TIME SERIES ANALYSIS
AND
MARKET MICROSTRUCTURE ASPECTS
ON SHORT TIME SCALES

Zur Erlangung des akademischen Grades eines Doktors der
Wirtschaftswissenschaften

(Dr. rer. pol.)

von der Fakultät für
Wirtschaftwissenschaften
des Karlsruher Institut für Technologie

genehmigte

Dissertation

von

ALEXANDER BECK

Tag der mündlichen Prüfung: 07.12.2011
Referent: ....Prof. Dr. S.T. Rachev
Korreferent: ....Prof. Dr. M. Feindt

Contents

List of Figures

List of Tables

9

Introduction

This work is a result of a collaboration between the Karlsruhe Institute of Technology's chair of Statistics, Econometrics and Mathematical Finance and Phi-T, a financial research company. In its key role, Phi-T operates as a signal provider for the *Lupus Alpha Neurobayes Short Term Trading Fund*, a hedge fund that relies on complex mathematical trading strategies.

In recent years, an increasing market share of stock trading falls into the responsibility of high-frequency and algorithmic trading (Itai and Sussmann (2009)). In the recent past, especially automated trading was frequently accused of enforcing price shocks at stock exchanges.[1]

As a consequence, increasing academic interest arises in order to understand the complex phenomena and processes that dominate financial markets on short time horizons. Not only are high-frequency phenomena important for the pure theoretical understanding of the working principles of financial markets. Financial firms that trade assets on high-frequency time scales also seek to extend their knowledge about financial processes on short time intervals.

The possibility to analyze the procedures on financial markets on a tick-by-tick basis, that is, every single move in a financial asset's price is taken into account, has only emerged during the past years. Up to that time, the analysis of daily recorded prices (such as daily closing prices) was the most prevalent frequency on which financial data was available. One of the most famous studies on financial data is the work of Mandelbrot (1963), an analysis of cotton prices based on daily records that led to a first discussion of heavy-tailedness in financial data, a phenomenon that is still subject to today's research (Rachev *et al.* (2005); Kim *et al.* (2008a)). The widespread digitalization of the market places around the globe and recent developments in computer technology opened the doors for the availability of seamless streams of financial data.

This work consists of two parts dealing with financial aspects on high-frequency

---

[1]See, for example, "Calls for oversight of automated trading systems", *Financial Times*, March 23, 2011

time scales.[2] The first part is concerned with intraday time series modeling and risk forecasting whereas the second part deals with limit order execution probabilities on short time horizons.

Part I

The development of econometric models that describe financial time series has been subject to research for decades. In particular, conditional mean and variance models were introduced and further developed in numerous studies (see, for example, Engle (1982); Bollerslev (1986); Baillie *et al.* (1996)). Their main target is not only to describe and understand the underlying stochastic processes of asset price developments. They provide a means to forecast future return expectations and their involved volatility.

Standard models in the financial industry include ARMA and GARCH models as well as further developments thereof, such as FIGARCH models. In this thesis, the focus is on the application of such models on high-frequency time series, such as the S&P 500 Index on the shortest possible time intervals. Financial time series features such as heavy tailedness and skewness need to be explicitly considered, because these characteristics appear more intensified on high-frequency data (Sun *et al.* (2008)).

Tempered infinitely divisible distributions are applied as potential candidates to capture high-frequency time series idiosyncrasies in the framework of ARMA-GARCH and FIGARCH models. Their descriptive capabilities are compared to frequently applied innovation distribution assumptions in the financial industry, the normal- and the Student-t distribution.

Risk forecasting is an essential task for financial institutions and portfolio managers (Basel Committee on Banking Supervision (2006)). Tempered infinitely divisible distributions are compared in their risk forecasting capability to the normal- and the Student-t innovation distribution assumption in the framework of ARMA-GARCH models on short time intervals. Here, two common risk measures, Value-at-Risk and Average Value-at-Risk are constructed from these models and compared in a backtest (Beck *et al.* (2011)).

Part II

Order books on a tick-by-tick basis not only offer a way to obtain a much more detailed picture about the continuous procedures in the stock trading process. In

---

[2]Throughout this thesis, the term high-frequency is used in order to denote phenomena on time scales of seconds. It is not used to denote ultra-high-frequency, that is, sub-millisecond trading activity.

their essence, real-time order books provide the possibility to place passive limit orders in a way such that the trading of stocks can be carried out most efficiently in terms of real costs and opportunity costs (Harris and Hasbrouck (1996)). Developers of trading strategies that rely on the two fundamental order types, passive limit orders and aggressive marketable orders, need to have a profound understanding of the prevailing advantages and disadvantages of the respective order type in all market phases. In order to establish trading strategies that rely on these order types, the execution probabilities of passive limit orders need to be quantified accurately (Cho and Nelling (2000)).

However, besides the large amount of information related to limit order executions that is already contained in the order book, the execution behavior of limit orders will in addition depend on quantities reflecting the market sentiment at the present time of the order insertion. In order to deal with this enormous amount of available information in an efficient way, a model is constructed and presented in this thesis, capable of forecasting a limit order's execution on a one minute time horizon conditional on an arbitrarily large set of descriptive variables.

The derivation of this model was made possible by the development of a software package as a part of this thesis. It provides the possibility to construct and analyze limit order books and the limit order flow from financial data streams. Besides its application in the scope of this thesis, it is used in various other research projects at Phi-T.

Part I

Financial Time Series Analysis and Risk Modeling on Short Time Scales

Chapter 1

Introduction

Trading volume in intraday and high-frequency trading has grown strongly over the past few years (Itai and Sussmann (2009); Zhang and Riordan (2011)). This statement is reflected in the market share of high-frequency equity trading, which accounts for 73% in the United States.[1] Along with the increasing market activity in intraday trading comes greater demand for risk assessment on short time scales. This entails research with the aim of providing robust and reliable methods for the quantitative assessment of risk on high-frequency time intervals.

Quantitative risk assessment is a key task in all sorts of financial institutions. A proper risk assessment is not only crucial for classifying assets into risk groups and making investment decisions; it is especially important to understand and quantify risk in order to build sustainable portfolios. Among the most common methods for quantitatively assessing risk are Value-at-Risk (VaR) and Average Value-at-Risk (AVaR). VaR backtesting is considered a standard tool for assessing risk which is required by regulators for determining capital requirements (see, for example, Basel Committee on Banking Supervision (2006)).

Sophisticated mathematical approaches are necessary in order to compute reliable VaR forecasts from historic financial data. Over the past years, a lot of effort has been put into modeling financial time series with the goal of a better understanding of the complex phenomena and structures which are prevailing in financial markets. Conditional mean and variance models have been extensively studied and have become standard tools in financial modeling (Engle (1982); Bollerslev (1986); Nelson (1991); Engle (2001)). Due to the limited scope of financial data available to researchers, most of those studies were carried out on a daily data basis (on daily closing prices, for example). As a result of the digitalization of market places and plummeting costs for data storage facilities, tick-by-tick data has become widely available to researchers. As the name indicates, tick-by-tick

---

[1] See "High-frequency trading under scrutiny", *Financial Times*, July 28, 2009.

data contains every single change in a financial asset's price and offers the most microscopic picture of price movements.

As many studies have shown since, intraday financial time series show a substantially different behavior from time series based on daily recorded data (see, for example, Sun *et al.* (2008)). Prevailing phenomena in financial time series, known as stylized facts, include fat-tails, skewness and volatility clustering. These effects get pronounced in high-frequency time series and are observed along with additional features which typically arise, such as intraday volatility profiles (see, for example, Lee *et al.* (2011)).

Researchers have been attempting to tackle the challenges in high-frequency financial modeling by modifying and extending the standard methods. One essential question is how relative price changes (returns) can be described with respect to the stylized facts on short time scales.

The normal distribution hypothesis for modeling logarithmic returns was frequently applied in financial models. However, the empirical findings persistently rejected this assumption (see, for one of the earliest examples, Mandelbrot (1963)). As a consequence, several alternatives have been proposed and applied in empirical studies (see, for example, Bollerslev (1987); Rachev and Mittnik (2000); Hansen (1994); Rachev *et al.* (2005)). Recently, models based on fat-tailed infinitely divisible distributions have been applied in finance (see, for example, Kim *et al.* (2008c, 2010a) and Kim *et al.* (2008b)). Not only is this type of distribution capable of modeling daily and intra-daily stylized facts, it additionally exhibits desirable attributes such as finite moments.

In this thesis, risk assessment on short time scales is carried out by comparing three different innovation distributions in the framework of conditional mean and variance models. More precisely, the commonly applied normal and Student-t return distribution assumptions are compared against the class of tempered fat-tailed infinitely divisible distributions. It is empirically shown that the latter exhibits an improved modeling and risk forecasting capability over the classical models. The study is carried out on S&P 500 Index data, which is reported in a 15 seconds periodicity during 9:30 - 16:00 Eastern Standard Time.

The remainder of Part I of this thesis is organized as follows. First, the idiosyncratic behaviors of financial time series, especially high-frequency phenomena, are discussed. Financial models with a special focus on conditional mean and variance models are then presented, followed by the innovation distribution assumptions which are relevant in the context of this thesis. The theoretical survey finishes with an introduction to commonly applied risk measures.

Following this, after empirically examining the S&P 500 Index' high-frequency idiosyncracies, the findings of modeling this stock index on short time scales with different time series models are presented. Here, the focus is on ARMA-GARCH and FIGARCH models with different innovation distribution assumptions. Model

parameter dependencies on the return time interval are further investigated with the aim of observing a characteristic scaling behavior. A Value-at-Risk backtest is carried out in which the applied models are compared in their risk forecasting capabilities. Finally, the risk measure Average Value-at-Risk is discussed and compared between the applied time series models.

Chapter 2

Theory of Time Series Modeling and Risk Estimation

2.1 Financial Time Series' Idiosyncracies

Financial Time Series have been subject to researchers in the financial industry for decades. An overview of their idiosyncratic properties is provided in this section. After presenting the most prevalent phenomena in financial time series from daily recorded data (such as closing prices), high-frequency characteristics are further presented and discussed.

Financial time series exhibit properties which seem to be persistently present between different asset classes and markets. Those repeatedly observed phenomena are known as *stylized facts* (see, for example, Cont (2001) and Bouchaud (2002)). A selection of the most prevailing ones is presented in the following.

- *Fat-Tail behavior of return distributions*
  Financial return distributions show a significant excess kurtosis (commonly denoted as fat-tails, an exact definition hereof is given in Appendix A.2). In other words, extreme price movements have a non-negligible likelihood to occur. This phenomenon was first observed by Mandelbrot (1963) and Fama (1965), who demonstrated that cotton price and common stock return distributions possess fatter tails than the normal distribution allows for. The expression *tail-risk* is used to refer to extreme negative price movements which exceed the 3 standard deviation boundary of the underlying return distribution and possess a higher likelihood to occur than expected from a normal distribution assumption.[1]

---

[1] Since the financial markets meltdown in 2008, so-called *tail-risk hedging* was a widely discussed topic among market participants and has even lead to a new class of financial products, see, for example, "Fat-tail attraction", *The Economist*, March 24th, 2011.

- *Skewed return distributions*
  This property refers to the asymmetry which many financial return distributions exhibit.[2] A negative skew means that the distribution's left tail is *longer* than its right tail (i.e., there is more probability mass on the left hand side of the density's center). Negative Skewness is a frequently observed phenomenon in financial time series.

- *Volatility Clustering*
  Volatility clustering denotes the phenomenon that, if volatility is high, it is likely to remain high. On the contrary, if it is low, it is expected to stay low. The more general phenomenon *heteroskedasticity* (White (1980)) denotes changing variance over time (in contrast to homoskedasticity) and is a prevailing feature in financial time series.

It is crucial in financial modeling and risk forecasting to find models that are capable of properly describing the observed stylized facts. Omitting them may potentially lead to unreliable and oversimplified models which may potentially lead to implausible figures in risk assessment or other applications.

Especially the observation of fat-tailed and skewed return distributions leads to the assumption that the normal distribution is generally not well suited in order to model financial processes. In fact, many researchers have rejected the normal hypothesis in their studies (see, for example, Mandelbrot (1963)).

2.1.1   Intra-daily Data and High Frequency Data

The term *high-frequency* is used by researchers in a broad range of meanings. Generally, it is used in order to distinguish data that are recorded on intra-daily frequencies from data that are recorded on daily, weekly, monthly or yearly periods. For example, Bollerslev *et al.* (2006) denotes 5 min return time series as high-frequency data. Engle (2000) uses the term *ultra-high-frequency* in order to describe the maximum possible level of disaggregation in financial time series, such as provided by tick-by-tick data.[3] Tick-by-tick data offers the most microscopic picture of the price process. Especially, if order book information is avail-

---

[2]The skewness of a return distribution is reflected in its third central moment. The sample skewness is constructed as

$$\hat{\beta}_3 = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n}\sum_{i=j}^{n}(x_j - \overline{x})^3}{\left(\frac{1}{n}\sum_{i=j}^{n}(x_j - \overline{x})^2\right)^{3/2}} \qquad (2.1)$$

[3]The term "tick-by-tick" is used to refer to records which contain every price move for a financial asset. Such data is generally available for electronic exchanges and some OTC trading platforms.

able on a tick-by-tick basis, a large number of time series can be constructed and investigated in addition to the commonly applied examination of log-return time series.[4] This depth of information offers great research opportunities, which has lead to a new field of research known as high-frequency-econometrics. Goodhart and O'Hara (1997) summarize important features of intra-daily data and discuss issues and methodologies which have to be considered when high-frequency data is employed and analyzed.

2.1.2 Stylized Facts on Intra-Daily Data

After having examined stylized facts on daily time series, this section focuses on phenomena observed on intra-daily time series. The most important stylized facts observed on intra-daily data are listed below (see, as a reference, Sun *et al.* (2008) and Goodhart and O'Hara (1997).

- *Pronounced Volatility Clustering*
  Volatility clustering, which is also observed on low-frequency financial time series, is intensified in high-frequency data.

- *Autocorrelation*
  In contrast to data which is aggregated over relatively long time periods (such as daily), high-frequency data may exhibit significant autocorrelation (Cont (2009)). On the most disaggregated level, this can be explained through the *Bid-Ask-Bounce*, a micro-structure effect which leads to negative first order correlation (Roll (1984)). The minimum price difference between subsequently executed sell- and buy initiated trades is in the dimension of the bid-ask spread. Under the assumption that buy and sell initiated trades occur randomly, the observed last transaction price fluctuates between best bid and best ask price, leading to a negative first order autocorrelation.

- *Strong heavy tails*
  The heavy tail property of unconditional return distributions in intra-daily data is even more pronounced than in daily data.

- *Random Durations*
  High frequency data from single assets will usually be unequally spaced in time (Engle and Russell (1998)). This results from the market participants' independent behavior, i.e., their actions do not underlie any fixed

---

[4]A list of examples could contain time series on the bid-ask-spread, the size of order book levels, asymmetries in the level size between the bid- and the ask-side, etc.

time frame and occur randomly distributed over time. This is an important feature which has to be considered in time series modeling since many models rely on equally spaced time series.

Time series on indices (such as the S&P 500 Index) don't show such irregularities because they are reported on fixed time schedules.

- *Long Range Dependency*

  Long range dependency (LRD) in time series has first been described by Hurst (Hurst (1951)). It denotes persistence in time series, i.e. non-negligible autocorrelation between instances over long lags. LRD is observed if the autocorrelation coefficients γ for a discrete time series $\{x_t,\ t = 1...T\}$ satisfy $\sum_{k=0}^{\infty} \gamma_k \to \infty$, where k denotes the lag (Racheva-Iotova and Samorodnitsky (2011)). Typically, the γ_k decay hyperbolically in this case. On the contrary, if the γ_k decay exponentially, LRD is non-existent. For financial time series, it is widely assumed that LRD effects increase with increasing frequency over which market data is collected (Sun *et al.* (2008)).

  LRD is observed in return time series as well as in time series of squared returns (Cont (2011)). The latter indicates that volatility changes at earlier times $t-i, i = 1...N_l$, where N_l denotes the number of lags, will have a non-negligible influence on the volatility at time t. To be more precise, changes in volatility will have a long lasting impact on future volatility observations.

- *Intraday Volatility Profile*

  Intra-daily data typically show a U-like volatility pattern (see, for example, Lee *et al.* (2011)). This periodic pattern is comparable to seasonality effects, as, for example, observed in daily data with repeating patterns over long time spans. Intraday volatility is especially high during the morning hours, due to the overnight information flow and the resulting market movements. It is lowest around lunch time and increases again towards the end of the trading day. Figure 2.1 shows a typical intraday volatility profile.[5],[6]

2.2 Financial Econometrics

Financial Econometrics denotes the mathematical and empirical examination of economic data. One import sub-field within econometrics is financial time series analysis. As opposed to the laws of nature, findings in the field of finance

---

[5]This profile has been constructed from 15 seconds S&P 500 Index log-returns over a period of 10 months (April, 2010 to January, 2011)

[6]This chart-type is known as *profile plot* and is explained in Appendix A.1.
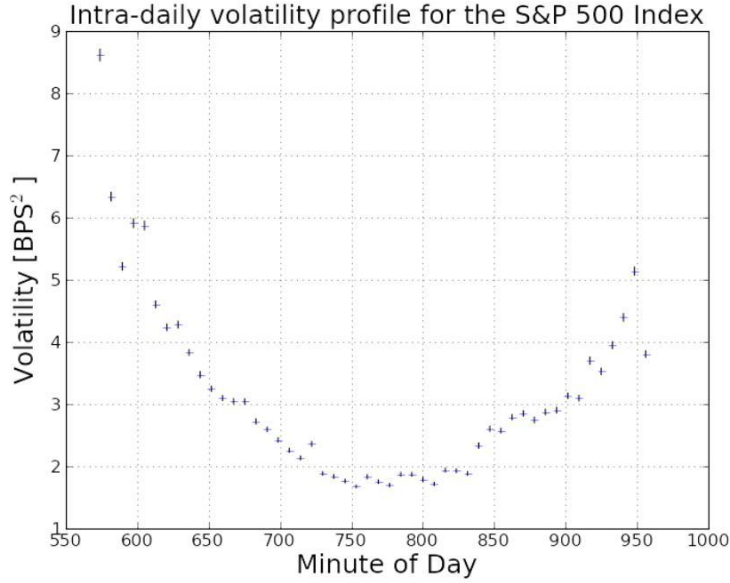
Figure 2.1: Intra-daily volatility profile for the S&P 500 index on 15 second log-returns. The squared returns serve as a proxy for the unobserved volatility in this chart.

do not have the intrinsic property of remaining valid over time. As a consequence, financial models are subject to continuous revision and re-adoption.

One of the main targets of financial time series analysis is to find mathematical descriptions for price processes. The term *Data Generating Process* is used in order to denote a mathematical formulation, depending on variables at earlier times $t' < t$, which explains the realization of a stock price at a time t, denoted as S_t.

Most financial time series studies focus on the *net return* or the *log-return* of a financial asset between two subsequent points of time, t and $t-1$. The net return is given by

$$R = \frac{S_t - S_{t-1}}{S_t}.$$ (2.2)

It describes relative price changes between times t and $t-1$. The log-return is given by

$$r_t = \ln \frac{S_t}{S_{t-1}} \approx \frac{S_t}{S_{t-1}} - 1 = \frac{S_t - S_{t-1}}{S_t},$$ (2.3)

where the Taylor expansion is valid if $S_t/S_{t-1} \approx 1$. This is especially true on short time horizons where asset prices do not show large price movements. The log-return is usually preferred in practical applications because it has two great ad-

vantages over the net return. First, subsequent log-returns can simply be summed up to construct the log-return over a longer time period, that is: $r_T = \sum_{t=1}^{T} r_t$. Second, when modeling returns with probability distributions that allow for negative values (such as the normal distribution), using the log-return ensures that the stock price S_t naturally remains positive at all times (this is not the case with the net return R).[7] The log-return can be motivated from another point of view. If continuous compounding is assumed between two points of time t and $t-1$, the rate r at which the asset price grows in continuous time is given by $S_t = S_{t-1} \exp(r)$. The log-return is applied throughout this thesis. It is therefore also simply denoted as *return*.

2.2.1 Randomness in the Data Generating Process

Randomness plays a crucial role in financial time series analysis. Consequently, the theory of *stochastic processes* is an important element in time series modeling. A stochastic process is a sequence of random variables $\{x_t, t \in T\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$. If the parameter set $\{T\}$ is countable, i.e. $T \in \mathbb{N}$, the resulting process is denoted as *discrete process* whereas a *time continuous process* results from $T \in \mathbb{R}^{\geq 0}$. In the case of time series, the parameter t denotes the time and x_t the state of the process at time t. A stochastic process is denoted as *real-valued*, if $x_t \in \mathrm{R}$. In practical examples, such as stock prices, only one *realization* or *trajectory* of a stochastic process is given. The mean and covariance function of a real-valued process with $t \geq 0$ and the additional property $\mathbb{E}[x_t^2] < \infty$ are given by

$$
\begin{aligned}
\mu_x(t) &= \mathbb{E}[x_t] \\
\gamma_k &= cov(x_t, x_{(t-k)}),
\end{aligned}
\tag{2.4}
$$

and, as a natural consequence, the unconditional variance is given by

$$
\sigma^2 = cov(x_t, x_t) = \gamma_0.
\tag{2.5}
$$

The *autocorrelation function*, that is, the linear correlation of the process with itself at *lag k*, is given by

$$
\rho_k = \frac{\gamma_k}{\gamma_0}.
\tag{2.6}
$$

As a reference, see, for example Rachev *et al.* (2007).

---

[7]Consider the net return R modeled with, for example, a normal distribution. Consequently, there is certain likelihood that the asset price will become negative.

Lévy Processes

Lévy processes are a type of time continuous stochastic processes which have independent and identical (IID) increments, start at zero and are right continuous with left limits. Independence indicates that the value of an increment at one point of time t_1 is in no way related to other values at times t_i. Additionally, the probability distribution of the increments must have the property of *infinite divisibility*.

A real-valued random variable x is said to have an *infinitely divisible* distribution, if, for any integer number n there exists a sequence of independently and identically distributed (IID) variables such that the sum

$$s_n \stackrel{d}{=} x_{n1} + x_{n2} + x_{n3} + ... + x_{nn} \tag{2.7}$$

is distributed like the x_{ni}. That is, the variable x can be decomposed into a sum of n equally distributed variables. The stable property implies infinite divisibility.

The IID property indicates that the process is *stationary* in its increments, that is, sequences of increments within a time series are equally distributed. This is expressed as

$$(d_t, d_{t+1}, ..., d_{t+n}) \stackrel{d}{=} (d_{t+h}, d_{t+h+1}, ..., d_{t+h+n}), \tag{2.8}$$

for any $h \in \mathbb{N}$, where d_t denotes the increment at time t. In practical applications, stationarity is often used to refer to stationarity with regard to the first and second order moments of the increments' distribution only. That is,

$$\mathbb{E}[x_t] = \mu \quad \forall t \tag{2.9}$$
$$cov(x_t, x_{t-h}) = \gamma_h \quad \forall t, h. \tag{2.10}$$

This form of stationarity is commonly referred to as *weak stationarity*. Throughout this thesis, the term *stationary* is used in order to refer to weak stationarity.

The Wiener Process The *Wiener Process*, a member of the class of Lévy-Processes, is an important example of time-continuous stochastic processes. It is given by the conditions:

1. $W_0 = 0$

2. W_t is continuous

3. W_t has independent increments with $W_t - W_s \sim \mathcal{N}(0, t - s)$,

where $\mathcal{N}(0, t - s)$ denotes a normal distribution with a mean of zero and a variance of $t - s$. A stochastic process based on the Wiener Process with *drift* μ and *volatility* σ^2 is given by

$$x_t = \mu + x_{t-1} + \sigma^2 z_t, \quad z_t \sim \mathcal{N}(0, 1). \tag{2.11}$$

The Wiener Process fulfills the *Markov property*. That is, the conditional probability distribution of the following realization depends only on the present state x_{t-1}. In mathematical terms, this is expressed as

$$F_{t|t-1, t-2, \dots, t-k}(x_t | x_{t-1}, x_{t-2}, \dots, x_{t-k}) = F_{t, t-1}(x_t | x_{t-1}), \ k \geq 1 \tag{2.12}$$

where F_t denotes the conditional probability distribution for the state of the process at time t.

The Wiener Process with $\mu = 0$ additionally fulfills the martingale property, expressed as

$$\mathbb{E}[x_t | x_{t-1}, x_{t-2}, \dots, x_0] = x_{t-1}. \tag{2.13}$$

It states that the conditional expectation value of the future state is equal to the current state of the process. This is the definition of a fair game.

The Wiener Process is a frequently applied model in research concerning financial time series. Even though it does not take properties such as time varying volatility, autocorrelation and other stylized effects into account, it is still a helpful tool in order to understand fundamental characteristics of data generating processes. Note that the Wiener Process itself is not stationary. In order to achieve a stationary time series from a Wiener Process, one frequently applied measure in financial time series analysis is, to: (1) de-trend the process $x' = x_t - \mu$; (2) construct a time series of net- or log-returns. The stationarity property is one additional important reason for the utilization of the log-return in time series modeling instead of the asset price.

2.2.2   Linear Time Series Models

In order to address the above mentioned features such as autocorrelation and dependence, models which are more complex than the Wiener Process are considered. Especially if the markov property is not assumed to be fulfilled, that is, future returns may depend on past realizations, dynamic models are applied which take earlier returns into account. Very popular linear time series models include *Autoregressive* (AR) and *Moving Average* (MA) models as well as their combination, denoted as *ARMA* models. They can be applied to discrete, equally spaced and stationary time series in order to estimate the conditional mean of future return distributions.

The AR model of order p (commonly denoted as AR(p)), describes a data generating process as

$$
\begin{aligned}
y_t &= \sum_{i=1}^{p} a_i y_{t-i} + \epsilon_t \\
\epsilon_t &= \sigma z_t, \quad z_t \sim IID(0,1)
\end{aligned}
\tag{2.14}
$$

where the a_i are the parameters of the model. ϵ_t denotes a disturbance term which introduces randomness to the process. It consists of a random variable drawn from a distribution with zero mean ($E_{t-1}[x_t] = 0$) and unit variance ($Var_{t-1}[x_t] = 1$). $E_{t-1}[\cdot]$ denotes the conditional expectation value whereas $Var_{t-1}[\cdot]$ stands for the conditional variance, both at time $t - 1$. The unconditional variance σ of the process is used to scale the innovation z_t.

To summarize, y_t is described as a weighted sum over the past p realizations ($y_{t-1}, ... y_{t-p}$) plus an IID distributed disturbance term.

Similarly, MA processes of order q are defined as

$$
y_t = \sum_{i=0}^{q} b_{t-i} \epsilon_{t-i},
\tag{2.15}
$$

where y_t consists of a weighted sum of the past q disturbances $\epsilon_{t-1}, ..., \epsilon_{t-q}$, plus the disturbance term at time t: $b_0 \epsilon_t$. The distribution of the disturbances ϵ is defined in the same way as for the AR process. Generally, it is assumed that $b_0 = 1$.

