the GARCH and Actual models overpredicts ITM, FFM and CTM warrants.					

Warrant Name: 5OMLIB						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	0	In	-	-	-	-
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	41	In	2.13	0.00	2.43	0
	0	At	-	-	-	-
	0	Out	-	-	-	-
Description:	The GARCH model predicts ITM, FFM warrants better than the Actual model does. Both the GARCH and Actual models overpredicts ITM, FFM warrants.					

Warrant Name: 3SAPIB						
Time to maturity (days)	Observations	Moneyness	GARCH		Actual	
			Below	Above	Below	Above
Less than 70	18	In	0.23	0.15	0	0.03
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	103	In	1.31	0.21	0.72	0.35
	0	At	-	-	-	-
	0	Out	-	-	-	-
Description:	The GARCH model predicts ITM, CTM warrants worse than the Actual model does. The GARCH model predicts ITM, FFM warrants worse than the Actual model does. Both the GARCH and Actual models overpredicts ITM, FFM warrants.					

Warrant Name: 3SAPUB						
Time to maturity (days)	**Observations**	**Moneyness**	**GARCH**		**Actual**	
			Below	**Above**	**Below**	**Above**
Less than 70	47	In	2.46	3.26	0	6.21
	0	At	-	-	-	-
	0	Out	-	-	-	-
70 and Above	162	In	3.12	1.87	1.09	2.73
	0	At	-	-	-	-
	0	Out	-	-	-	-

Description:

The GARCH model predicts ITM, CTM warrants better than the Actual model does.
The GARCH model predicts ITM, FFM warrants worse than the Actual model does.
Both the GARCH and Actual models predicts ITM warrants poorly.

9.4.2 Conclusion to Results

In this study, the results due to implied volatility or actual observed market prices performed marginally better than the GARCH prices in the forecasting of market prices of 11 to 20 days in the future.

The forecast due to the actual observed market prices performed marginally better in both time to maturity classes for "in the money" warrants.

The GARCH option pricing forecasts where marginally better for "at the money" warrants with less than 70 days to maturity and "out of the money" warrants with more than 70 days to maturity.

9.4.3 Comments on Study and Results

No specific GARCH or ARMA process can ever be used to fully explain market dynamics. A GARCH process can for instance be useful only in forecasting options on certain assets, in certain market conditions, with a certain range of maturities. Thus plainly put, if an (implied) volatility process follows an approximate GARCH process, then use the GARCH process or option pricing methodology to forecast option prices, if not don't.

A general study, as done here defeats the purpose of GARCH processes to a certain extent, since a GARCH process must be tailor made to the specific market instrument and conditions.

This study does however show that GARCH series can be fitted to implied volatility with some success.

Chapter 10

Conclusion

This dissertation highlights some of the real world deviations from the Black-Scholes option pricing framework.

Unlike the assumption of constant volatility of increments in Brownian motion, volatility in the market is stochastic. Markets with stochastic volatility are no longer complete, as it is in the Black-Scholes structure. Options in incomplete markets are harder to price since investors demand higher returns for taking additional risk.

Duan [10] proposed a new measure under which to price options in incomplete markets, called the Local Risk-Neutral Valuation Relationship (LRNVR). The LRNVR and related option pricing methodology is discussed in detail in this dissertation. The neccesary measure theoretical and stochastic calculus background is given for a clear understanding of this relationship.

The stochastic volatility in this dissertation is assumed to be a statistical time-series process, the Generalized Autoregressive Conditional Heteroscedastic (GARCH) process. Time-series processes are discussed in this dissertation, to give readers who aren't familiar with these statistical methods a reasonable foothold therein.

Warrants are option-like instruments traded on the JSE Exchange. Warrants can't be sold short. This restriction adds to incompleteness in the market. In this dissertation the GARCH option pricing process is applied to the implied volatility of the warrant instead of the stock price process as done by Duan. This is because the standard deviation of the stock price and the implied volatility levels differ significantly because of the short selling restrictions and the illiquidity of the market.

Results of the application of the GARCH option pricing process to implied volatility, shows that it compares well to the use of implied volatility of current warrant prices in forecasting future warrant prices.

Chapter 11

Related Literature

The following advances to the GARCH literature have been published since Duan's 1995 paper:

- Heston and Nandi [22] published a closed-form solution to a GARCH option pricing problem similar to that of Duan's 1995 paper. This method makes use of the conditional moment generating function of the stock price at expiry.

- Duan and Simonato [11] proposed a numerical method for valuing American options with GARCH -like volatility in 2001. This method is based on approximating the underlying asset price process by a finite-state, time-homogeneous Markov chain.

- Ritchken and Trevor [30] in 1999 proposed a lattice approximation scheme for the pricing of GARCH and bivariate stochastic volatility frameworks.