The mixture of AR and MA processes results in an ARMA(p,q) process which is given by

$$
y_t = a_1 y_{t-1} + a_2 y_{t-2} + ... + a_p y_{t-p} + \epsilon_t + b_1 \epsilon_{t-1} + b_2 y_{t-2} + ... + b_q y_{t-q}.
\tag{2.16}
$$

With the definition of a lag-operator, $L(y_t) \equiv y_{t-1}$, the above expression can be written in the more convenient form:

$$
a(L) y_t = b(L) \epsilon_t,
\tag{2.17}
$$

where $a(L) = \sum_{i=1}^{p} a_i L^i$ (the same definition holds for $b(L)$). Compared to pure AR and MA models, ARMA models have the advantage of requiring less parameters (a stationary ARMA process has an MA(∞) representation). This is a crucial issue for parameter estimations.

Unit Roots in Time Series

A time series $\{x_t, \ t = 1...T\}$ with an autoregressive representation, given by

$$
x_t = \sum_{i=1}^{p} \alpha_i L^i x_t + \epsilon_t,
\tag{2.18}
$$

with $\epsilon_t = \sigma z_t$, where $z_t \sim IID(0,1)$, can be written in the more compact form

$$
\begin{aligned}
\phi(L)x_t &= \epsilon_t \\
\phi(\lambda) &= 1 - \sum_{i=1}^{p} \lambda^i \alpha_i.
\end{aligned}
\tag{2.19}
$$

The *characteristic polynomial* of an autoregressive process is given by

$$
1 - \alpha_1 \lambda - \alpha_2 \lambda^2 - ... - \alpha_p \lambda^p = 0.
\tag{2.20}
$$

The process is stationary, if the characteristic polynomial's roots lie outside the unit circle. Otherwise, the process exhibits a *unit root* and is not stationary. The process is then said to be *integrated*. The order of integration corresponds to the multiplicity of the unit root. Note that the roots of $\phi(\lambda)$ are in general complex numbers. A stochastic process is not stationary if it contains a unit root and, as a consequence, ARMA models can not be applied without differencing the process in a preliminary step. As an alternative, ARIMA models can be applied. One commonly applied method for testing a time series on unit roots is the Dickey-Fuller-Test (Dickey and Fuller (1979)).

ARIMA Models

A more general representation are ARIMA (Integrated Autoregressive Moving Average) processes which can be applied to non-stationary time series. In this case, the difference between subsequent elements of the stochastic process is constructed before the above describe ARMA model is applied. The original process is reconstructed by integrating the resulting ARMA process.

There is a variety of modifications on ARMA models in order to capture additional effects and to add additional information. For example, FARIMA models are applied to time series which exhibit long range dependency (Hosking (1981); Granger and Joyeux (1980)). Exogenous information, for example a time series on interest rates, is considered in so-called ARMAX models, in which the exogenous variables are regressed to the target variable.

2.2.3 ARCH/GARCH Models

As stated earlier, one frequently observed phenomenon in financial time series is volatility clustering. This characteristic is encompassed in the more general feature of *heteroskedasticity* in financial time series. Heteroskedasticity denotes changing variance over time. Engle (1982) introduced a new class of stochastic processes in order to capture time varying volatility and heteroskedasticity. These

autoregressive conditional heteroskedastic processes are known as ARCH models and the generalization hereof, introduced by Bollerslev (1986), is known as GARCH models.

The ARCH(r) model, as introduced by Engle (1982), describes the ϵ_t generating process as follows:

$$\epsilon_t = \sqrt{h_t} z_t, \quad z_t \sim IID(0,1) \tag{2.21}$$

$$h_t = w_0 + \sum_{i=1}^{r} w_i \epsilon_{t-i}^2, \tag{2.22}$$

where the w_i are the parameters of the model, which must satisfy $w_0 > 0$ and $w_i \geq 0$ for $i = 1, ..., r$. The function h_t denotes the conditional variance of ϵ_t, constructed from a weighted sum over the squared past r disturbances, $\epsilon_{t-1}^2, ..., \epsilon_{t-r}^2$. h_t denotes the *innovation* at time t and is assumed to be IID distributed over time, with mean zero and unit variance. The normal distribution assumption is in most cases not appropriate for financial time series modeling. Typical distributions assumptions are discussed in a later section of this thesis.

From Equation 2.22 it appears that volatility clustering can be captured with the ARCH model: If the past errors $\epsilon_{t-i}, i = 1, ..., r$ have been large, this will be reflected in a correspondingly large conditional variance h_t. This in turn implies that ϵ_t is expected to be large as well. Note that the sign of ϵ_t is not determined.

Note that the ARCH process with conditionally normal distributed errors leads to a leptokurtic unconditional distribution (i.e. a heavy tailed unconditional distribution). As a reference see, for example, Bollerslev (1987) and Bollerslev *et al.* (1992). However, as Bai *et al.* (2003) point out, this implied unconditional kurtosis turns out to be too small in order to reconcile the sample kurtosis, which is detected in financial time series. As as consequence, non-normal, leptokurtic distributions appear inevitable in order to properly model the innovation process.

Bollerslev (1986) extended the ARCH model by appending a weighted sum over the past s conditional variances h_{t_i} for $i = 1, ..., s$. The resulting model is known as *generalized autoregressive conditional heteroskedastic model*, or in short, GARCH-model. Hence,

$$\epsilon_t = \sqrt{h_t} z_t, \quad z_t \sim IID(0,1) \tag{2.23}$$

$$h_t = w_0 + \sum_{i=1}^{r} w_i \epsilon_{t-i}^2 + \sum_{j=1}^{s} v_j h_{t-j}, \tag{2.24}$$

where, again, h_t denotes the variance conditional on information until time t, $\mathcal{I}_{t-1}$. In order for the process ϵ_t to be stationary within the GARCH-model framework, the parameters w_i and v_i must satisfy the necessary and sufficient condi-

tions: $\sum_{i=1}^{r} w_i + \sum_{i=1}^{s} v_i < 1$ and $w_i > 0, v_j \ \forall i \in \{0, ...r\}$ and $j \in \{1, ..., s\}$ (Bougerol (1992)).

In order to apply a GARCH model to a financial time series, it has to be ensured that the conditional return distributions are centered around zero. This can either be achieved by de-trending the complete time series or by applying a conditional mean model (e.g. an ARMA model).

The conditional variance process may exhibit persistence effects as well. To capture this phenomenon, Engle and Bollerslev (1986) introduced the Integrated GARCH (IGARCH) process. A GARCH process is *integrated*, if $\sum_i w_i + \sum_i \nu_i = 1$. Here, the current information remains important on all future time horizons. As a consequence, volatility shocks have a permanent impact on the future conditional variance.

2.2.4 FIGARCH Models

Following the development of GARCH (Bollerslev (1986)) and IGARCH (Engle and Bollerslev (1986)) models, (Baillie *et al.* (1996)) introduced FIGARCH (Fractionally Integrated Generalized Autoregressive Conditional Heteroskedastic) processes, which are capable of describing time series exhibiting persistence in squared log-returns. They are useful tools when the correlations of subsequent log-returns in a time series appear negligible but the time series' volatility exhibits long range dependency.

The FIGARCH(p,d,q) process for $\{\epsilon_t\}$ is defined by

$$
\begin{aligned}
\phi(L)(1-L)^d \epsilon_t^2 &= w + [1 - \beta(L)]\nu_t, \quad \nu_t = \epsilon_t^2 - h_t^2, \\
\epsilon_t &= h_t z_t, \quad z_t \sim IID(0,1),
\end{aligned}
\tag{2.25}
$$

where $0 \leq d \leq 1$ as well as $E_{t-1}[z_t] = 0$ and $Var_{t-1}[z_t] = 1$. For $d = 0$, the above process corresponds to the well-known GARCH process whereas for $d = 1$, Equation 2.25 describes an IGARCH(1) process (So and Yu (2006)). The calculation of the $(1-L)^d$ term has to be approximated in practical applications. To this end, the term is Taylor-expanded around $L = 0$, leading to

$$
\begin{aligned}
(1-L)^d &= \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} L^k \\
&= 1 - dL + \frac{(1-d)(-d)}{2} L^2 + ...
\end{aligned}
\tag{2.26}
$$

where $\Gamma(\cdot)$ denotes the gamma function. Usually, the sum in Equation 2.26 is calculated up to the lag corresponding to $k = 1000$ (Baillie *et al.* (1996); So and

Yu (2006)). For large k, applying Sterling's formula to $\Gamma(k-d)/\Gamma(k+1)$ leads to

$$\Gamma(k-d)/\Gamma(k+1) \approx k^{-d-1}, \tag{2.27}$$

indicating a hyperbolic decay in the polynomials coefficients. This, in turn, implies that the effect of past innovations decays hyperbolically in the FIGARCH model, given $0 < d < 1$. In the case of $d = 0$, which corresponds to a GARCH(p,q) model, the influence of past innovations decays exponentially, indicating no long range dependency effects in financial market volatility.

2.2.5 ARMA-GARCH Models

In order to connect conditional mean and conditional variance models, ARMA and GARCH models are often used in combination. The joint ARMA-GARCH model is defined as

$$y_t = \sum_{i=1}^{p} a_i y_{t-i} + \sum_{i=1}^{q} b_i \epsilon_{t-1} + \epsilon_t \tag{2.28}$$

$$\epsilon_t = \sqrt{h_t} z_t, \quad z_t \sim IID(0,1) \tag{2.29}$$

$$h_t = w_0 + \sum_{i=1}^{r} w_i \epsilon_{t-1}^2 + \sum_{i=1}^{s} v_i h_{t-1}, \tag{2.30}$$

where the *innovation process* z_t is assumed to be IID, with zero mean and unit variance, i.e. $E_{t-1}[z_t] = 0$ and $Var_{t-1}[z_t] = 1$.

2.3 Model Choice and Validation

2.3.1 Detecting the Autocorrelation and Dependence Structure

Before selecting a model to describe a time series, it is reasonable to investigate the time series' autocorrelation and independence structure. Two important methods which allow to visually detect such patterns are the Sample Autocorrelation Function (SACF) and the Sample Partial Autocorrelation Function (SPACF). The respective structure of those two functions provide a means to estimate the lag order and to motivate the employment of ARMA and/or GARCH models. The SACF is defined as follows:

$$\hat{\rho}_k = corr(y_t, y_{t-k}) = \frac{cov(y_t, y_{t-k})}{\hat{\gamma}_0}, \quad k = 0, 1, \dots \tag{2.31}$$

where $\hat{\gamma}_0$ is the unconditional variance of the time series: $\hat{\gamma}_0 = Var(y_t)$. $\hat{\rho}_k$ denotes the sample correlation coefficient between y_t and y_{t-k}. The SPACF is constructed as

$$\hat{\alpha}_k = corr(y_t, y_{t-k}|y_{t-1}, y_{t-k+1}). \quad k = 1, 2, ... \quad (2.32)$$

The SPACF describes the correlation between y_t and y_{t-k} after eliminating the linear dependence between the realizations y_{t-1} through y_{t-k+1}. In simple words, the correlation between y_t and y_{t-k} is measured after removing the influence of the realizations which occur intermediately.

When applying pure AR or MA models, these two methods can be directly used in order to find the appropriate number of model parameters. In more general cases, the SACF and SPACF are useful tools in order to detect non-stationarity and long-range-dependence (LRD) effects in time series. LRD is indicated if the SACF is declining slower than exponentially.

In addition to investigate a time series for autocorrelation, the dependence structure between single instances may be of interest as well. One frequently employed measure is the SACF constructed from a time series of squared returns. SACF values which are significantly different from zero indicate that the squared returns are correlated and suggests the application of GARCH models. If the SACF of squared returns exhibits persistent behavior, that is, the correlation between squared returns is not declining exponentially with increasing lag, long range dependency is indicated which motivates the utilization of FIGARCH models.

2.3.2 Testing Time Series on ARCH/GARCH Effects

Engle (1982) proposed a Lagrange multiplier (LM) test in order to examine if a time series contains ARCH(q) disturbances. The null hypothesis in Engle's ARCH test is $w_1 = w_2 = ... = w_q = 0$, which is tested against the alternative hypothesis of $\{\epsilon_t^2\}$ following an ARCH(q) process.

A differentiable function $h(\cdot)$ is assumed which satisfies $h_t = h(z_t w)$, where $z_t = (1, \hat{\epsilon}_{t-1}^2, ..., \hat{\epsilon}_{t-q}^2)$. ϵ_t are the residuals after subtracting the mean process, and $w = (w_0, w_1, ..., w_q)'$ is the set of ARCH-parameters. Under the null hypothesis h_t is a constant, denoted as h_0.

The test statistic is constructed as

$$\xi = \frac{1}{2} f^{0'} z (z'z)^{-1'} z' f^0, \quad (2.33)$$

where $z' = (z_1', ..., z_T')$ and $f^0 = \left(\frac{\hat{\epsilon}_t^2}{h_0} - 1 \right)$ with $h_0 = \frac{1}{T} \sum_{t=1}^{T} \hat{\epsilon}_t^2$.

As Breusch and Pagan (1979) and Godfrey (1978) point out, the above test statistic does not depend on $h(\cdot)$. This is a necessary condition for the application of this test if the residuals are not assumed to be conditionally normal distributed. The test statistic is asymptotically χ_q-distributed.

Engle's ARCH-test is usually implemented in standard mathematical software such as MATLAB.

2.3.3 Identifying the Lag Order

The identification of the lag order determines the number of free parameters in a parametric model. In order to obtain a model with a good descriptive nature in addition to reliable and robust forecasting qualities, it is quintessential to properly adjust the set of free model parameters.

Simply speaking, a model with too many parameters will be able to capture every random fluctuation in the data. In this undesired case, the model will perfectly describe the underlying time series but with a complete loss of the ability to generalize and to produce reliable forecasts. This phenomenon is known as *overfitting*. As a consequence, the number of parameters should be as limited as possible.

There is a variety of selection criteria available in order to find the appropriate set free parameters. In the case of ARMA(p,q) models, the most conservative procedure is the Bayesian Information Criterion (BIC), proposed by Schwarz (1978), which is defined as

$$BIC_{p,q} = \ln \hat{\sigma}_{p,q} + \frac{\ln T}{T}(p+q), \tag{2.34}$$

where $\hat{\sigma}_{p,q} = \frac{1}{T}\sum_{i=1}^{T}\hat{\epsilon}_t^2(p,q)$ is the mean variance of the residuals. As Equation 2.34 shows, the BIC scales down with decreasing mean variances $\hat{\sigma}_{p,q}$ and scales up with an increasing number of parameters p and q. The best estimate for the number of free parameters under this criterion is received for the tuple $\{p,q\}$ that minimizes $BIC_{p,q}$.

2.3.4 Estimation of the Model Parameters

After a model has been selected its parameters have to be estimated. There are two common procedures in order to perform the parameter estimation, Least-Squares-Estimation (LSE) and Maximum-Likelihood-Estimation (MLE). This section focuses on the MLE, which has been used throughout this thesis.[8]

---

[8]Detailed information of the Least-Squares Estimation method can be found in a variety of textbooks about statistics or econometrics, such as in Rachev *et al.* (2007).

Let $\{x_t,\ t = 1...T\}$ be a series of observed values. The underlying conditional distribution function $f(x|\mathbf{a})$ is assumed to be known, where $\mathbf{a}$ denotes a set of parameters with unknown values.[9] The purpose of the Likelihood estimation is to find the unknown parameter set $\mathbf{a}$ which maximizes the likelihood of observing the values $\{x_t\}$.

The Likelihood function $\mathcal{L}$ is defined as

$$\mathcal{L}(\mathbf{a}) = \prod_{i=t}^{T} f(x_t|\mathbf{a}). \tag{2.35}$$

It is a measure for the probability to observe the given values $\{x_t\}$ for the parameter set $\mathbf{a}$. By optimizing this measure, is,

$$\mathcal{L}(\mathbf{a}) \rightarrow \mathrm{maximize}, \tag{2.36}$$

the parameter set $\mathbf{a}$ is chosen under which the observation of $\{x_t\}$ becomes most likely.

In practical applications, $-\log \mathcal{L}$ is minimized instead of $\mathcal{L}$ being maximized. Due to the application of the log-function, the likelihood function decays into a sum over all data points $\{x_t\}$. This approach is generally more practical than dealing with the product over many elements, especially when the product could lead to values extremely close to zero. A minimization is carried out instead of a maximization due to common conventions by which optimization algorithms operate.

In general, multi-parameter log-likelihood functions can be highly nonlinear and non-continuous. This complicates the optimization problem and classical algorithms, such as the gradient descent or the Newton-Raphson method, will lead to unsatisfying results since those methods will typically only find the next local minimum. To this end, sophisticated algorithms are needed in order to achieve reliable results. Nevertheless, it has to be noted that up to this date, no algorithm is known which could guarantee to find the global minimum.

When estimating financial models where the actual conditional innovation distribution is not normal, the Normal-Maximum-Likelihood-Estimation may still be used. The resulting parameter estimation is known as Quasi-Maximum-Likelihood-Estimation (QMLE) (see, as a reference, Bollerslev and Woolridge (1992)).

2.4 The Innovation Process

The innovation process has to be capable of taking the stylized facts, as presented in Section 2.1, into account.

---

[9]It is important to note that $\int_{\infty}^{\infty} f(x|\mathbf{a})\mathrm{d}x = 1$ must hold for all sets of $\mathbf{a}$.

At first appearance, a stable distribution is desirable. For this proposition can be argued from the point of view that the behavior of market participants, which eventually determine the price (hence, the return) by their actions, are influenced from a variety of quantitative and qualitative variables. The *central limit theorem*, a fundamental principle in statistics, states that a sum of independent random variables, drawn from different distributions, will converge towards a stable distribution. In the special (and desirable) case that those distributions possess a finite expectation value and a finite variance, the distribution of their sum will converge towards a normal distribution.[10]

Another strong argument for the stable hypothesis is based on the construction of log-returns. A log-return over a period of one day, can be constructed from minutely observed log-returns by simply summing them up:

$$r_d = \sum_{i=0}^{N} r_i^m, \tag{2.37}$$

where N is the number of trading minutes withing one trading day and r_i^m is the return of minute i. Consequently, if minutely returns r_i^m were independent and exhibiting finite variance, the daily log-return distribution should be consistent with the normal distribution assumption, following again the rules stated by the central limit theorem. The normal probability density function is given by

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{2.38}$$

It's two parameters, μ and σ determine the location and the width of the distribution. The normal distribution is employed in a wide scope of applications for several reasons. First, the normal distribution is readily implemented in basically every available statistical software package and there exist numerous ways of assessing models which are based on the normal distribution. In addition, it has been used by financial institutions for a long time, and as a consequence, its idiosyncratic behavior and its shortcomings are well known and well documented.

Even though this way of reasoning seems conclusive, empirical studies have strongly rejected the normal distribution hypothesis in various cases. This is not only true for unconditional return distributions but also for conditional distributions as obtained in the framework of ARMA-GARCH models. Excess kurtosis and skewness are still prevailing characteristics in empirical innovation distributions which are extracted from financial time series.

This persistent question has been tackled by researchers by applying various probability functions to empirically observed innovation distributions. Some of

---

[10]To be more precise, the *Lindenberg Condition* (Lindenberg (1921)) has to be satisfied.

the most important theoretical distribution assumptions will be discussed in this section. Their application in ARMA-GARCH models has been extensively studied and is well documented.

2.4.1 α-stable Distributions

In order to adhere to the stability approach as outlined above, the α-stable distribution family should be well-suited candidates for modeling the innovation process, as Mandelbrot (1963) proposed. Even though α-stable distributions are very adaptive and therefore capable of modeling excess kurtosis and skewness effects, they have the severe drawback of not possessing finite moments (the only exception is the normal distribution). This leads to fundamental difficulties when portfolio selection criteria are considered in which asset prices and the involved risk rely on finite moments of the underlying distributions, such as Harry Markowitz' modern portfolio theory (Markowitz (1952)) or the Black-Scholes equation for option pricing (Black and Scholes (1973)).[11]

Stable distributions are attractors in the framework of the central limit theorem. That is, as mentioned above, a sum of IID random variables, drawn from different distributions, will converge towards a stable distribution.

A distribution is defined as stable, if for two independent random variables, X_1 and X_2, drawn from the same distribution X, the following is true:

$$aX_1 + bX_2 \stackrel{d}{=} c \cdot X + d, \tag{2.39}$$

for some real constants a, b, c and d.

The α-stable probability density function has no closed form expression in the return space. It is defined in the frequency space by its characteristic function as

$$
\begin{aligned}
\phi_{\text{stable}}(u; \alpha, \sigma, \beta, \mu) &= \mathbb{E}[e^{iuX}] \\
&= \begin{cases} \exp\left(i\mu u - |\sigma u|^\alpha \left(1 - i\beta \operatorname{sign}(u) \tan \frac{\pi\alpha}{2}\right)\right), & \alpha \neq 1 \\ \exp\left(i\mu u - |\sigma u|^\alpha \left(1 + i\beta \frac{2}{\pi} \operatorname{sign}(u) \ln(|u|)\right)\right), & \alpha = 1 \end{cases}
\end{aligned}
$$

Its parameters α, σ, β and μ have the following interpretations and limits:

- $\mu \in (-\infty, \infty)$ is the localization parameter.

---

[11]Even though booth theories are based on the assumption of normally distributed returns their key statements regarding risk assessment based on second order moments are still valid in a wide scope.

- $\sigma \in (0, \infty)$ determines the width of the distribution.

- $\beta \in [-1, 1]$ describes the distribution's skewness.

- $\alpha < 2$ defines the asymptotic tail behavior.

The normal distribution is the special case for $\alpha = 2$ and $\beta = 0$. All α-stable distributions besides the normal one exhibit excess kurtosis and skewness (if $\beta \neq 0$).

Benoit Mandelbrot (Mandelbrot (1963)) first employed α-stable distributions in the context of financial modeling in order to describe empirical distributions based on cotton prices.

GARCH models with α-stable innovations were introduced by Rachev and Menn (Menn (2005) and Rachev and Mittnik (2000)).

2.4.2 Student-*t* Distribution

Due to the drawbacks and difficulties which come along with the use of the α-stable distribution family, the Student-*t* distribution is widely used in financial modeling. Even though this distribution is symmetric, that is, it is not capable of describing skewness, it exhibits fat-tail behavior.

The Student-*t* probability density function is defined as follows:

$$f(x; \nu) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\nu \pi} \frac{\nu}{2}} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \tag{2.40}$$

In this form, the Student-*t* distribution has only one parameter, namely, the number degrees of freedom (DOF), denoted by ν in the above equation.

It's second order moment, the variance, is given by

$$Var[X] = \frac{\nu}{\nu - 2}, \quad \nu > 2. \tag{2.41}$$

That is, the variance is only defined for $\nu > 2$. The Student-*t* distribution interpolates between the normal distribution (for $\nu \to \infty$) and the Cauchy-Distribution (for $\nu = 1$).

In order to apply the Student-*t* distribution to real world applications, such as financial markets modeling, a frequently used modification of Equation 2.40 is the non-central three parameter version:

$$f(x; \mu, \lambda, \nu) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{\nu}{2})} \sqrt{\frac{\lambda}{\nu \pi}} \left(1 + \frac{\lambda (x - \mu)^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \tag{2.42}$$

In this three parameter form, the Student-*t* distribution is scalable in width and can be located around the empirical distribution's observed mean.

2.4.3 Tempered Stable Distributions

Classical Tempered Stable (CTS) distributions are capable of describing both skewness and excess kurtosis. In contrast to α-stable distributions (except the normal distribution), CTS distributions possess finite moments. This quality is crucial for the employment of these distributions in asset pricing models which rely on finite first and second order moments. CTS distributions are derived from α-stable distributions with the essential difference that their tails fade out exponentially, as Figure 2.2 illustrates. Tempered stable distributions possess, like α-stable distribution, the infinite-divisibility property.

Tempered stable distributions have first been introduced by Koponen (1995) as *Truncated Levy Flights*. They are also known as KoBol distributions (Boyarchenko and Levendorskiǐ (2000)). In the same way as for the α-stable distribution, the tempered stable distribution family has no closed form expression in the return space. It is defined through its characteristic function, which is given by

$$
\begin{aligned}
\ln \phi(u; \alpha, C_+, C_-, \lambda_+, \lambda_-, m) \;=\; \\
ium \;-\; & iu\Gamma(1-\alpha)(C_+\lambda_+^{\alpha-1} - C_-\lambda_-^{\alpha-1}) \\
+\; & C_+\Gamma(-\alpha)((\lambda_+ - iu)^\alpha - \lambda_+^\alpha) \\
+\; & C_-\Gamma(-\alpha)((\lambda_- + iu)^\alpha - \lambda_-^\alpha). \quad (2.43)
\end{aligned}
$$

In the case when $C = C_+ = C_- > 0$, as throughout this thesis, these distributions are also known as CGMY distributions (Carr *et al.* (2000)).

For modeling time series, it is convenient to employ the standardized classical tempered stable distribution (stdCTS), with zero mean and unit variance. Due to the predefined model's mean and variance, the parameter set is reduced from six to three parameters (under the additional assumption that $C = C_+ = C_- > 0$). Since the probability density's representation in the return space has to be constructed by employing a Fourier transform on the characteristic function ϕ, the fixed second order moment is particularly advantageous for technical reasons, (see Scherer *et al.* (2010)).

The standardized characteristic function of tempered stable distributions is given by

$$
\begin{aligned}
\log \phi_{stdCTS}(u) = \frac{2iu}{\alpha - 1} \cdot \frac{\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}}{\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}} \\
+ \frac{1}{(a-1)a} \cdot \frac{(\lambda_+ - iu)^\alpha - \lambda_+^\alpha + (\lambda_- + iu)^\alpha - (\lambda_-)^\alpha}{\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}}
\end{aligned}
\quad (2.44)
$$

where

$$
\begin{aligned}
C &= C_+ = C_- = \left[\Gamma(2-\alpha)\left(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}\right)\right]^{-1} \\
m &= 0
\end{aligned}
$$
(2.45)

The interpretation and range of the parameters α, λ_+ and λ_- is given in the following enumeration:

- α determines the asymptotic tail behavior and the distribution's shape around the center region symmetrically.

- λ_- defines how the left hand tail decays.

- λ_+ defines how the right hand tail decays.

In combination, λ_+ and λ_- define the distribution's skewness behavior. $\lambda_- > \lambda_+$ will result in a negatively skewed distribution, that is, more probability mass will be located on the left hand side of the distribution's center. In order to improve the robustness of parameter estimations, the CTS-function can be re-parametrized. The re-parametrization affects the parameters λ_+ and λ_-. They are transformed as follows:

$$
\begin{aligned}
\lambda_+ &= \lambda - \frac{\Delta}{2} \\
\lambda_- &= \lambda + \frac{\Delta}{2}
\end{aligned}
$$

$$
\begin{aligned}
\longrightarrow \quad \lambda &= \frac{1}{2}(\lambda_- + \lambda_+), \quad \lambda > 0 \\
\Delta &= \lambda_- - \lambda_+, \quad \Delta < 2\lambda
\end{aligned}
$$

The interpretation of the new parameters λ and Δ is straightforward: λ describes the degree of symmetric tempering and Δ influences the skewness. If Δ is hold at zero, changing the parameter λ will result in symmetrically changing the shape of the density function. Vice versa, holding λ constant and changing the parameter Δ from negative to positive values will result in the skewness of the density function moving from negative values to positive values. The interpretation of α remains unchanged. How differences between λ_+ and λ_- affect the distribution's skewness is depicted in Figure 2.3.

Note that the degree of skewness depends on all three parameters, α, λ and Δ. That is, Δ itself only determines the direction, in which the probability density function is skewed, but not exclusively how pronounced the skewness is. This can be shown by considering the cumulants $c_n(X)$ of a tempered stable random variable X, which are given by
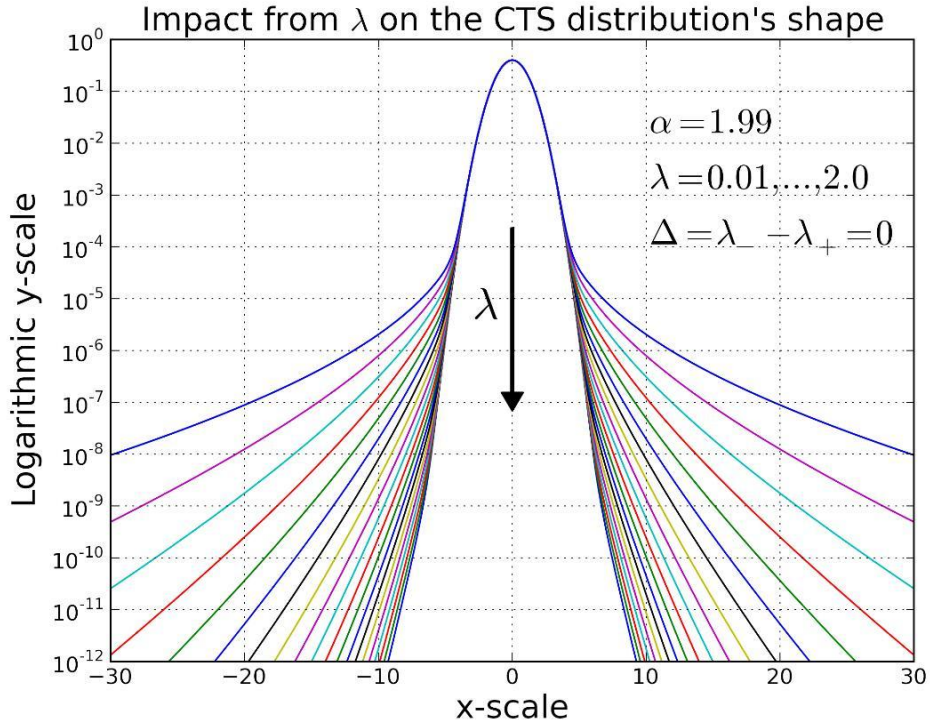
Figure 2.2: The above figure demonstrates the impact on changes in the parameter $\lambda = \lambda_+ = \lambda_-$. Increasing λ leads to a pronounced tempering effect. The y-log-scale has been chosen to demonstrate the differences in the tail behavior.

$$\begin{aligned} c_n(X) &= \Gamma(n-\alpha)C_1\lambda_+^{\alpha-n} + (-1)^n\Gamma(n-\alpha)C_2\lambda_-^{\alpha-n}, \text{ for } n \in \mathbb{N}, \ n \geq 2 \\ c_1(X) &= \mu + \Gamma(1-\alpha)C_1\lambda_+^{\alpha-1} - \Gamma(1-\alpha)C_2\lambda_-^{\alpha-1}. \end{aligned} \qquad (2.46)$$

Since the third cumulant c_3 corresponds to the third central moment (Abramowitz and Stegun (1965)), it is clear that $\Delta = \lambda_- - \lambda_+$ itself can only indicate the direction of the skewness.

Figure 2.2 illustrates how changing the parameter λ, as defined in Equation 2.46, affects the shape of the CTS distribution. The parameters α and Δ are held constant. The approximately linear tail decay in this chart (with a logarithmic y-scale) indicates an exponential decay on a linear y-scale. Note that the point at which the tail behavior starts to change also depends slightly on λ.

For $\alpha \to 2$ and $\lambda_{+/-} \to \infty$, the CTS distribution converges towards a normal distribution.

CTS innovations have been applied in financial modeling by, for example,
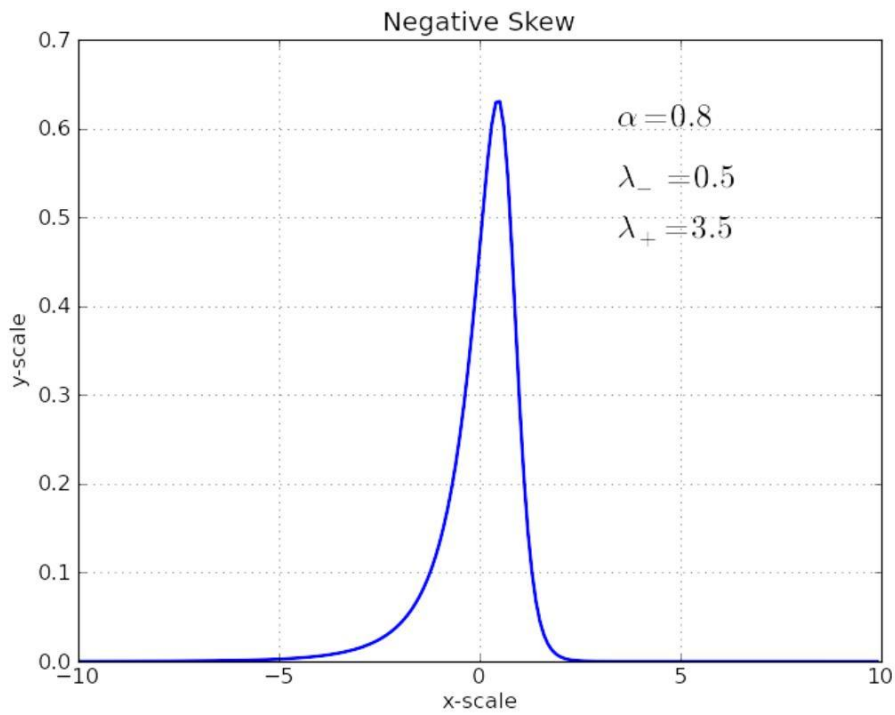
(a) Negative Skewness
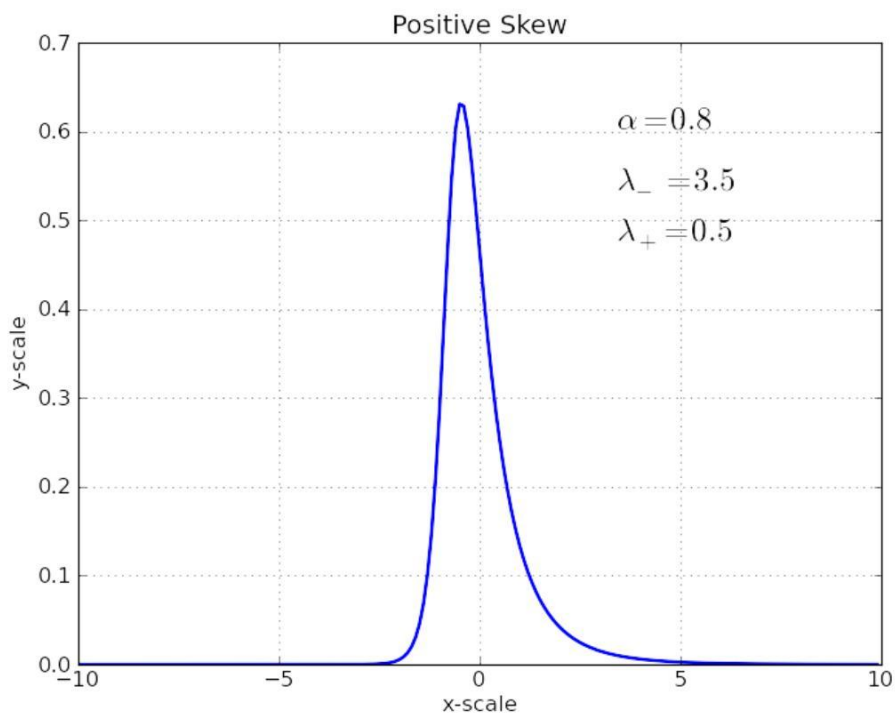
(b) Positive Skewness

Figure 2.3: The impact from differences in λ_+ and λ_- on the CTS-distribution's skewness. $\lambda_- > \lambda_+ \rightarrow \Delta > 0$ leads to a positive skewness (and vice versa).

Kim *et al.* (2008a,c), Kim *et al.* (2010b) and Beck *et al.* (2011). Comprehensive studies on tempered stable distributions and their applications in financial markets can be found in Kim *et al.* (2010c) and Bianchi (2009).

2.5 Goodness of Fit Tests

After a model has been chosen and applied to the data (i.e., the model's parameters have been estimated) it is crucial to verify the model's descriptive power. This includes especially a goodness-of-fit examination for the employed innovation distribution assumption. To this end, several statistical tools are available. The proposed goodness of fit tests will result in a p-value which has to be compared to a pre-defined confidence level, typically 95%, 99% or 99.9%. The hypothesis that the considered model is consistent with the observed data is usually denoted as *Null Hypothesis* (H_0). If the resulting p-value falls below the pre-defined level, the Null Hypothesis is rejected, otherwise, it is accepted.

A general issue which arises when using statistical tests are two specific error types:

Type I Error: Rejecting the Null Hypothesis when it is actually true.

Type II Error: Accepting the Null Hypothesis when it is actually false.

The choice of the confidence level is a means to find a suitable compromise between Type I and Type II errors. It is not possible to rule out both errors at the same time. Being conservative in the acceptance of a model will reduce the probability of Type II Errors but at the same time increase the probability of Type I errors and vice versa. The 95% confidence level is rather conservative, thus reducing the probability of Type II Errors whereas the 99.9% confidence level reduces the probability of Type I Errors.

The goodness-of-fit tests which have been used in this thesis are presented in this section.

2.5.1 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (KS) test is a nonparametric test that assess if two cumulative probability density functions (cdf) result from identical distributions within statistically allowed boundaries (Massey (1951)).

The Null Hypothesis is defined as

$$H_0 : F(x) = \hat{F}(x), \tag{2.47}$$

where $F(x)$ denotes the empirical and $\hat{F}(x)$ the theoretical distribution function. The alternative hypothesis can be formulated as

$$H_1 : F(x) \neq \hat{F}(x). \tag{2.48}$$

The maximum difference between the two functions is used as test statistic and is constructed by

$$KS = \sqrt{n} \cdot \sup_{x_i} |F(x_i) - \hat{F}(x_i)|, \tag{2.49}$$

where $\hat{F}(x_i)$ denotes the empirical and $F(x_i)$ the theoretical cumulated distribution function (cdf).

The probability of KS being smaller than some x is computed with the Kolmogorov-Smirnov cumulative distribution function, given by

$$Pr(KS \leq x) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 x^2} = \frac{\sqrt{2\pi}}{x} \sum_{i=1}^{\infty} e^{-(2i-1)^2 \pi^2/(8x^2)}. \tag{2.50}$$

In addition to the above formula, tabulated values are available in, for example, Miller (1956).

The resulting p-value is compared against the confidence level which is specified before the test is carried out. The test then either accepts or rejects the Null hypothesis.

2.5.2 Anderson-Darling Test

The Anderson-Darling (AD) test is carried out by following a similar procedure as compared to the KS test but with a modified test statistic (Anderson and Darling (1954)). The way in which the KS test statistic is constructed, the corresponding x-values will typically be found close to the center region of the distribution under investigation. That is, the KS test focuses on the goodness-of-fit in the center region. In cases, where proper tail modeling is essential, the AD-test statistic may be better suited for this purpose, since it focuses on the goodness-of-fit in the tail regions. The AD test statistic is given by (Stephens (1974))

$$A^2 = -n - \sum_{k=1}^{n} \frac{2k-1}{n} \left[\ln(z_k) + \ln(1 - z_k) \right]. \tag{2.51}$$

Here, the z_k result from transforming the ordered observed values x_k into a uniform distribution by applying the theoretical distribution function F. That is, $z_k = F(x_k)$.

The construction of the Anderson-Darling test's related p-values is not as straightforward as it is for the KS test. It can be constructed by following a procedure developed by Marsaglia and Marsaglia (2004).

2.5.3 Quantile-Quantile-Plot

Quantile-Quantile-Plots (QQ-Plots) are a statistical tool to visually detect if two distributions correspond each other. QQ-Plots for two sample distributions result from drawing the quantile positions for the two ordered samples against each other. In those cases where the two distributions are statistically identical, the corresponding quantile positions must coincide as well. The result is a QQ-Plot with all quantile pairs on one diagonal line. Deviations from the diagonal indicate that the two distributions do not coincide. The QQ-Plot allows to visually locate the quantiles in which the distributions do not coincide.

Figure 2.4 illustrates a QQ-Plot, where the normal distribution hypothesis is tested for its applicability on S&P 500 daily returns over the past ten years (2000-2010). The blue bullets deviate clearly from the diagonal which is indicated as a dashed red line, leading to the conclusion that the normal distribution is not well suited to model S&P 500 daily return distributions properly. Applying the QQ-Plot method allows to visually analyze a model's performance and especially to detect regions in which the empirical data is not well described.

2.6 Risk Measures Used by Financial Institutions

Financial Institutions which underlie the Basel Accords are obliged to determine their capital requirements by following the rules specified in the Basel II Accords,[12] a set of guidelines published by the Basel Committee of Banking Supervision. Firms holding large amounts of risky assets must follow these specified rules to assess the cumulated risk they are holding on their balance sheets. The general idea behind those rules is to control the risk to which financial institutions are exposed and, as a consequence, lower the risk of bankruptcies. According to the indicators received from quantitative risk estimations, firms in the financial industry are required to hold an adequate amount of capital. The Basel II Accord recommends the use of Value-at-Risk (VaR) methods in order to quantitatively estimate the risk in equity exposures (Basel Committee on Banking Supervision, 2006)(Part 2.III.H.11.ii.§527)).

2.6.1 Coherent Risk Measures

Risk measures used in order to determine capital requirements should satisfy a set of axioms, as suggested by Artzner *et al.* (1999). The future uncertainty of a portfolio $\mathcal{A}$ is usually described by a function $X : \Omega \to \mathbb{R}$, where Ω denotes a finite set of scenarios. In portfolio analysis, X will usually denote the value of a

---

[12]The development of Basel III is still in progress, see Moody's Analytics (2011).
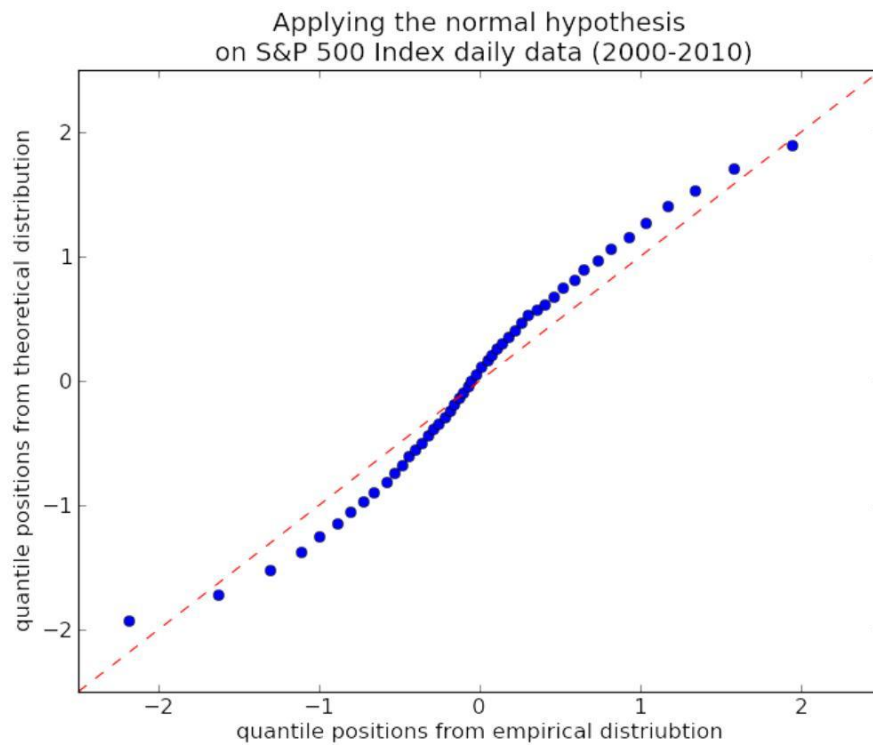
Figure 2.4: Application of a normal distribution to S&P 500 Index daily returns over a time period of ten years (2000-2010). The above QQ-Plot shows that the quantile positions of the empirical distribution and the normal assumption do not coincide. This indicates that the normal hypothesis is not well suited to model the empirical return distribution properly.

given portfolio which is constructed by a linear combination of functions taking into account each element in Ω. A function $\rho(X)$ is used to quantify the risk and cash requirements which have to be added to the portfolio. The linear space of functions $X : \Omega \to \mathbb{R}$ is denoted as $\mathcal{X}$.

In order to be defined as a coherent risk measure, $\rho : \mathcal{X} \to \mathbb{R}$ has to satisfy the following requirements, with $Y, Z \in \mathcal{X}$:

- *Monotonicity:* If $Y \leq Z$, then $\rho(Y) \geq \rho(Z)$.
  If portfolio Y leads to worse results than portfolio Z in all scenarios, the risk measure should be greater for Y.

- *Sub-additivity:* $\rho(Y + Z) \leq \rho(Y) + \rho(Z)$.
  If portfolios Y and Z are put together, the combined risk measure must not be greater than the sum of the separate risk measures.

- *Positive Homogeneity:* For $\lambda \geq 0$, $\quad \rho(\lambda Y) = \lambda \rho(Y)$.
  By changing all positions in a portfolio by the factor λ while keeping the relative proportions constant, the risk measure must change by the same factor λ.

- *Translation Invariance:* If $C \in \mathbb{R}$, then $\rho(Y + C) = \rho(Y) - C$.
  By adding the amount C of cash to the portfolio, the risk measure, hence the capital requirements, must decrease by the same amount.

2.6.2 Value-at-Risk

One of the most commonly applied risk measures by financial institutions is Value-at-Risk (VaR), which has been introduced by J.P. Morgan, an investment bank (Longerstaey and More (1995)). It depends on two parameters, a time horizon T and a confidence level η. With VaR, financial institutions have a feeling of what amount of money they are at least about to loose within a certain time period T at a given certainty level η. In simple terms, VaR is the mathematical description of the following formulation (cited from Hull (2007)):

"We are X percent certain that we will not loose more than V dollars in time T."

For example, assume for a portfolio $\mathcal{A}$ a VaR($\eta = 5\%$,$T = 1$ day) figure of 10 Million Dollars. This means that portfolio $\mathcal{A}$ is expected to loose an amount of at least 10 Million Dollars on one out of 20 days.

In mathematical terms, the definition of VaR at a confidence level of $\eta\%$ for a random variable X is given by

$$\text{VaR}(X)_\eta = -\inf\{x \in \mathbb{R} | P(X \leq x) > \eta\}. \qquad (2.52)$$
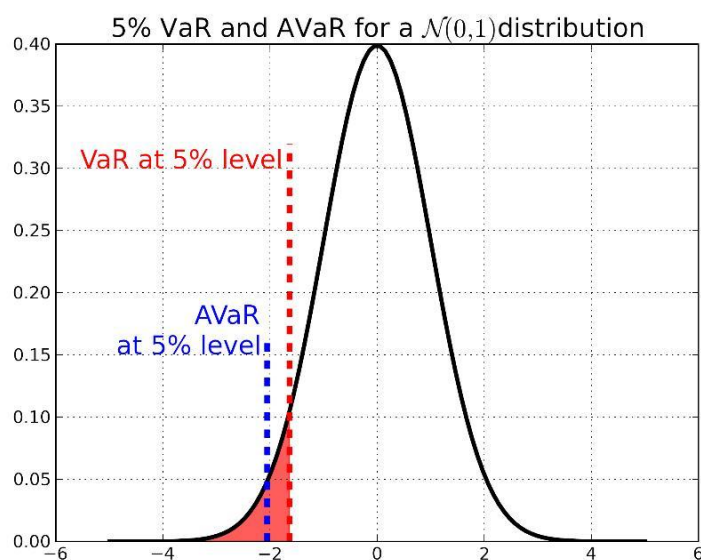
Figure 2.5: Visualization of the 5% VaR and AVaR levels for an exemplary $\mathcal{N}(0,1)$ return distribution.

Figure 2.5 illustrates the location of the 5% VaR level for a standardized normal distribution $\mathcal{N}(0,1)$. In the more general case of normal distributions $\mathcal{N}(0, \sigma^2)$, this level corresponds to -1.64σ, where σ denotes the standard deviation.

Value-at-Risk figures for a portfolio can be constructed in several ways. Common methods include historical simulation, monte-carlo methods or variance-covariance based approaches (Hull *et al.* (1998)). VaR figures can also be determined in combination with conditional mean and variance models as carried out in this thesis.

Even though Value-at-Risk is a frequently used risk measure by financial institutions, it's application must be considered under the awareness that it exhibits some severe drawbacks. Following the definition in Section 2.6.1, VaR is not a coherent risk measure since the sub-additivity requirement is not generally satisfied.[13]

A risk measure, which satisfies all four criteria for being coherent is Average Value-at-Risk (Acerbi and Tasche (2001)).

---

[13]Examples to demonstrate this fact can be easily constructed. See, for example, Hull (2007).

2.6.3 Average Value-at-Risk

Average Value-at-Risk (AVaR), also known as Conditional Value-at-Risk (CVaR), Expected Shortfall (ES) or Expected Tail Loss (ETL), is a risk measure which can be used as an alternative to Value-at-Risk. Instead of just determining the minimum expected loss, as in Value-at-Risk, it is applied in order to calculate the expected value of losses over a given period of time.

The definition of AVaR at the confidence level η is given by

$$\text{AVaR}_{\eta}(X) = \frac{1}{\eta} \int_0^{\eta} \text{VaR}_{\epsilon}(X)\, d\epsilon, \qquad (2.53)$$

where $\text{VaR}_{\epsilon}(X)$ denotes the Value-at-Risk figure for a random variable X at the confidence level ϵ. Figure 2.5 shows the location of AVaR for a $\mathcal{N}(0,1)$ distribution at the $\eta = 5\%$ confidence level.

Besides being a coherent risk measure, AVaR has one further significant advantage over VaR: By using VaR in estimating monetary losses, the only information received from this measure is the amount of money, which will at least be lost for a certain confidence level η in a given time period. VaR does not indicate what would happen in the worst possible scenarios. This information is provided by Average Value-at-Risk since it is the weighted mean over all losses exceeding the pre-defined confidence level.

Chapter 3

Empirical Analysis of High Frequency Data

An empirical analysis of time series constructed from S&P 500 Index data on high-frequency time scales is presented in this chapter.[1] First, the dataset used for this study is introduced. The transformation of the data which is applied to remove the intraday volatility profile is described and empirically observed stylized facts are reported. Subsequently, ARMA-GARCH and FIGARCH models with different innovation distribution assumptions are compared in their modeling capabilities. The special case of tempered infinitely divisible innovation distributions is investigated further to find dependencies between the corresponding parameter values and the frequency on which log-returns are constructed. The forecasting performance of the respective models is investigated in a Value-at-Risk backtest.

The empirical analysis was carried out by employing MATLAB, Python (Choirat and Seri (2009)) and ROOT.[2]

3.1   The Dataset - S&P 500 Index

The S&P 500 Index (short for Standard&Poor's 500 Index) is a North American stock market index which consists of the 500 biggest U.S. companies in accordance to their market capitalization. It is generally regarded as being representative for the whole U.S. economy and therefore widely used in portfolio optimizations. The S&P 500 Index is published by Standard&Poor's, a U.S. based financial services company. It is maintained by the S&P 500 Index Committee

---

[1]This chapter is to a great extent based on the publication Beck *et al.* (2011) which resulted from the empirical analysis carried out in the context of this thesis.

[2]ROOT is a data analysis framework developed by scientists at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland.
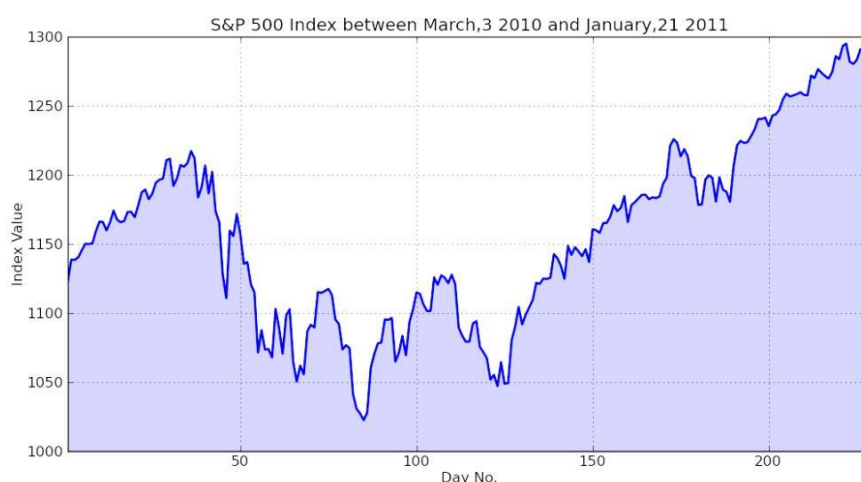
Figure 3.1: Illustration of the S&P 500 Index during the analyzed time period from March 3, 2010 to January 27, 2011.

who's task is to ensure that the index reflects the risk and return characteristics for the U.S. large cap market segment. To be regarded as a potential candidate for being taken into the S&P 500 Index, companies need to fulfill a number of specified requirements. Likewise, companies can be removed from the S&P 500 Index if they violate certain criteria.[3]

The S&P 500 Index' value is constructed by applying a weighted sum of each company's market capitalization. The weight used in the index' construction is proportional to the number of publicly available shares. To have the index' time series free of distortions resulting from company removals, rights issues, issuances etc., the index' value is re-adjusted after such events.

The S&P 500 Index belongs to the group of price return indices. That is, dividend payments are not reinvested, as it is the case for total return indices. It was first published in 1957 and is nowadays reported every 15 seconds between 9:30 and 16:00 Eastern Standard Time.

The dataset used in this study covers 15 second index values from March 3, 2010 to January 27, 2011 (227 trading days). Figure 3.1 depicts the S&P 500 Index over this time period with daily closing prices.