- Duan, Gauthier, Sasseville and Simonato [12] proposed an efficient approach to pricing in the GARCH framework by combining lattice methods and moments approximation in 2002.

Other stochastic volatility option pricing models (see Chriss [8]):

- Implied volatility trees. A model by Derman and Kani and similar models by others. This is a lattice system that use the implied volatility surface of a stock price as input to price an option. This model can also be adapted to price American options.

- Implied binomial trees. A lattice system that uses the implied volatility of European options of all strikes at a fixed expiration date to price nonstandard and exotic options.

Bibliography

[1] Alexander, C., 'Market Models', John Wiley & Sons, 2001.

[2] Bain, L.J. & Engelhardt, M., 'Introduction to probability and Mathematical Statistics', Duxbury Press, 1987.

[3] Bartle, R.G., 'Elements of Integration', John Wiley & Sons, 1966.

[4] Bjork, T., 'Arbitrage Theory in Continuous Time. Oxford University Press', 1999.

[5] Black, F. & Scholes, M. 'The Pricing of Options and Corporate Liabilities', The Journal of Political Economy, 1973.

[6] Bollerslev, T., 'Generalized Autoregressive Conditional Heteroscedasticity', Journal of Econometrics, 1986.

[7] Brzezniak, Z. & Zastawniak T., 'Basic Stochastic Processes', Springer-Verlag, 2000.

[8] Chriss, N.A., 'Black-Scholes and Beyond: Option Pricing Models', McGraw-Hill, 1997.

[9] Crouhy, M.& Galai, D.& Mark, R., 'Risk Management', McGraw-Hill, 1999.

[10] Duan, J.C., 'The GARCH Option Pricing Model', Mathematical Finance, 1995.

[11] Duan, J.C. & Simonato, J.G., 'American option pricing under GARCH by a Markov chain approximation', Journal of Economic Dynamics & Control, 2001.

[12] Duan, J.C. & Gauthier, G. & Sasseville, C. & Simonato, J.G., 'Seize the Moments: Approximating American Option Prices in the GARCH Framework', 2002.

[13] Elliott, R.J. & Kopp P.E., 'Mathematics of Financial Markets', Springer-Verlag, 1999.

[14] Engle, R.F., 'Autoregressive Conditional Heteroscedasticity with estimates of the variance of U.K. inflation, Econometrica.

[15] Engle, R.F. & Ng, V.K., 'Measuring and testing the impact of news on volatility', Journal of Finance, 1993.

[16] Ferreira, M.G., 'Financial time series with heteroscedastic volatility in the South African financial markets', M.Sc. dissertation, University of Pretoria, 1999.

[17] Fouque, J.P. & Papanicolaou, G. & Sicar K.R., 'Derivatives in Financial Markets with Stochastic Volatility', Cambridge University Press, 2000.

[18] Franses, P.H. & Van Dijk, D., 'Nonlinear Time Series Models in Empirical Finance', Cambridge University Press, 2000.

[19] Gerber, H.U. & Pafumi, 'G. Utility function: From Risk Theory to Finance', North American Actuarial Journal.

[20] Greene,W.H., 'Econometric Analysis', Macmillan, 1990.

[21] Haggarty, R., 'Fundamentals of Mathematical Analysis', Addison-Wesley, 1993.

[22] Heston S.L. & Nandi S., 'A Closed-Form GARCH Option Pricing Model', Federal Reserve Bank of Atlanta, www.frbatlanta.org/publica/work_papers/, 1997.

[23] Hull, J.C., 'Options, Futures & Other Derivatives', 4th edition, Prentice-Hall International, 2000.

[24] James, J. and Webber, N., 'Interest Rate Modelling', John Wiley and Sons, 2000.

[25] Klebaner, F., 'Introduction to Stochastic Calculus with Applications', Imperial College Press, 1998.

[26] Lamberton, D. & Lapeyre, B., 'Introduction to Stochastic Calculus Applied to Finance', Chapman and Hall, 1997.

[27] Oksendal, B., 'Stochastic Differential Equations', Springer-Verlag, 1995.

[28] Pliska, S.R., 'Introduction to Mathematical Finance Discrete Time Models', Blackwell, 1999.

[29] Poon, S.H. & Granger, C., 'Forecasting Financial Market Volatility A Review', 2001.

[30] Ritchken, P. & Trevor, R., 'Pricing options under generalized GARCH and stochastic volatility processes', Journal of Finance, 1999.

[31] Rudin, W., 'Real and Complex Analysis ', McGraw-Hill, 1996.

[32] Steele, J.M., 'Stochastic Calculus and Financial Applications', Springer-Verlag, 2001.