If gaps occur in the data, the missing values are obtained by linear interpolation (see, for example, Sun *et al.* (2008)). Taking the criteria for excluding days into account,[4] the final dataset contains 166 trading days (that is roughly

---

[3]A detailed overview of these rules can be found at http://www.indices.standardandpoors.com.

[4]Some dates are excluded from the study where either the data was completely missing or large gaps occured in the dataset. The exclusion criteria for one day are fulfilled if either more than 3%

300 000 index values, which corresponds to about 1200 years of daily recorded data). From these data, a time series of log-returns, $r_t = \log S_t/S_{t-1}$, is constructed by concatenating the respective time series of subsequent trading days. To avoid overnight effects, the first return from each daily time series is cut off. Subsequently, time series with returns on different time-intervals, ranging from 15 seconds to 300 seconds in steps of 15 seconds, are created. In the following, the time intervals over which returns are calculated are denoted as δt_r.

3.2 Transforming the Data by Removing the Volatility Profile

Intraday return time series exhibit a slightly U-shaped volatility pattern, similar to the well-known shape in intraday trading-volume (Fraenkle (2010)). That is, the volatility is generally higher in the opening and closing periods compared to the time around noon. Before ARMA-GARCH models can be applied to the intraday return series, the intraday volatility pattern, which can be thought of as a seasonality effect on a daily basis, must be removed. Such a pattern cannot be captured by applying an ARCH or a GARCH approach. On the contrary, if this pattern is ignored, it could potentially destroy an underlying GARCH process as demonstrated by Andersen and Bollerslev (1997).

The time series on different return time intervals δt_r have to be treated individually. For each of these time series, every trading day's volatility profile has to be forecasted.

For this purpose, Bollerslev *et al.* (2006) employ the sum over past squared returns for each intraday time interval as a proxy for the unobserved volatility. In this thesis, a slightly modified approach is applied. The volatility $v_{d0,i}$ for return r_i on day d_0 is estimated by the following expression:

$$v_{d_0,i}^2 = \frac{\sum_{k=1}^{N} r_{i,d_{-k}}^2 \cdot \exp(-\tau \cdot k)}{\sum_{k=1}^{N} \exp(-\tau \cdot k)}, \tag{3.1}$$

where $r_{i,d_{-k}}^2$ is a proxy for the unobserved volatility[5] belonging to the ith return on the kth day before day d_0. Due to this modification, the resulting volatility measure will depend more heavily on those previous days, which are close to the day under investigation. By this, the volatility measure adjusts more quickly to persistent changes in the volatility profile's structure and magnitude. In this

---

of the values are missing or if gaps occured which exceeded 1% of the expected number of values for one day.

[5]Volatility itself cannot be observed. A common proxy for the unobserved volatility is the squared return or absolute return, see, for example, Lee *et al.* (2011).

exponential smoothing model, the parameters $\tau = 0.1$ and $N = 10$ are used.[6] The weights on the past days are such that the older days are weighted less than the days that are closer to the day under investigation. The parameter set as described above will give the 10th day in the past a relative weight of $1/e$.

Subsequently, all returns are transformed by $\hat{r}_{d,i} = \frac{r_{d,i}}{v_{d,i}}$, where $\hat{r}_{d,i}$ denotes the transformed return and $r_{d,i}$ denotes the ith observed return on day d in a time series with timespan δt_r. $v_{d,i}$ is the forecasted volatility for day d and is applied to the ith return. This transformation is performed for each return time interval δt_r. The result are regularly spaced time series on different time scales δt_r which are free of intraday volatility profiles.

Figures 3.2a to 3.2c depict the transformed time series for the δt_r = 15, 75 and 300 seconds return time interval with their corresponding return distributions. In these figures, $\hat{r}$ denotes the transformed return, that is, the return with removed volatility profile. The following analysis are carried out on log-return time series with removed volatility profiles.

3.2.1 Volatility Scaling Behavior

The volatility estimators defined in Equation 3.1 are further utilized in order to investigate the volatility scaling behavior. Figure 3.3 shows the average estimated volatility in dependence of the return time interval in a *profile plot*.[7] The blue curve in this figure depicts a square-root model whereas the red curve shows a power law model with an estimated power of 0.58 ± 0.001. Figure 3.3 demonstrates that the empirically observed volatility scaling structure exhibits slight deviations from the square-root model. In the framework of a Wiener Process, the volatility would exhibit a square-root dependency on the return time interval.

3.3 Stylized Facts in High Frequency Data

It is essential to have a profound understanding of the prevailing features of a given dataset in order to select and apply time series models appropriately. Section 2.1.2 presents an overview of stylized facts which are observed in intra-daily and high-frequency data. In this section, a comprehensive analysis for these stylized facts is carried out on the dataset described above.

---

[6]$N = 10$ is chosen such that a time horizon of two business weeks is used to forecast the volatility. $\tau = 0.1$ has been chosen such that the 10th day still has a non-vanishing influence on the volatility.

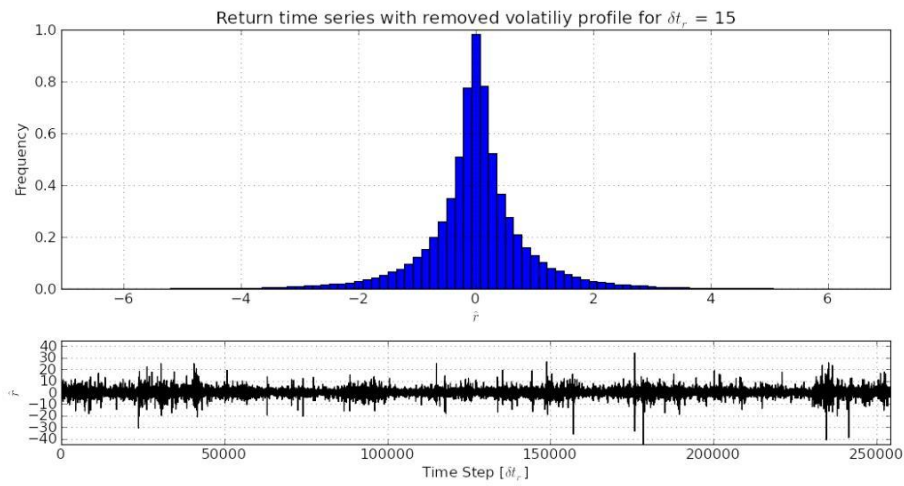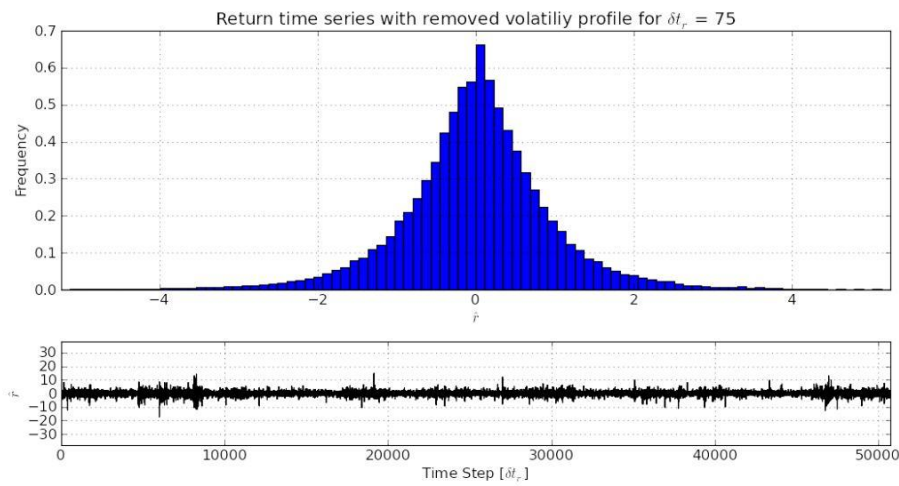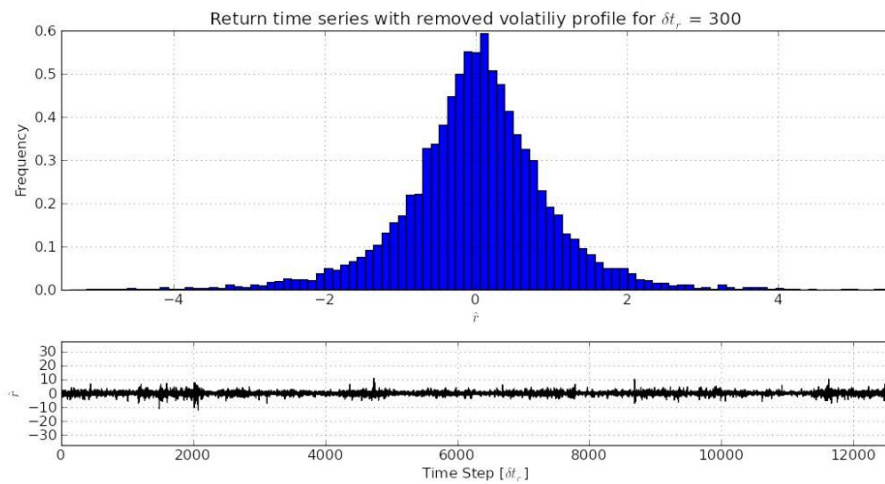[7]Profile plots are explained in Appendix A.1.

(a) $\hat{r}$ for $\delta t_r = 15$

(b) $\hat{r}$ for $\delta t_r = 75$

(c) $\hat{r}$ for $\delta t_r = 300$

Figure 3.2: The time series (black) and the corresponding distributions (blue) for the log-return time series with removed intraday volatility profiles for the return time intervals $\delta t_r = 15, 75$ and 300 seconds.
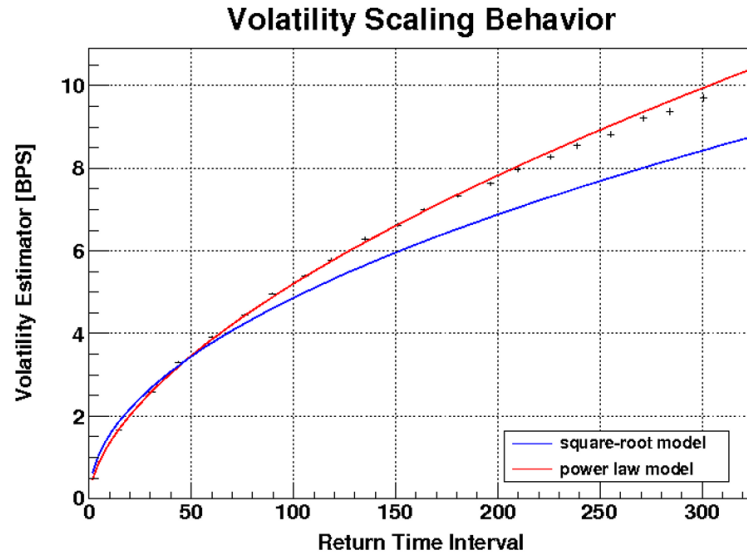
Figure 3.3: The volatility scaling behavior as a function of the return time interval δt_r. The blue curve shows a square-root model whereas the red curve depicts a power law model.

3.3.1   Pronounced Excess Kurtosis

The excess kurtosis has been constructed from the realized returns for every δt_r and for every day in the dataset after the removal of the intraday volatility profile. A profile plot is provided in order to illustrate the average excess kurtosis for each δt_r. This is shown in Figure 3.4.

The error bars for each data point result from averaging over each day's excess kurtosis. They indicate the estimated error on the mean of the excess kurtosis distribution which is observed for each δt_r.[8] The red line in Figure 3.4 indicates the daily excess kurtosis constructed from closing prices over the past ten years (2000-2010). Its value has been found to be $K_d = 7.58$.

A power law model is suggested in order to describe the δt_r-dependency of the excess kurtosis:

$$K = \kappa \cdot \delta t_r^{-\alpha} + c \tag{3.2}$$

The estimated model parameters are presented in Table 3.1.[9]

---

[8]The estimated error on a distribution's mean is given by $\hat{\sigma}_{\hat{\mu}_x} = \hat{\sigma}/\sqrt{N}$, where N is the number of values from which the mean $\hat{\mu}_x$ is constructed and $\hat{\sigma}$ denotes the distribution's standard deviation.

[9]The quantity χ^2/NDF denotes the mean variance for the model per data point and is an indicator for the goodness-of-fit. In case when the measurements are independent, the expectation value for this quantity is one. Here, the kurtosis estimations are positively correlated. Thus, the

Figure 3.4: The excess kurtosis is shown in dependency of the return time interval δt_r. A power law model is suggested to describe this relation. The corresponding parameter estimations are given in Table 3.1.

It is clearly evident that the excess kurtosis is strongly intensified on short timescales. The value of the suggested mod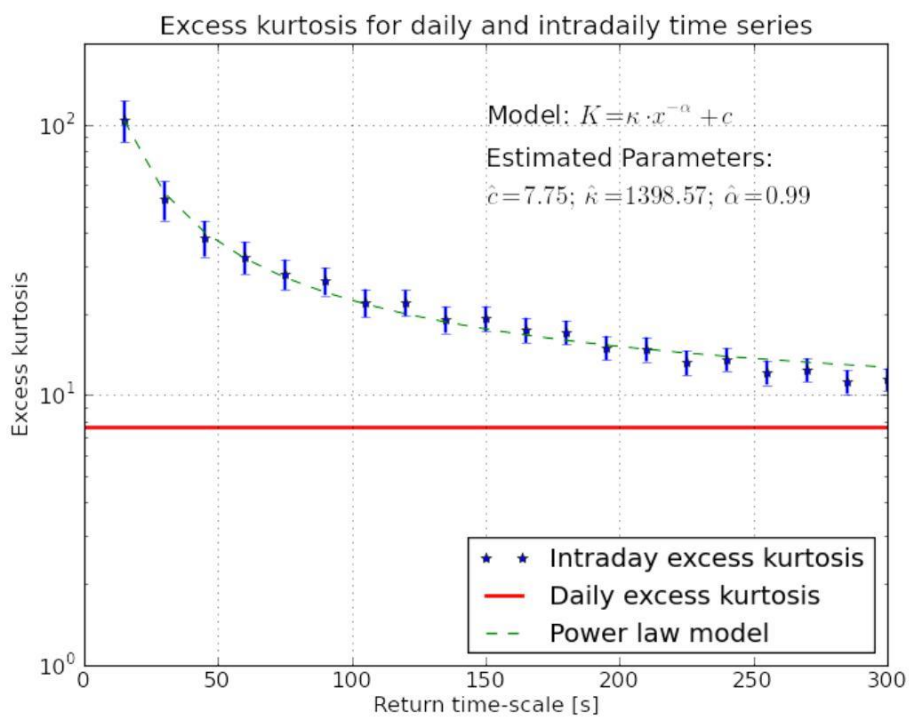el's constant c is very close to the daily excess kurtosis K_d ($c = 7.75$). This implies that the suggested model will approximately decay to K_d with increasing δt_r.

κ	α	c	NDF	χ^2	$\chi^2/$ NDF
1398.57	0.99	7.75	17	8.89	0.52

Table 3.1: Estimated parameters for model 3.2 in order to describe the excess kurtosis dependency on δt_r as illustrated in Figure 3.4.

3.3.2   Autocorrelation and Dependence

The sample autocorrelation function (SACF), the sample partial autocorrelation function (SPACF) and the sample autocorrelation function of squared returns of S&P 500 Index high-frequency data are discussed in this section. The examination of these functions allows to visually analyze autocorrelation and dependency patterns in the time series.

Figures 3.5 and 3.6 illustrate the SACF and the SPACF for log-returns as well as the SACF for squared returns for different time scales δt_r.

Autocorrelation appears to be significantly present on the first few lags especially on the shortest return time intervals. Despite being significant, their values remain quite low, thus only indicating weak autocorrelation effects.

Long range dependency effects do not seem to be existing since none of the SACFs constructed from log-return time series exhibit persistent behavior.

The SACFs of squared returns show persistent instead of exponentially decreasing behavior. This leads to the conclusion that the squared log-returns, which are proxies for the unobserved volatility, exhibit long range dependency effects. This motivates the employment of FIGARCH models which are capable of capturing those effects. This phenomenon is observed in all cases between $\delta t_r = 15$ and $\delta t_r = 300$.

3.4   Modeling Time Series with ARMA-GARCH Models

In this section, three different ARMA(1,1)-GARCH(1,1) models (as defined in Section 2.2.5) are empirically compared in their capability in modeling intraday

---

value of $\chi^2/$NDF is expected to be smaller than one.
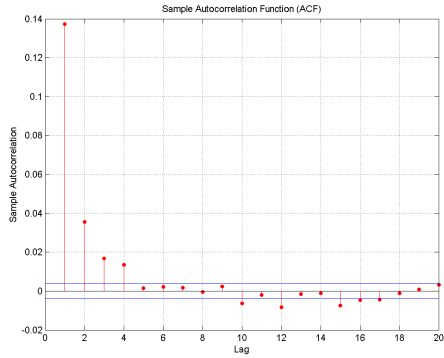
(a) SACF for $\delta t_r = 15$
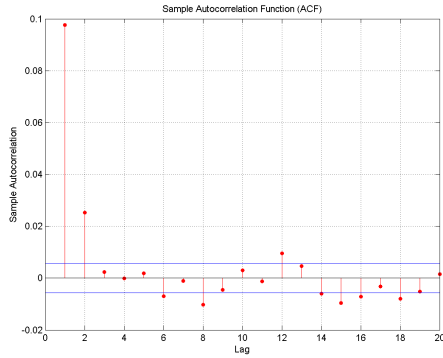
(b) SACF for $\delta t_r = 30$
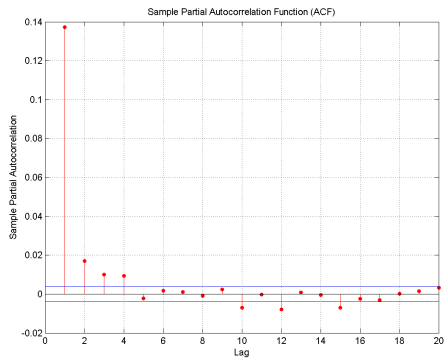
(c) SPACF for $\delta t_r = 15$
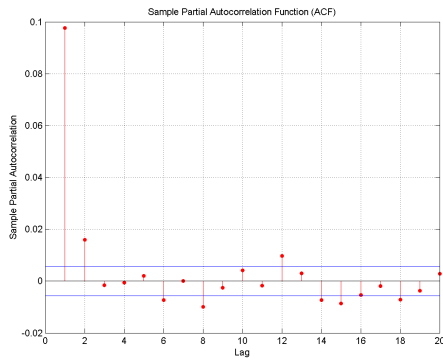
(d) SPACF for $\delta t_r = 30$
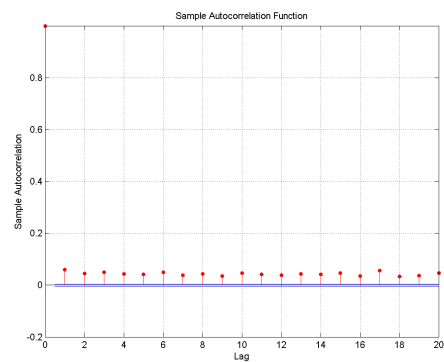
(e) SACF of squared returns for $\delta t_r = 15$
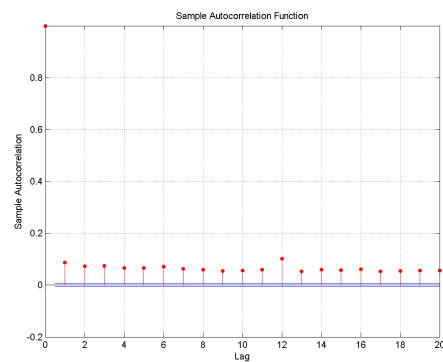
(f) SACF of squared returns for $\delta t_r = 30$

Figure 3.5: The SACF and the SPACF for log-returns as well as the SACF for squared log-returns for S&P 500 Index time series with $\delta t_r = 15$ (left hand side) and $\delta t_r = 30$ (right hand side). They both exhibit significant autocorrelation as well as significant and persistent dependency effects.
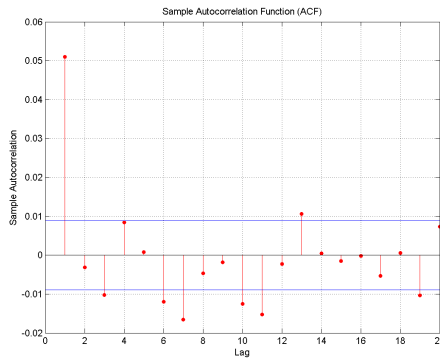
(a) SACF for $\delta t_r = 75$
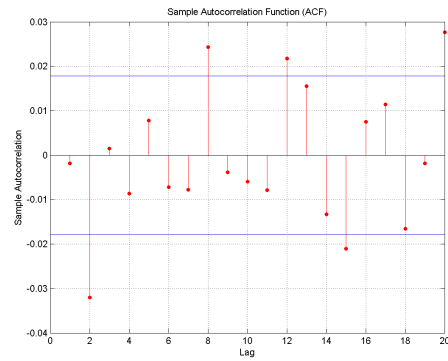
(b) SACF for $\delta t_r = 300$
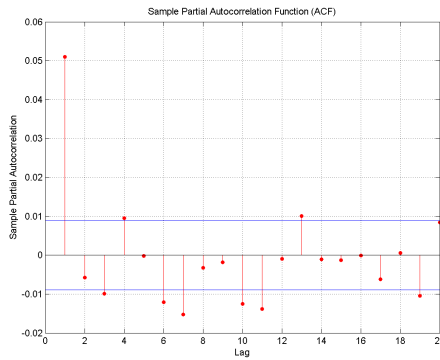
(c) SPACF for $\delta t_r = 75$
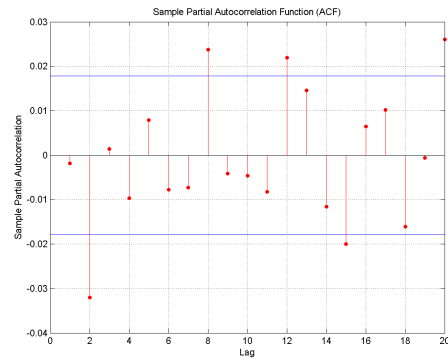
(d) SPACF for $\delta t_r = 300$
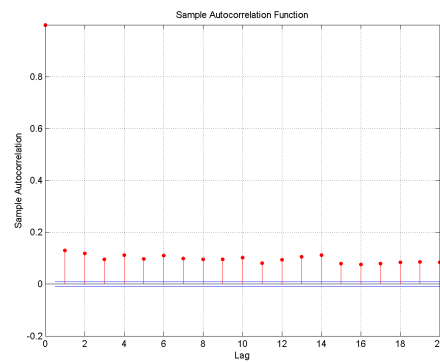
(e) SACF of squared returns for $\delta t_r = 75$
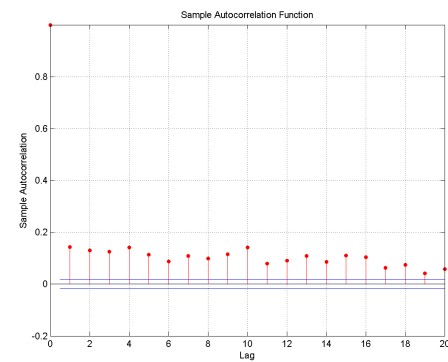
(f) SACF of squared returns for $\delta t_r = 300$

Figure 3.6: The SACF and the SPACF for log-returns as well as the SACF for squared log-returns for the S&P 500 Index time series with $\delta t_r = 75$ (left hand side) and $\delta t_r = 300$ (right hand side). The autocorrelation effects on log-returns are rather weak whereas the dependency effects appear pronounced as compared to the charts for $\delta t_r = 15$ and $\delta t_r = 30$.

S&P 500 Index log-return time series on high-frequency time scales. As shown in Bollerslev *et al.* (1992), GARCH(1,1) models are an adequate choice in modeling financial time series. Therefore, the focus in this analysis is on ARMA(1,1)-GARCH(1,1) models. Commonly used innovation distributions were introduced in Section 2.4. Here, the normal- and the Student-t distribution are compared to the classical tempered stable (CTS) distribution. As mentioned in Section 2.1.2, several stylized facts are pronounced in high-frequency data. Even based on daily data, the classical distribution assumptions, such as normal and Student-t, have been argued against in several studies. Hence, it is to assume that the statistical evidence for rejecting these hypothesis on high-frequency data will be even stronger. The CTS innovation distribution is employed in the ARMA-GARCH framework to present an alternative to the conventional models, especially for high-frequency applications. For this analysis, time series on return time intervals from $\delta t_r = 15$ seconds to $\delta t_r = 300$ seconds were used. The intraday volatility profile is removed in all time series as described in Section 3.2.

Engle's arch-test is carried out in order to detect conditional heteroskedasticity in the log-return time series (Engle (1982)). The null hypothesis of no conditional heteroskedasticity was rejected in all cases ($\delta t_r = 15, ..., 300$) at the 5% confidence level.

The charts in Figures 3.5 and 3.6 as well as Engle's arch test results motivate the employment of ARMA-GARCH models.

The (quasi-) maximum-likelihood method is used to estimate the model parameters (as introduced in Section 2.3.4).[10] The CTS-ARMA-GARCH approach is carried out in a way as described in Kim *et al.* (2010c): (1) estimate ARMA and GARCH parameters using the t-ARMA-GARCH model; (2) extract innovations, and; (3) fit the standard CTS parameters using the extracted innovations.

A comprehensive overview of the estimated parameters is presented in Table 3.2. The corresponding 1σ parameter confidence intervals are given in brackets.

In order to estimate the goodness-of-fits, Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests are applied. The results are summarized in the following enumeration.

- The normal-ARMA-GARCH model is rejected for all δt_r even at the 99.9% confidence level.

- The t-ARMA-GARCH model is rejected for all δt_r at the 99% level (except one case where $\delta t_r = 195$).

- In both AD and KS tests, the CTS-ARMA-GARCH model is accepted in all cases at the 99% confidence level and in almost all cases at the 95%

---

[10]For these studies, MathWorks MATLAB was employed.

confidence level for time series with $\delta t_r > 60$ s. For $\delta t_r \leq 60$ s, the CTS-ARMA-GARCH model is rejected in all tests at the 99 % confidence level.

Table 3.3 provides a complete overview of each model's KS and AD statistics as well as the corresponding p-values.

The normal-ARMA-GARCH and t-ARMA-GARCH models clearly fail to describe the underlying return time series. The models based on tempered stable distributions lead to significantly higher acceptance rates in the goodness-of-fit tests than the normal- and t-ARMA-GARCH models. The CTS-ARMA-GARCH model is capable of describing log-return time series in the case of the S&P 500 Index for δt_r's from roughly one minute on to lower frequencies.

Figures 3.7a and 3.7b show the achieved p-values for the KS-Test and for the AD-Test in dependence of the return time interval δt_r. The 1% confidence level is indicated by a black dashed line.

Table 3.3 as well as Figures 3.7a and 3.7b clearly indicate that the CTS distribution is a significantly more suitable model for the innovation distribution for high-frequency financial data. Nevertheless, on the shortest times scales ($\delta t_r = 15$ s to $\delta t_r = 60$ s), the CTS models is rejected at the 1% confidence level in both tests.

Figures 3.8a to 3.8f show the employed innovation distribution models together with the received parameter estimations for the normal-, t- and CTS-ARMA-GARCH model for the return time intervals $\delta t_r = 75$ and $\delta t_r = 300$ seconds. The corresponding QQ-Plots are depicted in Figures 3.9a to 3.9f.

3.4.1   Investigating the Significance of Mean Forecasts

The significance of the conditional mean forecasts which were utilized in the VaR-backtest are analyzed in this section. The forecasted conditional mean values are generated by the ARMA part of the employed ARMA(1,1)-GARCH(1,1) model, as given by Equation 2.30. If the forecasted conditional mean values are significantly different from zero, the ARMA part provides a substantive influence on the model's performance. Usually a 95% or a 99 % confidence interval is used in order to test this hypothesis. As a rule of thumb, the forecasted conditional mean values should be at least three times their conditional standard error apart from zero. Figure 3.10 illustrates the conditional mean forecasts in combination with their corresponding conditional standard errors. It is clearly evident from these charts that the conditional mean forecasts are not significantly non-zero.
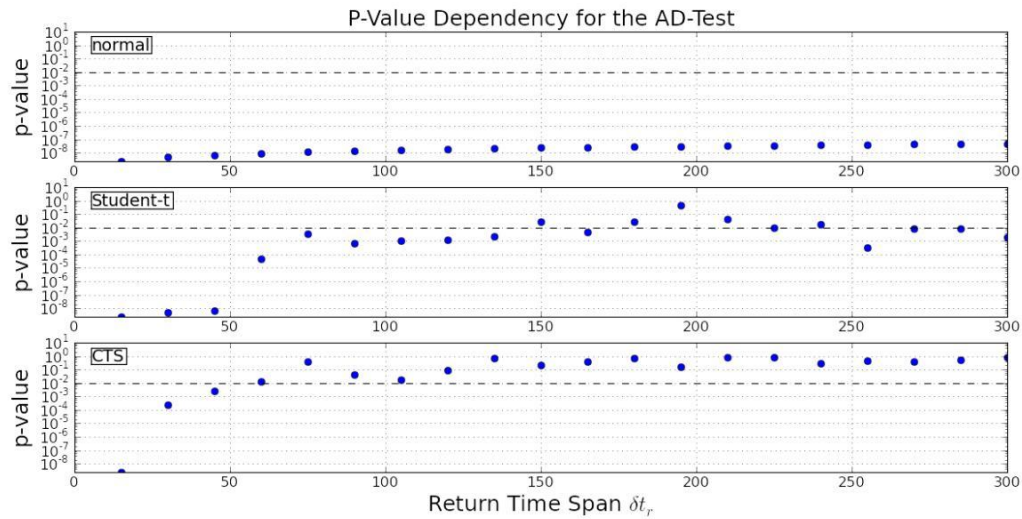
(a) AD Test
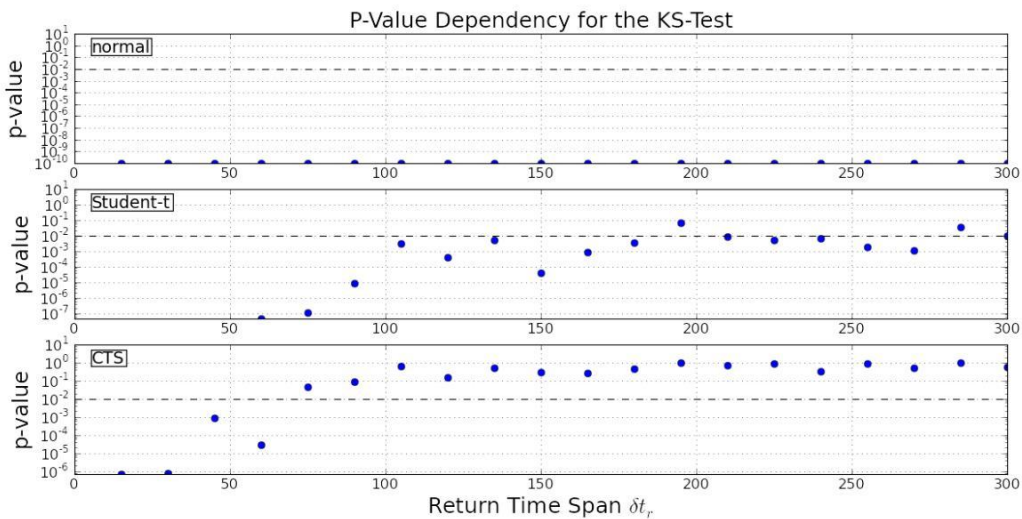
(b) KS Test

Figure 3.7: Figure 3.7a and Figure 3.7b show the achieved p-values for the Kolmogorov Smirnov (KS) as well as for the Anderson Darling (AD) test in dependence of the return time interval δt_r for the employed ARMA(1,1)-GARCH(1,1) models. The black dashed line indicates the 1% confidence level.
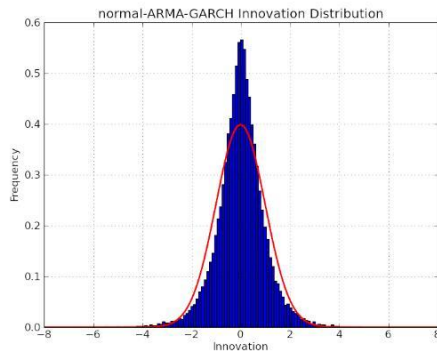
(a) normal, $\delta t_r = 75$
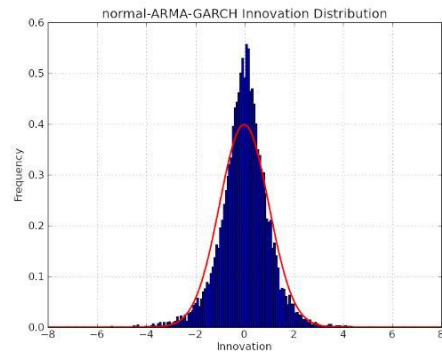
(b) normal, $\delta t_r = 300$
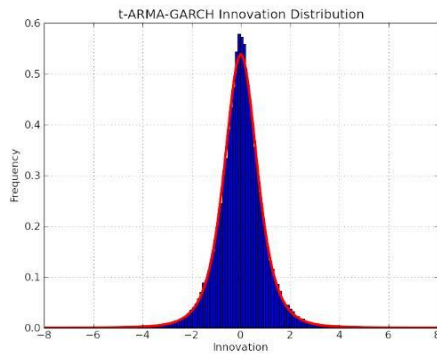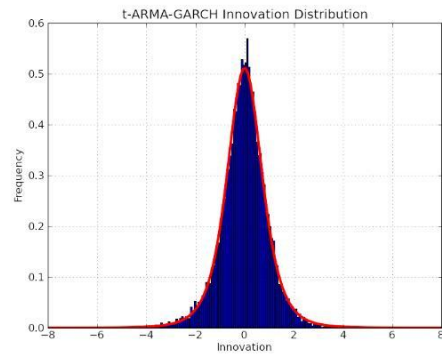
(c) Student-t, $\delta t_r = 75$
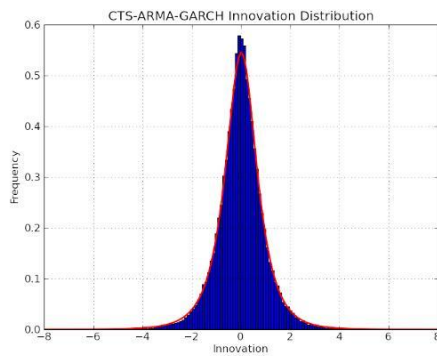
(d) Student-t, $\delta t_r = 300$
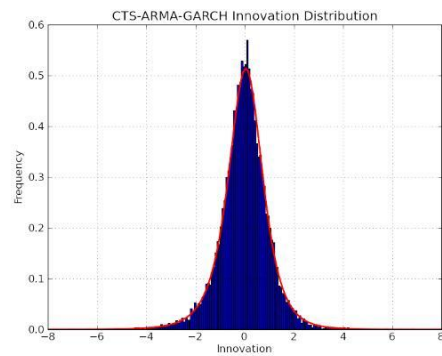
(e) CTS, $\delta t_r = 75$

(f) CTS, $\delta t_r = 300$

Figure 3.8: The above charts show the conditional innovation distributions for the respective ARMA-GARCH models. The distributions are shown for $\delta t_r = 75$ and $\delta t_r = 300$ seconds.
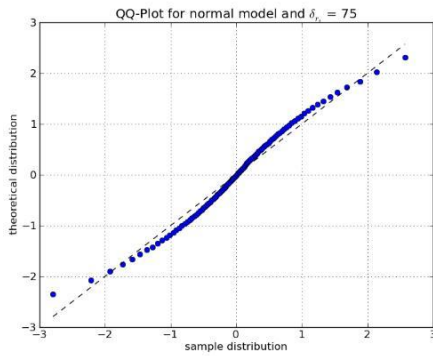
(a) normal, $\delta t_r = 75$
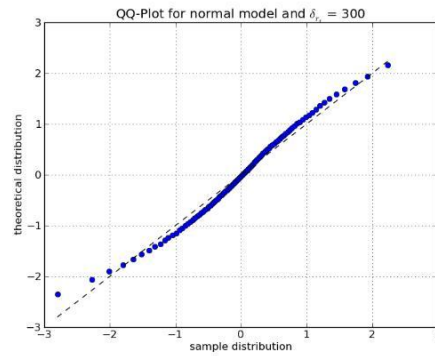
(b) normal, $\delta t_r = 300$
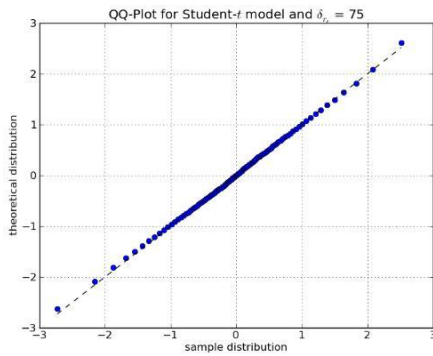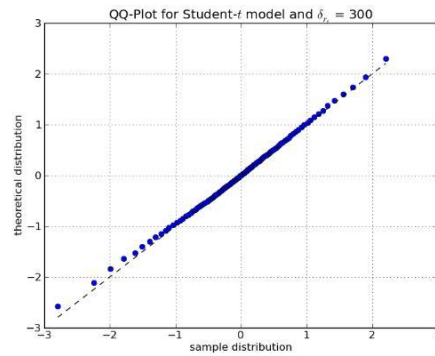
(c) Student-t, $\delta t_r = 75$
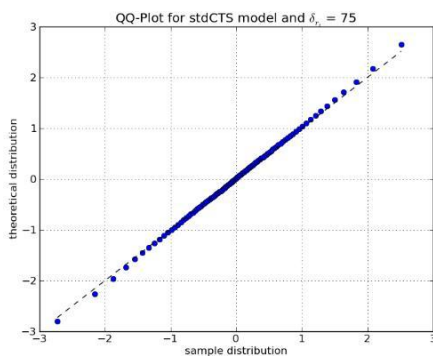
(d) Student-t, $\delta t_r = 300$
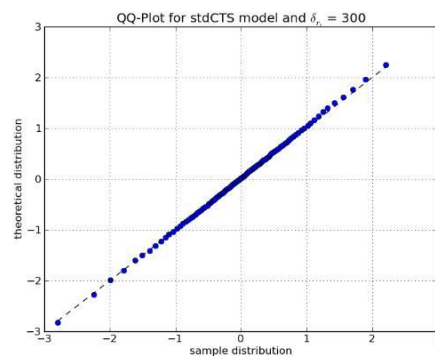
(e) CTS, $\delta t_r = 75$

(f) CTS, $\delta t_r = 300$

Figure 3.9: The above graphs show QQ-Plots for $\delta t_r = 75$ and $\delta t_r = 300$ seconds return time intervals for the three compared innovation distribution assumptions: Normal, Student-t and CTS.
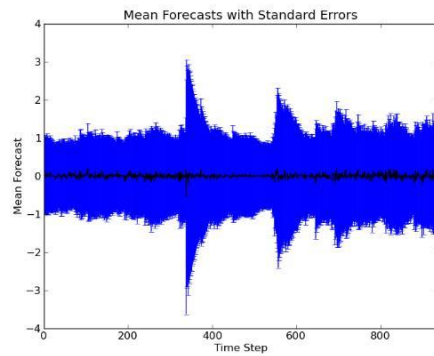
(a) normal, $\delta t_r = 75$
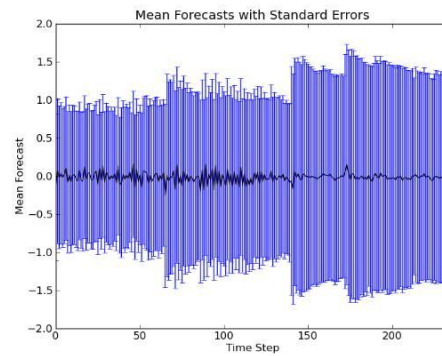
(b) normal, $\delta t_r = 300$
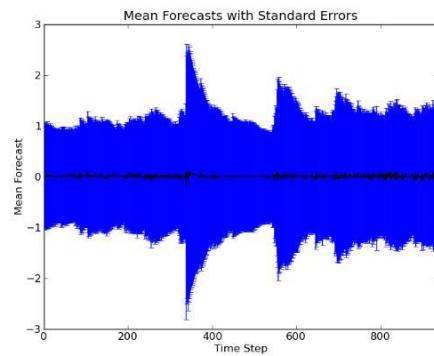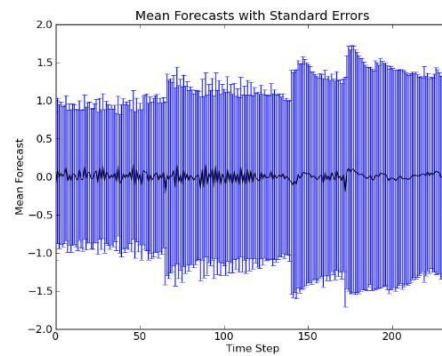
(c) Student-t, $\delta t_r = 75$

(d) Student-t, $\delta t_r = 300$

Figure 3.10: The above figures illustrate the conditional mean forecasts with their corresponding conditional standard errors. This is shown for the return time intervals $\delta t_r = 75$ and $\delta t_r = 300$ and the normal- and t-ARMA-GARCH models.

3.5 Model Parameter Dependencies on the Return Time Interval

Using the CTS-ARMA-GARCH parameter estimation results from the previous section (summarized in Table 3.2), it is possible to investigate potential parameter dependencies on the return time interval δt_r. This analysis is carried out for δt_r in the range from 75s to 300s, which corresponds to the δt_r-range in which the CTS-ARMA-GARCH model is accepted based on the AD and KS test results from the previous section (see Table 3.3).

The re-parametrization of the CTS parameters, as described in Equations 2.46, is used in this analysis in order to simplify the interpretation of the results. In the following, the CTS and the GARCH parameters are investigated for potential dependencies on δt_r. The contribution of the ARMA-part to the forecasting capability of the employed ARMA(1,1)-GARCH(1,1) model turns out to be negligible. This is indicated by the relationship between the ARMA parameters: $a_1 \approx -b_1$, as shown in Table 3.2 and the conditional mean return forecasts which are not significantly different from zero as indicated in Figure 3.10. Hence, the focus in this analysis is on the GARCH parameters only.

3.5.1 CTS-Parameter Dependencies

The dependencies from the CTS-Parameters α, Δ and λ on the return time interval δt_r are examined in this section. The parameters are received from an optimization algorithm, which not only provides the estimated parameter values but also computes the respective standard errors.[11] Error weighted trend lines are added to each chart in order to highlight potential linear dependencies. Note that the error bars' size increases with growing δt_r. This is due to the fact the number of return values decreases with growing δt_r over a fixed period of time. This change in sample size is reflected in the uncertainty of the parameter estimations. The 95% confidence levels for the estimated parameters result from their corresponding 1σ standard errors.

The CTS parameter dependencies are illustrated in Figure 3.11. The trend line's regression parameters are given in Table 3.4.

The relationship between α and δt_r indicates that the heavy tailedness decreases with increasing δt_r. This behavior was as well empirically observed on

---

[11]The parameters' 1σ errors are illustrated with error bars and result directly from the optimization algorithm. The optimization algorithm Minuit has been used in order minimize the log-likelihood function. This optimizer is widely used in the field of high-energy physics, is open source and has been developed at the CERN (European Organization for Nuclear Research). For more details, see http://www.dnp.fmph.uniba.sk/cernlib/asdoc/minuit/minmain.html.
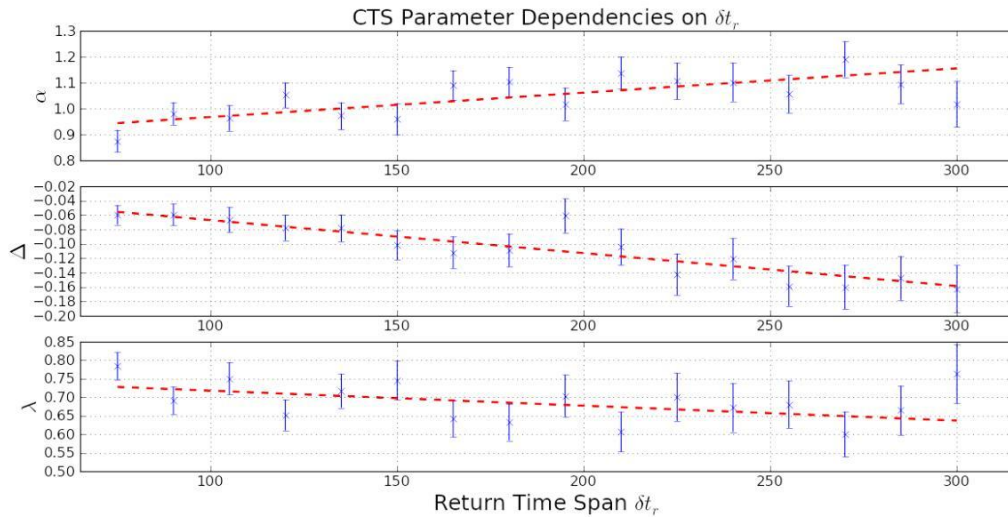
Figure 3.11: Dependency from the CTS-Parameters α,Δ and λ on the return time interval δt_r using the results from Table 3.2.

unconditional returns in Section 3.3.1, where the dependency of the sample excess kurtosis over the return time interval was examined. The parameter Δ remains negative for all δt_r indicating a negative skewness. The linear correlation between Δ and δt_r is not straightforwardly interpretable since the magnitude of the skewness depends on all three CTS-parameters. Thus, a linear dependency does not necessarily indicate an increasing negative skewness. As the t-statistic and R^2 for the dependency of λ on δt_r indicate, there is, if any, only a very slight dependency. The slope's 95% confidence interval indicates that the statistical evidence for the slope being significantly different from zero is very weak. Thus, one could argue that a constant model would be almost as well suited as a linear model in order to describe the relationship between λ and δt_r.

3.5.2 GARCH Parameter Dependencies

In the following, the CTS-GARCH parameters are examined for a potential dependency on δt_r, resulting from the model:

$$\epsilon_t = \sqrt{h_t}z_t, \quad z_t \propto stdCTS(\alpha, \lambda, \delta) \tag{3.3}$$

$$h_t = w_0 + \sum_{i=1}^{r} w_i\epsilon_{t-i}^2 + \sum_{j=1}^{s} v_jh_{t-j}. \tag{3.4}$$

The GARCH parameter dependencies are shown in Figure 3.12. Error weighted trend lines are added to each chart in order to visualize potential parameter dependencies. The respective linear regression parameters are given in Table 3.4. The
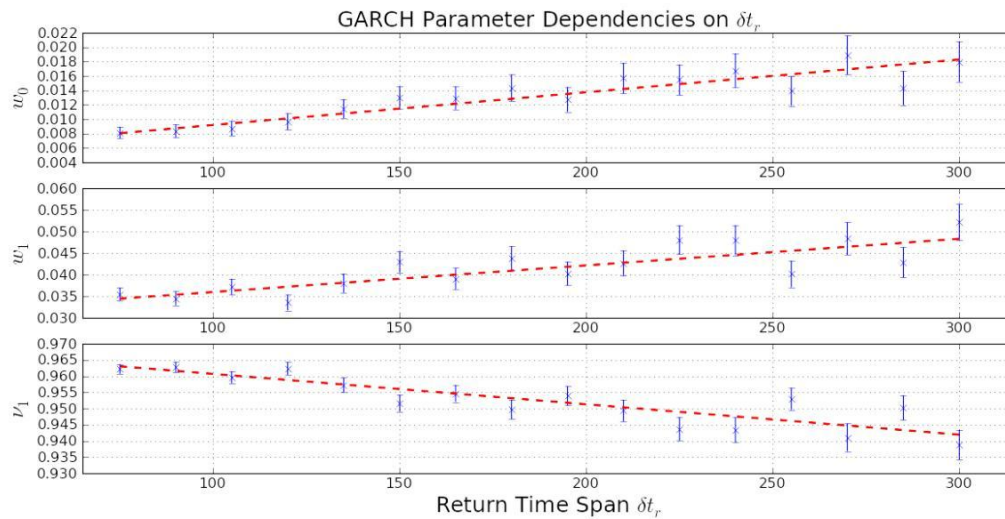
Figure 3.12: Dependency from the GARCH(1,1)-Parameters w_0, w_1 and ν_1 on the return time interval δt_r using the results from Table 3.2.

model's constant w exhibits a slightly increasing dependency. Since all time series have been normalized to an unconditional variance of one before the application of the econometric models, this parameter is an indicator of how much of the conditional volatility can be described with a constant. Even though the parameter w_1 increases slightly with growing δt_r, it remains quite small ($w_1 < 0.1 \; \forall \; \delta t_r$). Consequently, ARCH effects play are rather subordinate role in this GARCH(1,1) model. The slight decrease of ν_1 over increasing δt_r implies slightly decreasing GARCH effects. Nevertheless, the parameter ν_1 remains significantly high and continues to be the most important model parameter. Thus, the model is dominated from GARCH effects.

δ_{t_r}	Model	C	a_1	b_1	K	u_1	v_1	DOF	α	λ_-	λ_+
15	t	0.057(0.018)	−0.704(0.337)	0.713(0.332)	0.018(0.003)	0.052(0.004)	0.939(0.005)	1.898(0.010)			
	$\mathcal{N}$	0.044(0.019)	−0.727(0.384)	0.735(0.379)	0.024(0.002)	0.056(0.002)	0.929(0.002)		0.497(0.008)	0.638(0.006)	0.640(0.006)
30	t	0.006(0.002)	0.814(0.051)	−0.843(0.047)	0.014(0.002)	0.043(0.003)	0.950(0.004)	2.799(0.028)			
	$\mathcal{N}$	0.029(0.048)	−0.176(1.871)	0.181(1.870)	0.021(0.002)	0.048(0.002)	0.939(0.002)		0.432(0.020)	0.900(0.017)	0.915(0.017)
45	t	0.067(0.016)	−0.946(0.107)	0.944(0.109)	0.019(0.003)	0.048(0.004)	0.941(0.004)	3.289(0.043)			
	$\mathcal{N}$	0.030(0.046)	−0.138(1.737)	0.143(1.736)	0.024(0.001)	0.052(0.002)	0.933(0.002)		0.695(0.031)	0.801(0.026)	0.830(0.025)
60	t	0.006(0.002)	−0.818(0.048)	0.845(0.044)	0.014(0.002)	0.040(0.003)	0.953(0.003)	3.707(0.061)			
	$\mathcal{N}$	0.032(0.021)	−0.269(0.702)	0.280(0.699)	0.022(0.001)	0.051(0.001)	0.936(0.002)		0.835(0.039)	0.762(0.033)	0.801(0.032)
75	t	0.005(0.002)	0.832(0.051)	−0.855(0.047)	0.017(0.002)	0.048(0.003)	0.943(0.004)	3.865(0.073)			
	$\mathcal{N}$	0.029(0.022)	−0.098(0.757)	0.108(0.756)	0.026(0.001)	0.052(0.002)	0.932(0.002)		0.875(0.043)	0.754(0.038)	0.814(0.037)
90	t	0.051(0.014)	−0.993(0.003)	0.996(0.002)	0.015(0.002)	0.048(0.003)	0.944(0.004)	3.889(0.080)			
	$\mathcal{N}$	0.032(0.013)	−0.457(0.309)	0.475(0.305)	0.019(0.001)	0.047(0.002)	0.941(0.002)		0.981(0.044)	0.661(0.038)	0.721(0.036)
105	t	0.004(0.001)	−0.863(0.039)	−0.883(0.036)	0.016(0.002)	0.043(0.003)	0.949(0.003)	4.098(0.095)			
	$\mathcal{N}$	0.042(0.015)	−0.597(0.263)	0.612(0.259)	0.030(0.001)	0.050(0.001)	0.931(0.002)		0.963(0.049)	0.717(0.044)	0.784(0.043)
120	t	0.035(0.028)	−0.482(1.127)	0.486(1.124)	0.013(0.002)	0.040(0.003)	0.954(0.003)	3.985(0.096)			
	$\mathcal{N}$	0.034(0.018)	−0.194(0.546)	0.207(0.545)	0.021(0.001)	0.050(0.001)	0.938(0.002)		1.052(0.049)	0.613(0.043)	0.691(0.041)
135	t	0.043(0.012)	−0.746(0.190)	0.736(0.193)	0.014(0.002)	0.044(0.003)	0.950(0.003)	3.940(0.100)			
	$\mathcal{N}$	0.019(0.026)	−0.204(1.078)	−0.198(1.079)	0.022(0.001)	0.047(0.001)	0.939(0.002)		0.972(0.053)	0.678(0.047)	0.756(0.045)
150	t	0.003(0.001)	0.876(0.029)	−0.897(0.026)	0.013(0.002)	0.039(0.002)	0.955(0.003)	4.019(0.109)			
	$\mathcal{N}$	0.024(0.016)	−0.039(0.664)	0.050(0.664)	0.018(0.001)	0.045(0.001)	0.944(0.001)		0.959(0.061)	0.695(0.052)	0.796(0.052)
165	t	0.003(0.001)	−0.873(0.031)	−0.892(0.029)	0.013(0.002)	0.043(0.002)	0.952(0.003)	4.096(0.117)			
	$\mathcal{N}$	0.022(0.012)	−0.005(0.476)	0.019(0.476)	0.019(0.001)	0.050(0.001)	0.939(0.001)		1.090(0.057)	0.586(0.051)	0.698(0.048)
180	t	0.003(0.001)	0.888(0.025)	−0.907(0.022)	0.011(0.001)	0.038(0.002)	0.957(0.002)	4.073(0.121)			
	$\mathcal{N}$	0.020(0.017)	0.163(0.669)	−0.154(0.671)	0.013(0.000)	0.039(0.001)	0.954(0.001)		1.103(0.058)	0.578(0.051)	0.687(0.049)
195	t	0.034(0.009)	−0.550(0.142)	0.571(0.139)	0.010(0.001)	0.034(0.002)	0.962(0.002)	4.098(0.129)			
	$\mathcal{N}$	0.039(0.010)	−0.561(0.079)	0.598(0.076)	0.017(0.001)	0.041(0.001)	0.948(0.001)		1.017(0.064)	0.673(0.057)	0.734(0.055)
210	t	0.034(0.008)	−0.582(0.111)	0.605(0.109)	0.009(0.001)	0.037(0.002)	0.960(0.002)	4.141(0.135)			
	$\mathcal{N}$	0.029(0.009)	−0.451(0.113)	0.485(0.110)	0.015(0.001)	0.039(0.001)	0.951(0.001)		1.138(0.062)	0.554(0.055)	0.659(0.052)
225	t	0.026(0.007)	−0.345(0.152)	0.369(0.150)	0.008(0.001)	0.034(0.002)	0.963(0.002)	4.552(0.165)			
	$\mathcal{N}$	0.022(0.006)	−0.082(0.131)	0.120(0.131)	0.016(0.000)	0.038(0.001)	0.953(0.001)		1.106(0.071)	0.629(0.067)	0.771(0.063)
240	t	0.020(0.005)	−0.254(0.094)	0.293(0.093)	0.008(0.001)	0.035(0.002)	0.962(0.002)	4.359(0.158)			
	$\mathcal{N}$	0.017(0.005)	−0.045(0.080)	0.102(0.079)	0.014(0.000)	0.040(0.001)	0.952(0.001)		1.101(0.074)	0.611(0.067)	0.732(0.064)
255	t	0.017(0.004)	−0.014(0.070)	0.065(0.070)	0.007(0.001)	0.034(0.001)	0.965(0.001)	4.134(0.149)			
	$\mathcal{N}$	0.015(0.004)	0.121(0.059)	−0.054(0.060)	0.010(0.000)	0.033(0.000)	0.961(0.001)		1.056(0.073)	0.601(0.065)	0.760(0.062)
270	t	0.015(0.003)	−0.065(0.041)	0.136(0.041)	0.006(0.001)	0.032(0.001)	0.968(0.001)	4.347(0.165)			
	$\mathcal{N}$	0.015(0.004)	0.031(0.039)	0.058(0.039)	0.009(0.001)	0.034(0.001)	0.966(0.000)		1.190(0.069)	0.520(0.063)	0.680(0.059)
285	t	0.009(0.002)	0.151(0.023)	−0.053(0.023)	0.006(0.000)	0.025(0.000)	0.972(0.000)	4.194(0.165)			
	$\mathcal{N}$	0.012(0.003)	0.166(0.023)	−0.045(0.023)	0.006(0.000)	0.034(0.001)	0.966(0.001)		1.094(0.075)	0.591(0.068)	0.739(0.065)
300	t	0.004(0.001)	0.357(0.010)	−0.232(0.010)	0.017(0.001)	0.036(0.001)	0.964(0.001)	4.382(0.179)			
	$\mathcal{N}$	0.007(0.002)	−0.311(0.011)	−0.158(0.012)	0.006(0.000)	0.019(0.000)	0.979(0.000)		1.017(0.088)	0.683(0.081)	0.845(0.078)

Table 3.2: Estimated parameters for the t-ARMA-GARCH model with Student-t and CTS distributed innovations (denoted as model t) as well as for the normal-ARMA-GARCH model (denoted as model $\mathcal{N}$). The 1σ confidence intervals are shown in brackets.

δ_{tr}	normal-ARMA-GARCH		t-ARMA-GARCH		CTS-ARMA-GARCH	
	KS(p-value)	AD(p-value)	KS(p-value)	AD(p-value)	KS(p-value)	AD(p-value)
15	0.1137(0.0000)	7821.7116(0.0000)	0.0192(0.0000)	210.1317(0.0000)	0.0054(0.0000)	36.4156(0.0000)
30	0.0772(0.0000)	1908.0071(0.0000)	0.0185(0.0000)	63.3128(0.0000)	0.0076(0.0000)	7.2474(0.0003)
45	0.0638(0.0000)	944.9048(0.0000)	0.0132(0.0000)	18.5280(0.0000)	0.0068(0.0008)	5.1294(0.0025)
60	0.0566(0.0000)	542.2022(0.0000)	0.0128(0.0000)	11.4793(0.0000)	0.0094(0.0000)	3.5887(0.0139)
75	0.0545(0.0000)	366.8550(0.0000)	0.0131(0.0000)	8.1980(0.0001)	0.0061(0.0465)	0.9355(0.3930)
90	0.0571(0.0000)	328.7996(0.0000)	0.0115(0.0000)	7.7425(0.0001)	0.0061(0.0905)	2.5964(0.0441)
105	0.0521(0.0000)	247.2568(0.0000)	0.0107(0.0005)	2.9164(0.0302)	0.0039(0.6251)	3.4353(0.0165)
120	0.0529(0.0000)	237.0457(0.0000)	0.0125(0.0001)	2.2052(0.0710)	0.0064(0.1486)	2.0855(0.0825)
135	0.0519(0.0000)	204.6005(0.0000)	0.0103(0.0050)	6.0676(0.0009)	0.0049(0.5181)	0.5479(0.6985)
150	0.0510(0.0000)	164.0908(0.0000)	0.0160(0.0000)	5.0447(0.0027)	0.0061(0.3112)	1.3405(0.2196)
165	0.0533(0.0000)	152.9976(0.0000)	0.0137(0.0004)	3.1110(0.0240)	0.0067(0.2561)	0.9654(0.3760)
180	0.0538(0.0000)	141.2094(0.0000)	0.0126(0.0024)	2.9452(0.0292)	0.0058(0.4701)	0.5262(0.7203)
195	0.0540(0.0000)	137.7518(0.0000)	0.0083(0.1368)	1.8875(0.1060)	0.0035(0.9736)	1.6215(0.1499)
210	0.0511(0.0000)	121.9901(0.0000)	0.0122(0.0094)	3.2723(0.0200)	0.0052(0.7243)	0.4725(0.7750)
225	0.0483(0.0000)	95.1065(0.0000)	0.0118(0.0193)	4.0850(0.0079)	0.0043(0.9210)	0.4099(0.8389)
240	0.0482(0.0000)	86.8460(0.0000)	0.0144(0.0030)	2.8650(0.0321)	0.0076(0.3159)	1.1855(0.2731)
255	0.0527(0.0000)	99.6204(0.0000)	0.0171(0.0004)	5.1335(0.0025)	0.0075(0.3770)	0.4208(0.8279)
270	0.0525(0.0000)	85.3274(0.0000)	0.0149(0.0040)	4.3917(0.0056)	0.0071(0.4745)	0.7424(0.5245)
285	0.0519(0.0000)	78.9733(0.0000)	0.0152(0.0043)	3.0174(0.0268)	0.0066(0.6208)	0.2557(0.9673)
300	0.0493(0.0000)	77.2590(0.0000)	0.0154(0.0054)	3.2163(0.0213)	0.0057(0.8119)	0.8189(0.4676)

Table 3.3: Goodness-of-Fit summary for the KS and AD tests. The three models were each tested on 20 time series. The respective p-values are shown in brackets.

Parameter	Slope	Slope t-stat.	Intersect	Intersect t-stat.	R^2
CTS-Parameters					
α	$9.4 \cdot 10^{-4} (4.32 \cdot 10^{-3})$	4.27	0.87 (0.074)	22.89	0.49
Δ	$-4.6 \cdot 10^{-4} (1.6 \cdot 10^{-4})$	5.75	-0.021 (0.025)	1.6	0.92
λ	$-4 \cdot 10^{-4} (3.7 \cdot 10^{-4})$	2.1	0.76 (0.065)	23	0.21
GARCH-Parameters					
w_0	$4.6 \cdot 10^{-5} (1.1 \cdot 10^{-5})$	7.9	0.0046 (0.0017)	5.28	0.86
w_1	$6.2 \cdot 10^{-5} (1.8 \cdot 10^{-5})$	6.7	0.03 (0.003)	20	0.79
ν_1	$-9.4 \cdot 10^{-5} (1.9 \cdot 10^{-5})$	9.4	0.97 (0.003)	646	0.87

Table 3.4: Summary of CTS and GARCH(1,1) parameter dependencies over δt_r. The respective estimation errors are shown in brackets.

3.6 Figarch Modeling of the S&P 500 Index

A FIGARCH(1,d,0) model, as defined in Section 2.2.4, is applied to the S&P 500 Index time series on return time intervals ranging from 15 to 300 seconds. As shown in Figures 3.5 and 3.6, the SACF and SPACF of squared log-returns show persistent behavior, indicating long range dependency effects and motivating the employment of FIGARCH models. The corresponding function to calculate the conditional variance for the proposed FIGARCH(1,d,0) model is given by

$$\sigma_t^2 = w + \beta_1 \sigma_{t-1}^2 + (1 - \beta_1 L - (1 - L)^d)\epsilon_t^2. \tag{3.5}$$

By investigating Equation 3.5, it appears possible for the conditional variance to result in negative values. Such a set of parameters is clearly unwanted. Baillie *et al.* (1996) showed that the conditional variance is positive almost surely for all t, if the conditions $w > 0$ and $0 \le \beta_1 \le d \le 1$ are fulfilled.

In this empirical analysis, normal, Student-t as well as CTS distributions are employed as potential innovation distributions. In the case of Student-t innovations, the corresponding log-likelihood is given by

$$l_t = T \ln \frac{\Gamma((\nu + 1)/2)}{\sqrt{(\nu - 2)\pi}\Gamma(\nu/2)} - \frac{1}{2} \sum_{t=1}^{T} \left[\ln \sigma_t^2 + (\nu + 1) \ln \left(1 + \frac{\hat{\epsilon}_t^2}{(\nu - 2)\sigma_t^2} \right) \right], \tag{3.6}$$

whereas the log-likelihood function for normal innovations is given by

$$l_n = \frac{1}{2}T \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \left[\ln \sigma_t^2 + \frac{\hat{\epsilon}_t^2}{\sigma_t^2} \right]. \tag{3.7}$$

The respective model's parameters are received by optimizing the corresponding log-likelihood functions.

In the same way as CTS-ARMA-GARCH models were applied to time series, CTS-FIGARCH models are employed by following the steps: (1) Estimate FIGARCH parameters with Student-t innovations; (2) Extract residuals; (3) Apply the CTS model to the residuals.

The results of the parameter estimations are given in Table 3.5. The goodness-of-fit test results from Anderson-Darling and Kolmogorov-Smirnov tests together with the corresponding p-values are shown in Table 3.6. By comparing the number of acceptances between the three models, the CTS-FIGARCH model has the highest rate of acceptances. The achieved AD and KS p-values for the respective models are shown in Figure 3.13.

The significance of the parameter d is an indicator for the existence of long range dependency effects in the time series of squared residuals. The values of

d are significantly different from zero in all cases, as shown in Table 3.5. Thus, the model actually captures effects of long range dependency. Note that the parameter d exceeds persistently the parameter β_1 which ensures that the modeled conditional variance is almost surely positive on all tested time series and models.

Potential parameter dependencies for the t-FIGARCH model with CTS innovations (CTS-FIGARCH) is shown in Figure 3.14. Error weighted trend lines are added to each chart in order to emphasize potential parameter dependencies. The trend line's parameter values are given in Table 3.7. Similar to the parameter dependencies in the CTS-ARMA-GARCH model, the CTS parameter α increases with growing δt_r whereas the parameter Δ declines. The parameter λ also shows a slightly negative correlation to the return time interval, as observed in the analysis on CTS-ARMA-GARCH parameter dependencies.

The FIGARCH parameter d appears rather constant but exhibits stronger fluctuation than a linear model allows for (under consideration of the observed parameter's standard errors). The parameter w, the model's constant, shows a decreasing behavior on growing δt_r whereas the parameter β increases at the same time. The two parameters β and d are comparable in magnitude.

Finally, the FIGARCH model's descriptive performance is compared to the previous ARMA-GARCH approach. More precisely, the CTS-FIGARCH model is compared to the CTS-ARMA-GARCH model. Those model's exhibited the best performance in describing the underlying time series in the respective FIGARCH and ARMA-GARCH approaches.

The comparison is graphically carried out by visualizing the achieved p-values on the utilized set of return time horizons. This is shown in Figure 3.15. The black dashed line in these charts indicates the 1% confidence interval. The CTS-FIGARCH model appears slightly better suited in describing time series on return time intervals below 75 seconds. For lower frequencies however, the two compared model exhibit the same acceptance rate when considering a 1% confidence level.

3.7   Comparison of Risk Measures for ARMA-GARCH Models

The findings from Section 3.4 and 3.6 indicate that conditional mean and variance models based on both the normal and the Student-t distribution assumption fail in describing high-frequency S&P 500 Index time series in almost all tested cases. In contrast to this, models based on the tempered stable assumption show significantly improved descriptive power for the examined time series.

One of the most import applications in financial time series modeling is risk

(a) AD Test

(b) KS Test

Figure 3.13: The above charts show the descriptive performance of the analyzed FIGARCH models with normal, Student-t and CTS innovations by depicting the achieved p-values in the AD and the KS tests. The black dashed lines indicate the 1% confidence level.
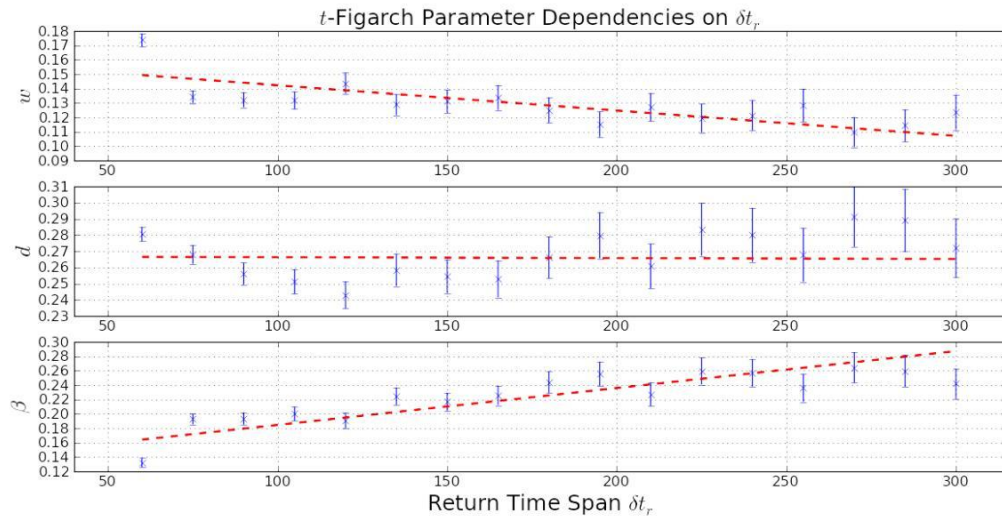
(a) FIGARCH Parameters
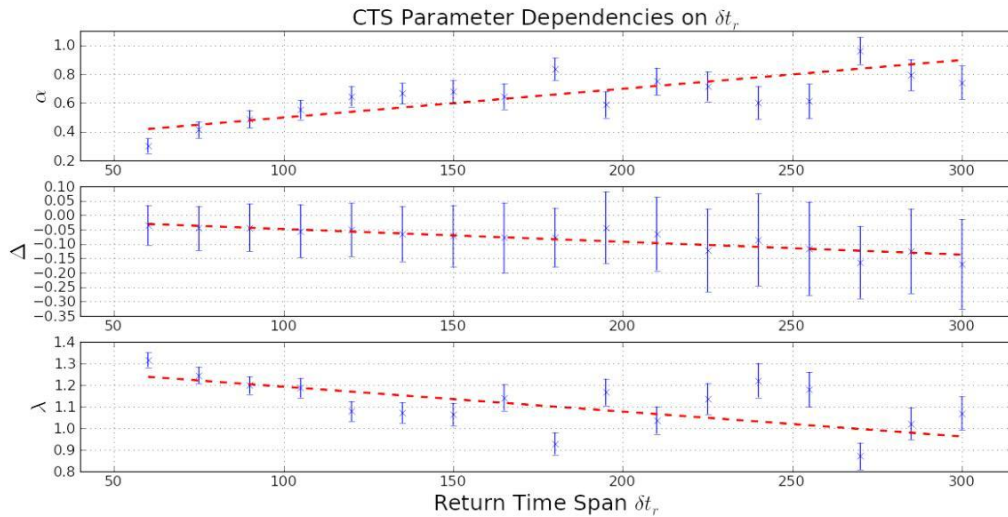
(b) CTS Parameters

Figure 3.14: The above charts show the estimated CTS-FIGARCH model parameters in dependence of the return time interval δt_r. Error weighted trend lines are added in order to visualize potential parameter dependencies.

(a) AD Test

(b) KS Test

Figure 3.15: The above charts compare the achieved p-values between the CTS-FIGARCH model and the CTS-ARMA-GARCH model at the 1% confidence level in the AD and KS tests. The CTS-FIGARCH model yields in the AD as well as in the KS test slightly better goodness-of-fit results. The dashed black lines indicate the 1% confidence level.

estimation. The purpose of this section is to show the significant differences in the distinctive model's performance in estimating market risk. Here, the ARMA(1,1)-GARCH(1,1) approach is chosen to investigate this issue.[12]

3.7.1 Value-at-Risk Backtesting

In this section, VaR backtests for the normal-, t-, and the CTS-ARMA-GARCH model are discussed. Using ARMA-GARCH models, VaR for the return at time $t + 1$ with information given until time t can be defined as follows:

$$\text{VaR}_{t,\eta}(y_{t+1}) = -\inf\{x \in \mathbb{R} | P_t(y_{t+1} \leq x) > \eta\},$$

where $P_t(y_{t+1} \leq x)$ is the conditional probability of y_{t+1} based on the information until time t.

A VaR backtest is provided for a time period of 13 days (ranging from June 8, 2010 until June 24, 2010).[13] It is carried out for all time series ($\delta t_r = 75$ to $\delta t_r = 300$), to compare the respective models' performance on different time scales.

A moving time-window of 10 days is used for the VaR backtest. This results in 3120 return values in the case of $\delta t_r = 75$ (which corresponds to roughly 12.5 years in daily data), the number of predicted returns is 936. In the case of $\delta t_r = 300$, the moving time window consists of 802 returns and the number of predicted returns is 234.

The models' risk estimation performance is compared for the 1% VaR Level. In order to verify if the tested models are capable of forecasting VaR correctly, the Christoffersen-Likelihood-Ratio test (CLR) (see Christoffersen (1998)) is applied. A detailed explanation hereof is given in Appendix A.3. It consists of two parts, the unconditional convergence test (UC), which corresponds to the Kupiec test (Kupiec (1995)) and a test for independence (IND), which examines if violations occur subsequently.

In addition to the CLR test, the Berkowitz-Likelihood-Ratio (BLR) test is applied (see Berkowitz (2001)). The Berkowitz test examines, how well an empirical tail distribution is described under a proposed theoretical distribution function. A detailed explanation of the BLR test is provided in Appendix A.4.

The CLR and BLR test results are summarized in Table 3.8 where the corresponding p-values are provided for each test. A graphical summary of the acceptances is given in Figure 3.16. This charts show the resulting p-values from the CLR test and the BLR tests at different VaR levels. The results state that the

---

[12]This section is part of my publication Beck *et al.* (2011).

[13]The period has been chosen under the constraint that the period contains no dates with gaps in the data.

CTS-ARMA-GARCH model is the most reliable model in order to estimate VaR for intraday U.S. market time series at the 1% level. It is important to note that the CTS-ARMA-GARCH model shows the best performance in the BLR tests, indicating that it models the empirical return distributions' negative tail better than the other two models. This is an important characteristic when Average Value-at-Risk is considered, as in the following section.

The normal-ARMA-GARCH model fails in all BLR tests. That is, it fails in describing the tail behavior correctly and is therefore not a reliable model to forecast VaR on U.S. intra-daily market data. Its performance in the CLR test is also well below those from the CTS- and t-ARMA-GARCH models.

The t-ARMA-GARCH model shows a similarly good VaR forecasting capability at the 1% level in the CLR as well as in the BLR tests. However, the 5% BLR test indicates that the tail modeling capability is not as good as for the CTS-ARMA-GARCH model. It is still significantly better than the normal-ARMA-GARCH model.

3.7.2 Average Value-at-Risk Analysis

In this section, the risk measure Average Value-at-Risk (AVaR) is investigated for the CTS-ARMA-GARCH model and compared to the AVaR for the normal-ARMA-GARCH and the t-ARMA-GARCH model. The risk measure AVaR provides, unlike VaR, the expected potential loss at a certain probability level η (VaR only gives an estimate for the smallest possible loss at a certain probability level η). AVaR is therefore strongly depending on the shape of the respective tail function. For this reason, the distribution function which is used to model the empirical return distribution must be well suited to describe the empirical distribution's tail. This property was investigated for each distribution with the BLR test in the previous section.

The risk measure AVaR is empirically compared on S&P 500 data between the CTS-ARMA-GARCH, the normal-ARMA-GARCH and the t-ARMA-GARCH model. The analysis is carried out for two return time intervals, $\delta t_r = 75$ and $\delta t_r = 300$ seconds. This corresponds to the shortest and the longest return time interval for which the CTS-ARMA-GARCH model was accepted in the AD and KS tests (see Section 3.4).

AVaR is defined as

$$\text{AVaR}_\eta(X) = \frac{1}{\eta} \int_0^\eta \text{VaR}_\epsilon(X)d\epsilon.$$

As shown in Kim *et al.* (2010c), the conditional AVaR for y_{t+1} with information until time t is given by $\text{AVaR}_{t,\eta}(y_{t+1}) = \frac{1}{\eta} \int_0^\eta VaR_{t,\epsilon}(y_{t+1})d\epsilon$. In the case of ARMA(1,1)-GARCH(1,1) models, $\text{AVaR}_{t,\eta}(y_{t+1})$ is obtained by
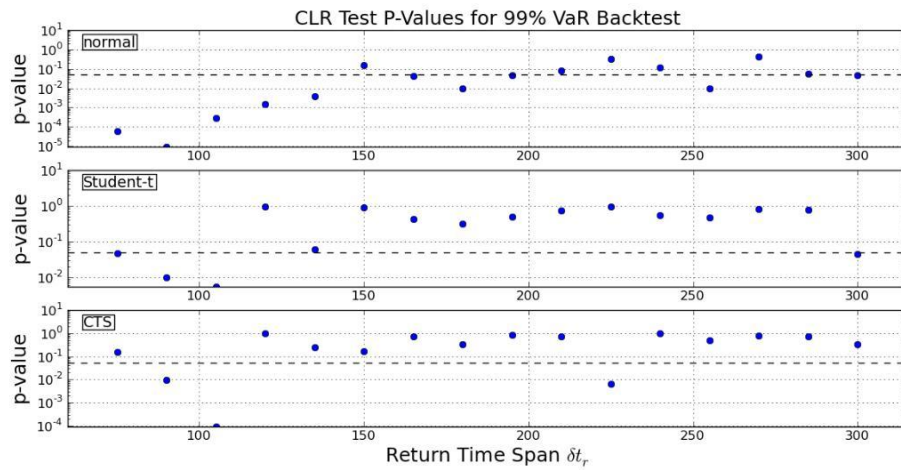
(a) 99% CLR Test
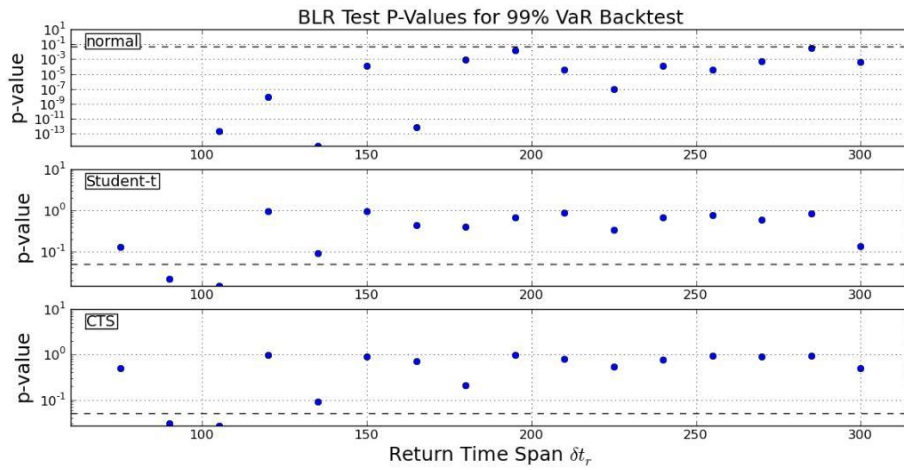
(b) 99%BLR Test
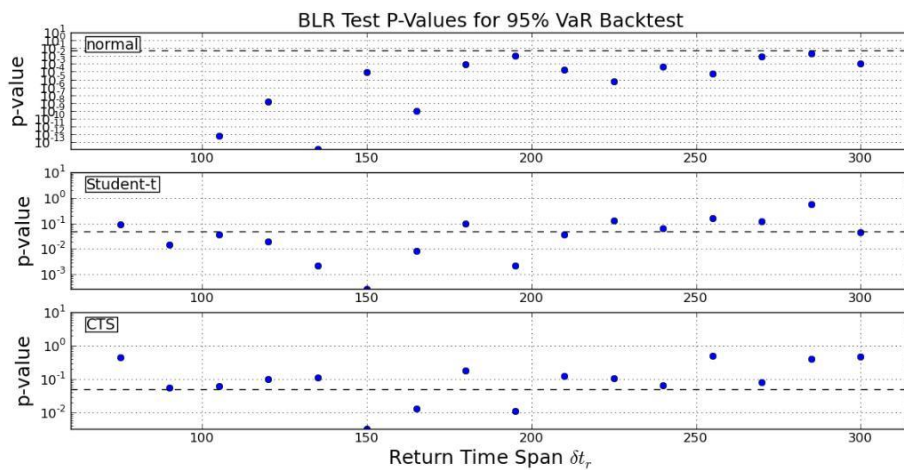
(c) 95 % BLR Test

Figure 3.16: The p-values in the CLR and BLR tests are shown in the above charts. The CLR test is shown for the 1% VaR level whereas the BLR test results are given at the 1% and 5% VaR levels. The black dashed line indicates the 5% confidence level.

$$\text{AVaR}_{t,\eta}(y_{t+1}) = -(c + ay_t + b\sigma_t\epsilon_t) + \sigma_{t+1}\text{AVaR}_{t,\eta}(\epsilon_{t+1}).$$

Since ϵ_{t+1} is independent to the information until t, $\text{AVaR}_{t,\eta}(\epsilon_{t+1}) = \text{AVaR}_\eta(\epsilon_{t+1})$. $\text{AVaR}_\eta(\epsilon_{t+1})$ can be obtained in the case of tempered stable distributions, as described in Kim *et al.* (2010c) and Kim *et al.* (2009), by

$$\begin{aligned}&\text{AVaR}_\eta(Y)\\&= \text{VaR}_\eta(Y) - \frac{\exp[-\text{VaR}_\eta(Y)\rho]}{\pi\eta}\Re\left(\int_0^\infty e^{-iu\text{VaR}_\eta(Y)}\frac{\phi_Y(-u + i\rho)}{(-u + i\rho)^2}du\right),\end{aligned}$$

where $\rho > 0$ has to be chosen such that $|\phi_Y(-u+i\rho)| < \infty$ for all $u \in \mathbb{R}$. ϕ_Y is the characteristic function of the random variable Y. The AVaR figures for Student-t and normal distributions can be computed with standard numerical methods.

Figure 3.17 shows VaR and AVaR at the 1% level for the ARMA-GARCH model with CTS distributed innovations over the backtesting period. As a natural consequence of the construction of AVaR, the line depicting AVaR estimations lies below the VaR line.

Figure 3.18 illustrates the AVaR estimations for the three ARMA-GARCH models used in this analysis, with normal, Student-t and CTS innovations at the 1% AVaR level.

To illustrate the idiosyncratic differences in the AVaR calculations for the different examined models, the following quantity is introduced to measure the difference:

$$d_t = \text{AVaR}_t^{CTS} - \text{AVaR}_t^{CM},$$

where AVaR_t^{CTS} denotes the AVaR which results from applying the CTS-ARMA-GARCH model and AVaR_t^{CM} refers to AVaR from either the normal-ARMA-GARCH model or the t-ARMA-GARCH model. This quantity is illustrated over the backtesting period from Section 3.7.1 in Figure 3.19 for $\eta = 1\%$.

From these empirical comparisons, the following conclusions can be drawn: The CTS-ARMA-GARCH is more conservative in estimating AVaR as the normal-ARMA-GARCH model for $\eta = 1\%$. The t-ARMA-GARCH model leads to slightly more conservative AVaR values as the CTS-ARMA-GARCH model at the 1% level. Additionally, the difference in AVaR resulting from t-ARMA-GARCH models fluctuates less than the difference in AVaR resulting from normal-ARMA-GARCH models, as shown in Figure 3.19.

The findings from Section 3.4 and 3.7.1 state that models based on the normal and the Student-t distribution assumption fail in modeling the innovation process on high-frequency S&P 500 Index time series. This statement is supported by the BLR test results from Section 3.7.1. This indicates that calculating AVaR from normal or Student-t innovation assumptions will most likely lead to unreliable
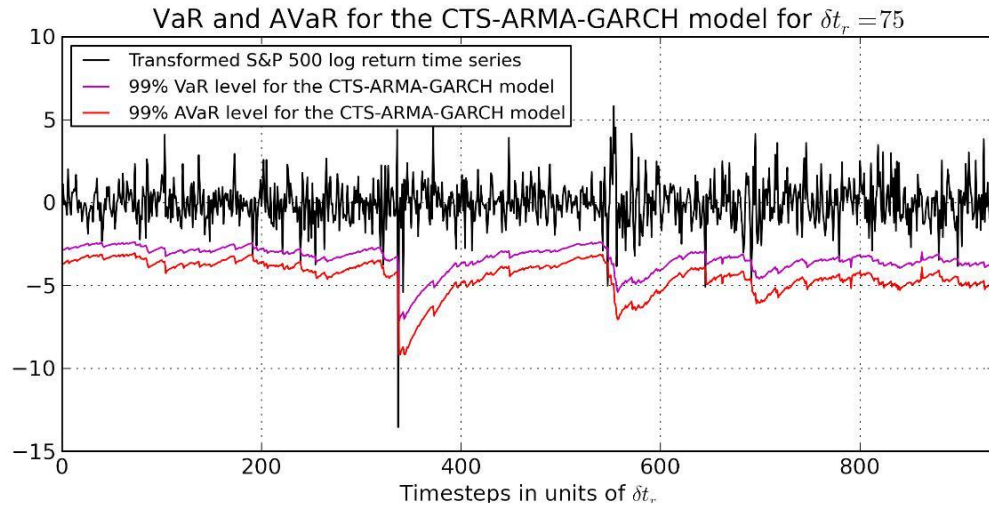
(a) 1% VaR-AVaR CTS, δt_r=75
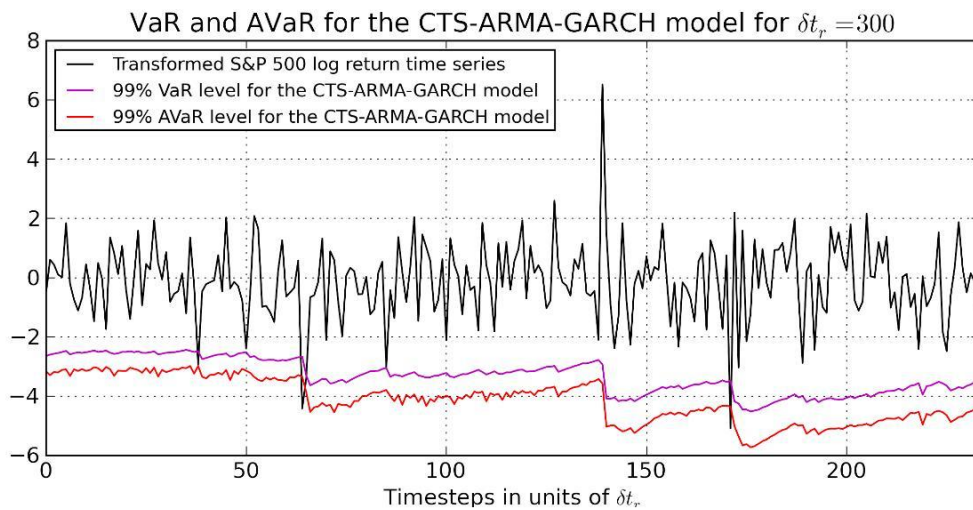
(b) 1% Var-AVaR CTS, δt_r=300

Figure 3.17: AVaR and VaR estimates are depicted in this figure for the backtesting period for the CTS-ARMA-GARCH model at the 1% level for $\delta t_r = 75$ and 300 seconds.

(a) 1% AVaR, δt_r=75
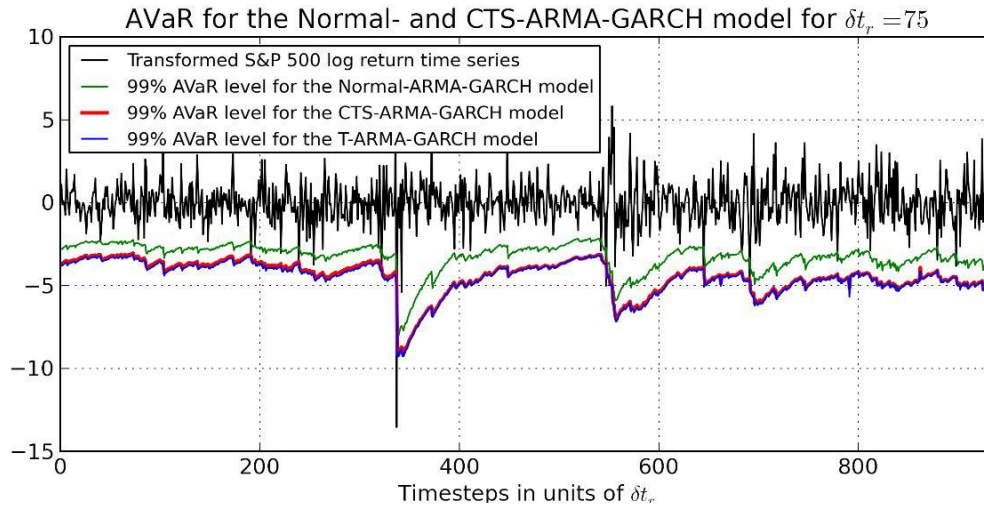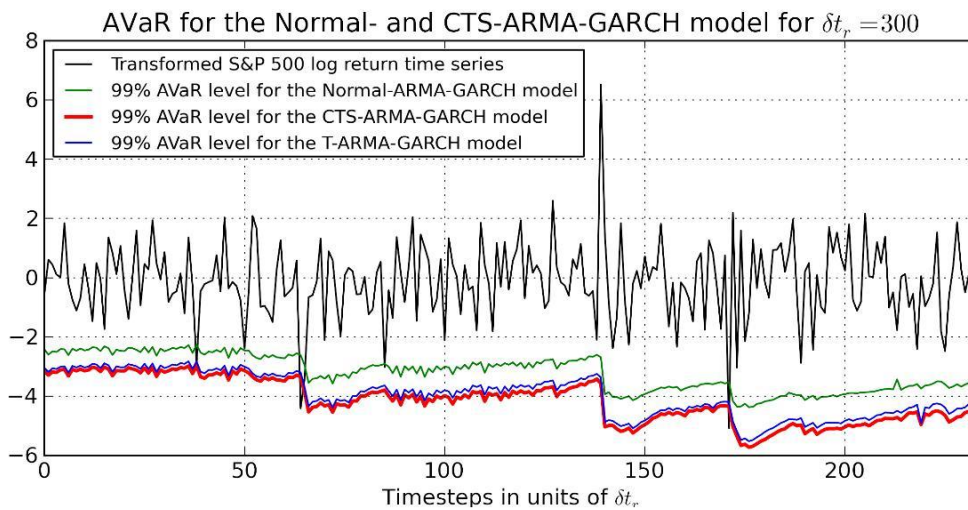
(b) 1% AVaR, δt_r=300

Figure 3.18: Illustration of 1% AVaR for three ARMA-GARCH models with normal, Student-t and CTS distributed innovations over the backtesting period for δt_r=75 and 300 seconds.
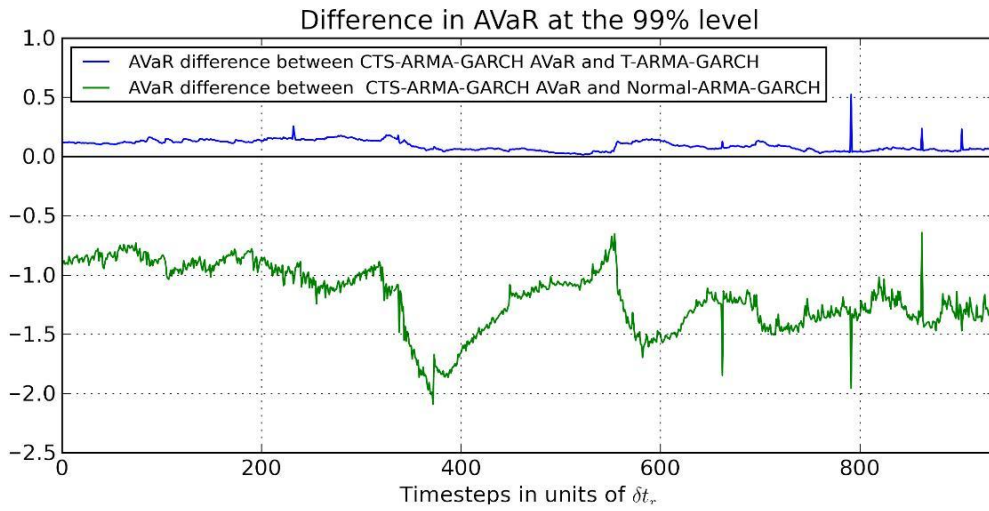
(a) 1% AVaR difference, δt_r=75
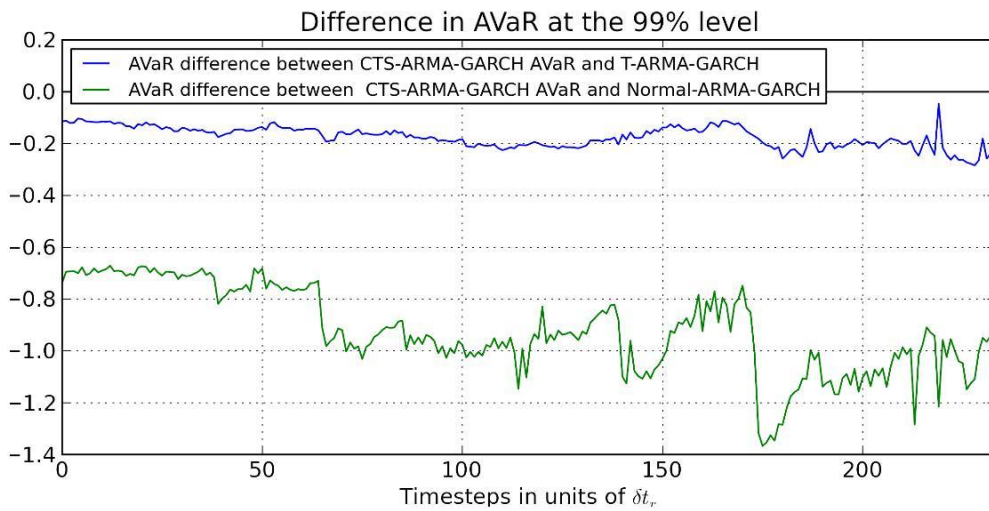
(b) 1% AVaR difference, δt_r=300

Figure 3.19: AVaR differences for the examined models at the 1% level during the backtest period where $\delta t_r = 75$ and 300 seconds.

estimates. This may potentially lead to under- or overestimated capital require-
ments.

δ_{t_r}	Model	w	d	β_1	DOF	α	λ_-	λ_+
15	t	0.174(0.004)	0.281(0.004)	0.132(0.006)	2.037(0.012)	0.100(0.000)	0.981(0.003)	0.971(0.003)
	$\mathcal{N}$	0.177(0.002)	0.206(0.003)	0.120(0.003)				
30	t	0.134(0.005)	0.268(0.006)	0.193(0.008)	2.993(0.031)	0.182(0.026)	1.185(0.024)	1.189(0.023)
	$\mathcal{N}$	0.149(0.003)	0.225(0.003)	0.153(0.005)				
45	t	0.132(0.005)	0.256(0.007)	0.194(0.009)	3.530(0.050)	0.312(0.036)	1.197(0.034)	1.217(0.033)
	$\mathcal{N}$	0.150(0.004)	0.226(0.004)	0.155(0.006)				
60	t	0.132(0.006)	0.251(0.008)	0.200(0.010)	3.976(0.071)	0.302(0.051)	1.299(0.049)	1.333(0.048)
	$\mathcal{N}$	0.152(0.005)	0.227(0.005)	0.172(0.007)				
75	t	0.144(0.007)	0.243(0.008)	0.191(0.011)	4.082(0.082)	0.414(0.058)	1.223(0.055)	1.268(0.054)
	$\mathcal{N}$	0.159(0.006)	0.224(0.006)	0.151(0.008)				
90	t	0.129(0.007)	0.258(0.010)	0.224(0.012)	4.201(0.095)	0.488(0.062)	1.178(0.059)	1.220(0.058)
	$\mathcal{N}$	0.158(0.006)	0.230(0.007)	0.188(0.009)				
105	t	0.131(0.008)	0.254(0.010)	0.217(0.013)	4.404(0.111)	0.553(0.068)	1.159(0.066)	1.214(0.064)
	$\mathcal{N}$	0.167(0.007)	0.219(0.007)	0.159(0.010)				
120	t	0.134(0.009)	0.253(0.011)	0.225(0.014)	4.291(0.112)	0.645(0.070)	1.053(0.066)	1.104(0.065)
	$\mathcal{N}$	0.143(0.007)	0.245(0.009)	0.212(0.010)				
135	t	0.125(0.009)	0.266(0.013)	0.244(0.015)	4.336(0.121)	0.665(0.072)	1.039(0.069)	1.104(0.067)
	$\mathcal{N}$	0.140(0.007)	0.245(0.009)	0.207(0.011)				
150	t	0.115(0.009)	0.280(0.014)	0.256(0.017)	4.353(0.128)	0.677(0.081)	1.029(0.076)	1.101(0.074)
	$\mathcal{N}$	0.135(0.007)	0.248(0.010)	0.189(0.012)				
165	t	0.127(0.010)	0.261(0.014)	0.227(0.016)	4.593(0.148)	0.644(0.089)	1.102(0.087)	1.181(0.085)
	$\mathcal{N}$	0.147(0.008)	0.241(0.010)	0.195(0.012)				
180	t	0.119(0.010)	0.283(0.016)	0.259(0.019)	4.371(0.140)	0.835(0.077)	0.890(0.073)	0.967(0.071)
	$\mathcal{N}$	0.150(0.009)	0.241(0.011)	0.189(0.013)				
195	t	0.121(0.011)	0.280(0.017)	0.257(0.019)	4.446(0.153)	0.587(0.092)	1.146(0.089)	1.189(0.087)
	$\mathcal{N}$	0.135(0.009)	0.257(0.012)	0.214(0.014)				
210	t	0.128(0.011)	0.268(0.017)	0.236(0.019)	4.598(0.167)	0.750(0.095)	1.005(0.091)	1.070(0.089)
	$\mathcal{N}$	0.159(0.010)	0.235(0.012)	0.180(0.014)				
225	t	0.110(0.011)	0.291(0.019)	0.264(0.021)	4.916(0.195)	0.714(0.103)	1.076(0.103)	1.198(0.101)
	$\mathcal{N}$	0.135(0.009)	0.258(0.013)	0.214(0.015)				
240	t	0.114(0.011)	0.289(0.019)	0.260(0.022)	4.829(0.196)	0.602(0.116)	1.179(0.114)	1.264(0.112)
	$\mathcal{N}$	0.140(0.010)	0.260(0.014)	0.209(0.016)				
255	t	0.124(0.012)	0.272(0.018)	0.242(0.021)	4.661(0.189)	0.611(0.120)	1.122(0.116)	1.238(0.114)
	$\mathcal{N}$	0.160(0.011)	0.235(0.012)	0.160(0.014)				
270	t	0.120(0.012)	0.282(0.020)	0.250(0.023)	4.712(0.194)	0.961(0.094)	0.789(0.091)	0.953(0.088)
	$\mathcal{N}$	0.143(0.011)	0.250(0.012)	0.204(0.017)				
285	t	0.124(0.013)	0.282(0.021)	0.255(0.024)	4.663(0.199)	0.793(0.108)	0.959(0.106)	1.084(0.103)
	$\mathcal{N}$	0.131(0.010)	0.253(0.014)	0.230(0.018)				
300	t	0.112(0.012)	0.304(0.023)	0.280(0.027)	4.663(0.204)	0.742(0.116)	0.984(0.112)	1.155(0.109)
	$\mathcal{N}$	0.136(0.011)	0.269(0.016)	0.203(0.018)				

Table 3.5: Estimated parameters for the t-FIGARCH model with Student-t and CTS distributed innovations (denoted as model t) as well as for the normal-FIGARCH model (denoted as model $\mathcal{N}$). The 1σ confidence intervals are shown in brackets.

δ_{t_r}	normal-FIGARCH		t-FIGARCH		CTS-FIGARCH	
	KS(p-value)	AD(p-value)	KS(p-value)	AD(p-value)	KS(p-value)	AD(p-value)
15	0.112 (0.000)	7817.191 (0.000)	0.025 (0.000)	219.927 (0.000)	0.029 (0.000)	83.388 (0.000)
30	0.078 (0.000)	1887.368 (0.000)	0.019 (0.000)	66.201 (0.000)	0.015 (0.000)	5.898 (0.001)
45	0.065 (0.000)	952.055 (0.000)	0.015 (0.000)	25.936 (0.000)	0.010 (0.000)	2.708 (0.039)
60	0.058 (0.000)	565.526 (0.000)	0.012 (0.000)	14.226 (0.000)	0.007 (0.005)	1.327 (0.224)
75	0.056 (0.000)	403.840 (0.000)	0.013 (0.000)	10.261 (0.000)	0.006 (0.056)	1.157 (0.285)
90	0.058 (0.000)	326.683 (0.000)	0.010 (0.001)	6.598 (0.001)	0.006 (0.122)	1.844 (0.112)
105	0.053 (0.000)	259.189 (0.000)	0.009 (0.003)	6.310 (0.001)	0.006 (0.156)	1.313 (0.228)
120	0.055 (0.000)	248.378 (0.000)	0.009 (0.008)	6.754 (0.000)	0.005 (0.426)	0.592 (0.655)
135	0.055 (0.000)	200.911 (0.000)	0.010 (0.009)	3.910 (0.010)	0.003 (0.929)	0.523 (0.724)
150	0.054 (0.000)	174.587 (0.000)	0.013 (0.000)	3.368 (0.018)	0.005 (0.485)	0.709 (0.551)
165	0.052 (0.000)	159.455 (0.000)	0.013 (0.001)	3.794 (0.011)	0.004 (0.802)	0.608 (0.641)
180	0.054 (0.000)	146.739 (0.000)	0.010 (0.024)	1.546 (0.166)	0.005 (0.644)	0.557 (0.690)
195	0.053 (0.000)	131.298 (0.000)	0.010 (0.049)	1.936 (0.100)	0.004 (0.956)	0.338 (0.907)
210	0.051 (0.000)	120.062 (0.000)	0.009 (0.107)	2.219 (0.070)	0.004 (0.980)	0.257 (0.967)
225	0.048 (0.000)	96.496 (0.000)	0.012 (0.021)	2.675 (0.040)	0.004 (0.966)	0.272 (0.958)
240	0.050 (0.000)	88.646 (0.000)	0.009 (0.125)	2.321 (0.062)	0.004 (0.979)	0.310 (0.931)
255	0.053 (0.000)	96.773 (0.000)	0.015 (0.004)	2.343 (0.060)	0.006 (0.565)	0.436 (0.813)
270	0.052 (0.000)	81.347 (0.000)	0.017 (0.001)	3.159 (0.023)	0.006 (0.618)	0.649 (0.603)
285	0.049 (0.000)	77.757 (0.000)	0.013 (0.017)	1.551 (0.165)	0.004 (0.972)	0.556 (0.691)
300	0.051 (0.000)	78.732 (0.000)	0.014 (0.014)	4.496 (0.005)	0.005 (0.878)	0.291 (0.944)

Table 3.6: Goodness-of-Fit summary for the KS and AD tests for three FIGARCH models. The models were applied on 20 time series. The respective p-values are shown in brackets.

Parameter	Slope	Slope t-stat.	Intersect	Intersect t-stat.	R^2
CTS-Parameters					
α	0.002(0.00053)	7.4	0.3(0.085)	6.9	0.5
Δ	-0.00044(0.00071)	1.2	-0.0037(0.11)	0.065	0.96
λ	-0.0012(0.00036)	6.4	1.3(0.057)	45	0.38
FIGARCH-Parameters					
w	$-0.00018(5 \cdot 10^{-5})$	6.8	0.16(0.0076)	41	0.27
d	$-5.5 \cdot 10^{-6}(7 \cdot 10^{-5})$	0.15	0.27(0.0094)	56	0.00047
β	$0.000051(8.6 \cdot 10^{-5})$	12	0.13(0.012)	21	0.42

Table 3.7.: The estimated linear regression parameters for the CTS-FIGARCH parameter dependencies on δt_{tr} are shown in the above table. The 95% confidence levels for the parameter estimations are shown in brackets.

δ_{t_r}	Model	UC 99%	ID 99%	BLR 95%	BLR99%
75	CTS	0.156	0.295	0.441	0.506
	normal	0.000	0.000	0.000	0.000
	Student-t	0.048	0.106	0.092	0.129
90	CTS	0.010	0.025	0.055	0.030
	normal	0.000	0.000	0.000	0.000
	Student-t	0.010	0.025	0.014	0.022
105	CTS	0.000	0.000	0.064	0.027
	normal	0.000	0.001	0.000	0.000
	Student-t	0.005	0.015	0.035	0.015
120	CTS	0.951	0.938	0.100	0.980
	normal	0.001	0.004	0.000	0.000
	Student-t	0.951	0.938	0.019	0.954
135	CTS	0.253	0.459	0.115	0.091
	normal	0.004	0.011	0.000	0.000
	Student-t	0.061	0.141	0.002	0.089
150	CTS	0.161	0.326	0.003	0.911
	normal	0.161	0.326	0.000	0.000
	Student-t	0.883	0.937	0.000	0.933
165	CTS	0.722	0.884	0.013	0.699
	normal	0.044	0.005	0.000	0.000
	Student-t	0.422	0.665	0.008	0.427
180	CTS	0.322	0.558	0.185	0.210
	normal	0.009	0.027	0.001	0.001
	Student-t	0.322	0.558	0.100	0.393
195	CTS	0.835	0.936	0.011	0.971
	normal	0.045	0.111	0.012	0.017
	Student-t	0.484	0.729	0.002	0.680
210	CTS	0.725	0.895	0.122	0.784
	normal	0.079	0.185	0.000	0.000
	Student-t	0.725	0.895	0.036	0.864
225	CTS	0.006	0.012	0.109	0.543
	normal	0.325	0.568	0.000	0.000
	Student-t	0.945	0.969	0.126	0.331
240	CTS	0.963	0.968	0.067	0.771
	normal	0.113	0.251	0.000	0.000
	Student-t	0.548	0.789	0.066	0.686
255	CTS	0.478	0.733	0.512	0.952
	normal	0.010	0.028	0.000	0.000
	Student-t	0.478	0.733	0.156	0.770
270	CTS	0.808	0.937	0.079	0.905
	normal	0.419	0.677	0.008	0.001
	Student-t	0.808	0.937	0.119	0.588
285	CTS	0.738	0.911	0.413	0.935
	normal	0.055	0.048	0.021	0.028
	Student-t	0.761	0.939	0.568	0.843
300	CTS	0.322	0.571	0.487	0.493
	normal	0.045	0.042	0.001	0.000
	Student-t	0.045	0.042	0.046	0.131

Table 3.8: The results from the Value-at-Risk backtest are shown in this table.

Chapter 4

Conclusion

Following a brief theoretical introduction to financial econometrics, the empirical analysis in this first part of the thesis encompassed a thorough investigation of the S&P 500 Index on high-frequency time scales.

Log-return time series were constructed from the highest possible frequency (15 seconds return time interval) up to 300 seconds. In order to apply conditional mean and variance models, the characteristic intraday volatility profile has been removed from the time series.

After a detailed examination of the prevailing stylized facts on these time series, conditional mean and variance models have been applied. Here, the focus was on ARMA(1,1)-GARCH(1,1) and FIGARCH(1,d,0) models. Three innovation distribution assumptions, the normal, the Student-t and the Classical Tempered Stable, have been employed in the framework of each model.

The assumption that the normal distribution is not capable of modeling the intraday innovation process was empirically proven in all tested cases. The CTS innovation assumption showed in both approaches, the ARMA-GARCH and the FIGARCH, the best descriptive properties.

This conclusion was empirically supported by comparing the respective p-values in statistical tests, such as the Kolmogorov-Smirnov and the Anderson-Darling tests.

On the basis of the time series on return time intervals between 15 seconds and 300 seconds, potential parameter dependencies were further investigated. This was carried out for both the CTS-ARMA-GARCH model as well as for the CTS-FIGARCH model. In both cases, the CTS parameters showed a similar behavior. Especially the parameter α increased with a growing return time interval, indicating that the heavy tailedness of financial data increases with shorter time scales. The investigation of the FIGARCH parameters led to the insight that the conditional variance process in the S&P 500 Index exhibits effects of long-range-dependence. This conclusion is indicated by the significance of the model param-

eter d that expresses the degree of long-range-dependence.

Following this, a Value-at-Risk backtesting was carried out in order to compare the risk forecasting capability of the three employed innovation distributions. Here, the ARMA(1,1)-GARCH(1,1) framework was chosen in accordance with my publication Beck *et al.* (2011).

The VaR backtesting was carried out on roughly 3000 high-frequency index values, which is comparable to roughly 12.5 years in terms of daily return data. Based on the Christoffersen-Likelihood-Ratio test and the Berkowitz-Likelihood-Ratio test, the findings suggest that neither the normal nor the Student-t assumptions lead to reliable Value-at-Risk estimates on short intraday time scales. The tempered stable innovation assumption, however, was accepted in most cases. In accordance with this finding, modeling high-frequency S&P 500 Index log-return time series with tempered stable innovation distributions leads to significantly improved VaR forecasts compared to the normal- or the Student-t innovation assumption.

The Berkowitz test results suggested that the CTS-ARMA-GARCH model exhibits the most reliable tail modeling capability among the three examined innovation distribution assumptions. This indicates that Average Value-at-Risk measures will be most reliably forecasted with the tempered stable assumption.

Finally, the risk measure Average Value-at-Risk was provided for ARMA-GARCH models with tempered stable innovations for the same time period as in the Value-at-Risk backtest. This measure was compared to Average Value-at-Risk values derived from ARMA-GARCH models with normal and Student-t innovations. By investigating the difference in Average Value-at-Risk between the distinct models, it turned out that both the models with normal and the Student-t innovations do not lead to reliable estimates of Average Value-at-Risk.

Part II

Limit Order Execution Probabilities

Chapter 5

Introduction

Stock trading around the world takes place on dedicated trading venues (mainly stock exchanges) and is nowadays mostly performed electronically. Even though each country provides an individual legal framework for the exchange operators and the trading of stocks,[1] the basic principles of the procedures on trading platforms are very similar. The general purpose of a stock exchange is to form a market place around potential buyers and sellers. A third party which is apparent on most trading platforms are market makers. They are important participants in the trading process in that they are obliged to provide liquidity to the market by continuously quoting buy and sell offers.

The focus in this study is on markets that enable stock trading on the basis of open limit order books. There exist two fundamental mechanisms in order to participate in the stock trading process. One way is through submitting market orders which are immediately executed (given enough liquidity in the market). In the case of buy (sell) market orders, the trader pays (receives) the currently available market price.

A second way to trade a stock is through placing passive limit orders. This type of order signals the willingness to buy or sell a stock for a certain price. The limit order may eventually be executed against an incoming market order but the execution is not guaranteed. The obvious drawback of limit orders is the uncertainty of execution. Passive limit orders are stored in the *order book* until they are executed or deleted.

As a consequence, the choice between trading by market versus trading by limit is one of the main issues in developing trading strategies (Cho and Nelling (2000)).

The decision whether to place a limit or a market order is severely influenced

---

[1]The analysis in the following are conducted on the UK and the French Stock Exchanges. In accordance to that, the Financial Services Authority (FSA) regulates and controls the UK financial markets. In France, the responsible institution is the Autorité des marchés financiers (AFM).

by the probability of the execution of a limit order. Recent research on limit orders involves the investigation of the economic role of limit orders in electronic order books, such as, for example, Glosten (1994). Furthermore, order placement strategies have been extensively studied by, for example, Foucault (1999) and Harris and Hasbrouck (1996). Biais *et al.* (1995) provide an empirical analysis of the order flow at the Paris Bourse. Earlier studies attempting to derive models for limit order execution probabilities or, closely related, modeling limit order's *survival times* involve Omura *et al.* (2000); Cho and Nelling (2000) and Lo *et al.* (2000).

Trading strategies that are attempting to profit from intraday price movements are less concerned about the execution of a limit order until the end of the trading day. Particularly important for these strategies is the question whether a limit order is going to be executed within a certain time interval. One obvious approach for tackling this issue is by finding a mathematical model capable of calculating execution probabilities for limit orders for a given time horizon. In order to produce exact results, such an approach will have to take prevailing stock market characteristics as well as intrinsic order related quantities into account.

The goal of this thesis is to present an algorithm capable of forecasting conditional execution probabilities for limit orders on short time intervals. The approach chosen is a combination of two techniques. First, a basic model is derived containing the average execution probabilities for single stocks and their order book levels. This covers the idiosyncratic unconditional execution probabilities for stock-level pairs. Following on from this, a non-parametric approach (the machine learning algorithm NeuroBayes®) is applied which computes execution probabilities conditional on variables such as volatility, order book asymmetries, daytime, etc. Finally, the two models are combined and applied on two stock exchanges, the London Stock Exchange (LSE) and the Euronext Paris Stock Exchange (EP).

The second part of this thesis is organized as follows. First, the principles of stock trading are explained and the data set, on which the derived models are constructed and verified, is introduced. Subsequently, NeuroBayes® is introduced together with its operating principles. Following on from this, the basic model, denoted as Mean-Matrix-Model is derived. The application of NeuroBayes® is discussed and a selection of variables which exhibit strong dependencies on the execution probability are further examined and discussed. The combination of both models is presented and its forecasting capability is examined in an out-of-sample test. Following this, the phenomenon of a fractional execution of limit orders is further discussed and forecasted with NeuroBayes®.

Chapter 6

Financial Markets and Data Sources

6.1 Stock Market Mechanics and Market Microstructure

This section introduces the mechanics of the *open limit order book* (denoted in the following as *order book*) and discusses the rules and processes by which *order-driven-markets* operate. First, the concept of market(-able) and limit orders is introduced in a more comprehensive fashion.

As briefly stated in the introduction, market participants have two fundamental possibilities in order to participate in the trading process:

- Through passive *limit orders*
  Limit orders signal the intention to trade at a specific price. The limit order is a binding offer as long as it remains in the order book. This type of order may eventually be executed through an incoming marketable order. Otherwise, it can be canceled on the investor's demand. This order type is in the following denoted as *limit order* or just *order*.

- Through aggressive *market(-able) orders*
  A marketable order is either a market order or a bid (ask) limit order with a limit price that is above (below) the currently best ask (bid) price. Given enough liquidity, a marketable order is immediately filled at the currently best available price.

Each limit order submitted to trade a stock consists of at least four basic pieces of information: the intention to buy or sell, the limit price, the number of shares which are intended to trade and the exact time stamp, when the order arrived. The latter is crucial in order to determine which market participant's order is privileged when there exist two or more limit orders with the same limit price (this rule is known as *price-time-priority*).

In order to keep track of each market participants' intentions to buy or sell a specific stock, their *limit orders* are stored in two lists, one for potential buyers (the bid-side) and one for potential sellers (the ask-side). The bid-side is sorted in a descending order (the highest offer comes first) whereas the ask-side is sorted in an ascending order (the cheapest offer comes first). This particular order corresponds to the succession in which limit orders are executed when marketable orders arrive.

These two ordered lists for buy- and sell-orders form the *limit order book* (Cont *et al.* (2008)). Usually, limit orders with equal prices are aggregated in *order book levels*. Each level is defined through its price, the number of pending orders and the accumulated number of shares.

Once a marketable order arrives it is matched against the existing orders in the order book. As a natural consequence, the buyer's and seller's order lists never overlap (except during auctions). The price gap which consequently exists between the best-bid price and the best-ask price is called *bid-ask-spread* c_t at time t:

$$c_t = S_t^a - S_t^b, \tag{6.1}$$

where S_t^a denotes the best-ask price and S_t^b the best available bid-price at time t.

As a consequence, marketable orders are said to take liquidity whereas limit orders provide liquidity. The execution of a marketable order differs in price compared to the execution of a limit order by at least the bid-ask-spread. Additionally, the respective costs incurred may differ between marketable and limit orders since some exchange offer a rebate for providing liquidity.

Limit orders can be equipped with additional features such as, for example, a validity time or a trigger for automatic cancelation. The scope in which those options are available differs among exchanges.

Limit orders can be subject to partial execution. That is, if the size of a limit order exceeds the size of an arriving marketable order, only the overlapping amount of shares will change hands. The partially filled limit order will become modified to the remaining size while its priority status is preserved.

Limit prices cannot be chosen from a continuous scale but rather from a discrete price-grid. The minimal possible difference between two adjacent allowed price levels is denoted as *tick size* (TS) and generates a quantization of the price process. The tick size depends on the price at which a specific stock trades and can be found in designated tables which are provided from the exchange operators. The minimum possible bid-ask-spread for a certain stock corresponds to the stock's tick size.

Markets which operate in the fashion as described above are denoted as *order-driven-markets*. An illustration of the underlying market mechanics is presented in Figure 6.1 (Smith *et al.* (2003)). In addition to that market type, there exist
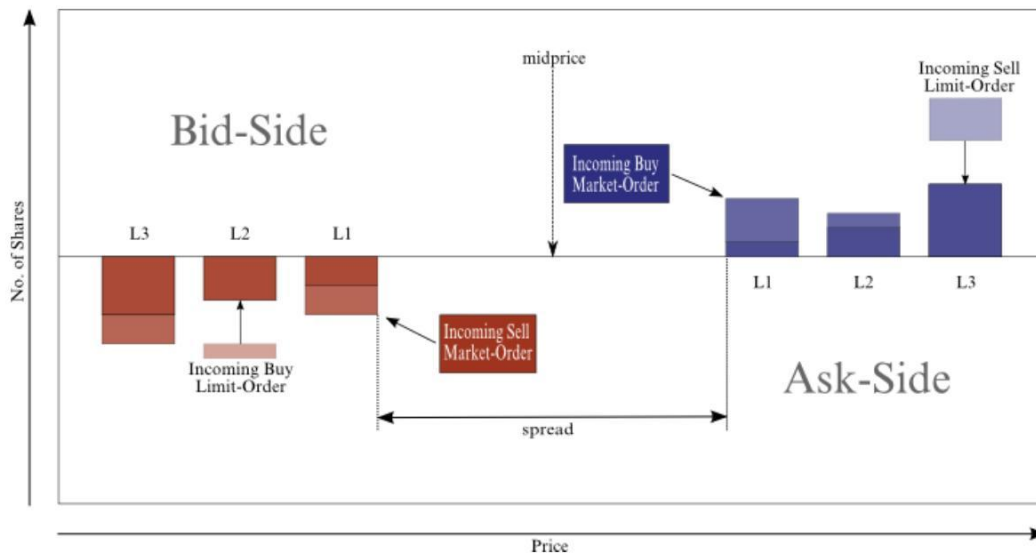
Figure 6.1: The above figure illustrates the mechanics of how an order-driven market operates. The red boxes on left hand side visualize bid-orders whereas the blue boxes on the right hand side represent ask-orders. The boxes' height implies the order size while the respective color's intensity indicates the order's priority. Arriving limit orders have a lower priority than existing orders. Incoming marketable orders are illustrated in the center region of the chart. They are matched against pending limit orders. Additionally, the mid-price and spread are indicated.

quote-driven-markets which are based on specialists through which the trading is centralized (Handa and Schwartz (1996)).

Time series created from this most disaggregated level face the question of how a price process is to be constructed, since several price measures are available (bid-prices, ask-prices, transaction prices). One commonly utilized compromise is to employ the *mid-price*, constructed as the mean between the best-bid and the best-ask price:

$$S_t^m = \frac{S_t^a + S_t^b}{2}. \tag{6.2}$$

Many market places (such as the London Stock Exchange and the Euronext Paris Stock Exchange) open and close the trading day with auctions in which the resulting price is determined such that the traded number of shares is maximized (see, for example, Deutsche Boerse Group (2010); Beltran-Lopez and Frey (2006); NYSE Euronext (2011)). The period in between the auctions is called *continuous trading period* through which market participants can insert, modify and delete their limit orders at any time. The following analysis on limit order executions is carried out during the continuous trading phase.

6.2 Data Sources and Historic Order Book Construction

The following analyses are based on the knowledge of the order book at all times.[1] In order to construct historic order books from recorded data, a tick-by-tick data record with all order book relevant information is required. There exist two common data formats in which order book information is published:

- *Market-By-Limit (MBL)*: The aggregated level information is submitted as soon as anything changes on a level (and those levels affected by this change). This corresponds to a status update.

- *Market-By-Order (MBO)*: Each individual *order action* is reported. This applies for order inserts, modifications and cancelations. That is, all changes in the order book are reported individually.

Obviously, the MBO format offers a more detailed picture about the microscopic procedures on a stock market. However, compared to the MBL format it is more prone to errors. As a consequence, some exchanges only offer their order book information in the MBL format (such as the Deutsche Boerse) whereas other exchanges offer both data formats (such as NYSE-Euronext).

Here, the focus is on order books constructed from MBO data. The analyzed stock universe consists of a selection of the most liquid stocks from the French and British stock exchanges. Appendix B.3 lists all stocks which were utilized. The time span over which the order information is analyzed is 6 months (Aug 2010 - Feb 2011), leading to a collection of millions of limit orders.

High data quality plays a crucial role when analyzing order book data since missing data, especially in the MBO format, cannot be interpolated or reproduced. In the worst case scenario, the loss of a few seconds of tick-by-tick data can cause that the remainder of a trading day cannot be utilized any more.[2] Hence, the data set on which analyses like the following are conducted, has to be carefully chosen.

6.2.1 Construction of Additional Order Related Information

The basic information delivered when a limit order is inserted into the order book is the order type (buy or sell), the price limit, the number of shares intended

---

[1]A software tool was developed in the scope of this thesis in order to construct and analyze order books.

[2]Imagine, for example, that the cancelations of a few limit orders near the mid-price were missed. This can easily lead to negative spreads and wrong mid-prices in the following order book development.

to trade and the arrival time. Additional order related information can be constructed from the knowledge of the structure of the order book at the time of the order insertion. However, some of the information can only be computed in an ex-post analysis of the data. The following order-specific information has been collected for the later analysis:

1. Whether the limit order has been canceled by the owner's intention or if it has been executed. This also provides information whether a trade was buyer- or seller initiated. Here, it is not distinguished between a partial and a full execution of a limit order.

2. Conditional on the execution, the traded fraction of the limit order.

3. The lifetime of limit orders with millisecond precision.

4. The level and the position[3] within the level in which the order has been inserted.

5. The level and the position within the level from which the order has been canceled.[4]

6. The number of shares which are in the order book with a higher price-time-priority than a newly inserted limit order.

7. The value of the shares which are in the order book with a higher price-time-priority than a newly inserted limit order.

Information 4) - 7) can easily be constructed from the current state of the order book at the time of the order action. Limit order lifetimes are received by tracking each order through its path through the order book. The lifetime is then received by subtracting the arrival time from the time of cancelation. The question whether a limit order has been canceled by the trader's intention or by being executed through an arriving marketable order requires the most attention, mainly due to technical obstacles. The method used in this thesis is outlined in Appendix B.1.

Additional Data Sources

In addition to the order specific information, a data source has been employed which contains information related to continuous trading characteristics, aggregated on a one minute time grid. The quantities stored in this dataset include

---

[3]Here, the position denotes the number of pending limit orders in the same order book level with a higher execution priority plus one.

[4]In the case when the order is executed, the limit order must have been canceled from the top position in the order book.

volatility measures, OHLC[5]-information for each minute, the number of trades over the past minute (split into buys and sells), mid-price changes and many further variables. The data source is constructed such that the quantities are computed precisely when a new minute begins. The trigger to process the past minute is the first incoming information of the following minute. If nothing happens during one minute, the respective quantities are filled with pre-defined null-values.

In addition to the quantities aggregated over past minutes, variables are constructed which describe the state of the order book precisely at the minute change. This information includes the bid-ask-spread, the population of the first bid- and ask levels, bid- and ask prices and the tick size.

Additional variables have been constructed which reflect prevailing asymmetries in the order book such as the quantity *Volume-Weighted-Mid-Price* ($VWMP$).[6] Two different versions of the $VWMP$ are introduced in this thesis, denoted as $VWMP_t$ and $VWMP_s$. They are defined as follows:

$$VWMP_t = \frac{\sum_{j=0}^{1} \sum_{i=1}^{10} e^{-i} v_{ji} p_{ji}}{\sum_{j=0}^{1} \sum_{i=1}^{10} e^{-i} v_{ji}} \tag{6.3}$$

$$VWMP_s = \frac{1}{2} \left(\sum_{j=0}^{1} \frac{\sum_{i=1}^{10} e^{-i} v_{ji} p_{ji}}{\sum_{i=1}^{10} e^{-i} v_{ji}} \right). \tag{6.4}$$

Here, the index j denotes the order book side ($j = 0$: bid-side, $j = 1$: ask-side) and the index i the order book level. v_{ij} stands for the number of shares in the ith level on the jth order book side whereas p_{ij} indicates the corresponding price level. The two quantities are comparable to a center of mass estimation where the outer levels have a smaller weight than the inner levels. $VWMP_t$ calculates this measure by taking both order book sides into account at once, whereas $VWMP_s$ constructs each order book side's *center of mass* individually and finally computes the mean of both.

6.2.2   Collection of Limit Orders

The collection of limit orders which is used in the following analysis is presented here. In order to relate limit orders to the information which is available on the one minute time grid, as described above, only limit orders are selected which

---

[5]OHLC stands for Open-High-Low-Close.

[6]The mid-price, a frequently applied measure, does not account for asymmetries within the order book and is, depending on the application, not necessarily a meaningful quantity. Imagine a situation where a very small-sized limit order is placed very aggressively in the order book. This would shift the mid-price to a higher or lower position even though the limit order should not have a significant economic impact on the price process due to its small size.

have been inserted into the order book within one second after a new minute began. By that, the state of the order book and the collected trading data from the previous minute will have a high probability to show significant contributions in statistical analysis.

In order to derive a model to forecast whether a limit order will be executed within one minute, the set of limit orders is filtered further by accepting only those limit orders which fulfill the following conditions: (1) The limit order has been executed within one minute, or (2) in the case it was not executed through an arriving marketable order, it remained in the order book for at least one minute. This selection is reasonable if the decision process of a trading strategy to cancel or modify limit orders is on a one minute time grid as well. Letting the above filtration aside, the resulting data set would contain a lot of noise, especially due to ultra-high-frequency limit order actions. This is indicated in Figure 6.2 which depicts the lifetime distribution of those limit orders which have been canceled on the trader's demand. The significant peaks in the left corners of these charts visualize the increased frequency of limit orders with a lifetime of less than 1 second.

Only limit orders are considered which have been inserted up to the 5th order book level. Additionally, all stocks in the collection are represented with roughly the same amount of limit orders to avoid drawing conclusions from only a subset of the analyzed stocks. Finally, the number of bid and ask orders is well balanced intrinsically. This feature is preserved in the filtered data set. Limit orders satisfying the above conditions are collected for the French Stock Exchange in Paris (resulting in 1.34 Mio limit orders) and the UK Stock Exchange in London (resulting in 2.04 Mio limit orders). This results in approximately 30,000 limit orders per stock on both exchanges.

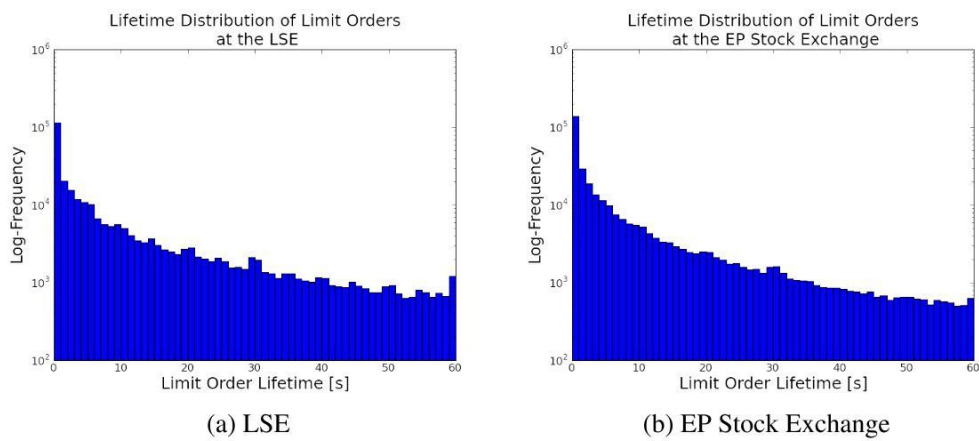(a) LSE   (b) EP Stock Exchange

Figure 6.2: The distribution of limit order lifetimes up to 60 seconds is shown in the above charts for the LSE and the EP Stock Exchange. The bin width in these histograms corresponds to one second. Note the logarithmic y-scale, thus the frequency of limit orders with a life time below one second strongly exceeds the frequencies on longer lifetimes.

Chapter 7

NeuroBayes® Technology

NeuroBayes® is a statistical software package, developed and maintained by Phi-T®, a research and development company.[1] It is constructed in order to solve multivariate statistical inference problems in a robust and efficient fashion (Feindt (2004)). Its main target is to carry out predictions under uncertainty, conditional on a set of endogenous variables. The predicted quantity is referred to as *target variable*. NeuroBayes® belongs to the class of supervised learning algorithms. That is, the algorithm's parameters are adjusted on a data sample where the target information is provided. This process is known as *training*. Subsequently, the *expertise* which results from the training is used in order to predict the target variable in cases where its actual value is not known. The field of applications for NeuroBayes® is manifold. It is used in high-energy physics experiments as well as for various business intelligence problems.

The focus in this section is on *classification problems*, that is, binary problems in which the target variable is either true or false. In mathematical terms, the conditional probability

$$p(t = 1|\mathbf{x}) \tag{7.1}$$

is predicted, where $\mathbf{x}$ denotes the vector of endogenous variables.

A typical classification problem for a financial institution could be, whether a bank loan will be fully repaid or not. NeuroBayes® is constructed such that it will produce the probability of the target event to occur. In the case of bank loans, NeuroBayes® would, based on historical customer information such as annual income, crime records, etc, compute the conditional probability that a loan to this customer will not be fully repaid.

The mechanisms by which NeuroBayes® operates are presented in this section. As stated above, NeuroBayes® detects patterns in data samples of variables

---

[1] This company was founded by Professor M. Feindt, Institut fuer Experimentelle Kernphysik, Karlsruhe Institute of Technology (KIT).

which are in some way related to the target variable. The endogenous variables on which the predictions are based are referred to as *input variables*. Generally, large data sets are required in order to construct the expertise, that is, to *learn* the relevant information from the input variables provided. A data set consists of instances denoted as *events*, that is, entities which contain values for each input variable as well as the target variable's value. If the target variable is true, the event is said to contain *signal*, otherwise, it contains *background*. Two important measures in this context are the *signal purity* $\mathcal{P}$ and the *signal efficiency* $\mathcal{S}$. The signal purity measures the proportion of selected signal compared to the total number of selected events. It is given by the following expression:

$$\mathcal{P} = \frac{N_{\text{signal}}}{N_{\text{signal}} + N_{\text{background}}}. \tag{7.2}$$

The signal efficiency denotes the proportion of a number of selected signal events to the overall number of signals:

$$\mathcal{S} = \frac{N_{\text{signal selected}}}{N_{\text{signal total}}}. \tag{7.3}$$

A crucial step in order to construct a robust algorithm is to transform the input variables in a way such that they are available in a pre-defined format, which allows for an efficient further processing. This preliminary step is known as *preprocessing* and is described in the following.

7.0.3   Preprocessing

Several obstacles arise when a set of raw input variables is to be compared and combined in order to produce a prediction. First, the values of distinct input variables may be present on hugely different numeric scales. This may eventually lead to numerical problems when those variables are to be combined. Additionally, outliers may be present within the set of detected values for one variable. Outliers are extremely disturbing because they can potentially overshadow information and dependencies which are contained in the bulk of the observed values. Frequently utilized measures, such as mean values or correlation coefficients are not robust against outliers (Blobel and Lohrmann (1998)). Furthermore, the relation between the values within one variable to the target usually contains statistical fluctuations which have to be smoothed out.

In order to have a set of input variables $\mathbf{x} = \left(x^{(1)}, x^{(2)}, ..., x^{(N)}\right)$ free of these disturbing features, each input variable is treated separately in a first preprocessing stage.

Generally, variables can exhibit fundamentally different characteristics which have to be considered individually. For example, input variables can be

- *continuous* ($x \in \mathbb{R}$). For example, a volatility measure on a certain stock.

- *discrete and unordered* ($x \in S, S = \{S_i | i \in 1, ..., N\}, N \in \mathbb{N}$), where the order of i does not indicate a certain sequence. For example, S_i could denote a color.

- *discrete and ordered* ($x \in S, S = \{S_i | i \in 1, ..., N\}, N \in \mathbb{N}$), where the order of i indicates a certain sequence. For example, S_i could denote an order book level.

- contain *delta values*: A variable may potentially be undefined in some events. In this case, the variable is set to a pre-defined *delta value*, which is recognized by the algorithm and treated separately.

The pre-processing is carried out for both discrete and continuous input variables.

7.0.4   Selecting Input Variables

A general approach when using NeuroBayes® for predictions is to provide the algorithm in a first step with as much information as possible. The algorithm is constructed such that it will automatically detect which input variables contain statistically irrelevant or redundant information in accordance with a pre-defined cut-off criterion. Those input variables will be ignored in the subsequent processes. This step is especially important in order to avoid over-training, that is, to prevent the algorithm from learning statistical fluctuations. Additionally, selecting only those variables which carry the most relevant information is crucial in order to construct a time efficient algorithm.

7.0.5   Final Classification

Once the input variables have been preprocessed, their individual correlation to the target is combined. The final result from the algorithm is a probability of the target event to occur. Depending on this probability and some user defined loss function, an event is classified to contain either signal or background. Taking a low probability as the threshold for classifying an event as signal will lead to a high signal efficiency. At the same time, the signal purity will remain rather low, because a significant number of actual background events will be classified as signal events as well. Adjusting the classification threshold to a high probability will cause the inverse effect. Only few signals will be selected resulting in a high signal purity.

The actual threshold depends on the application and the damage or loss that either of the above mentioned errors will cause.

7.0.6 Investigating the Algorithm's Forecasting Capability

Before an algorithm is used in practical applications, its true forecasting capability should be thoroughly examined. For this purpose, there exist a number of graphical and quantitative tools which are presented in this section.

Signal-Purity Plot

A typical visual tool which is applied in order to identify an algorithm's separating quality is the *Signal-Purity plot*, depicted in Figure 7.1. It measures the observed signal purity as a function of the forecasted probability. The probability range over which the data points in this chart are spread is an indicator for the separating capability between signal and background events. For the output of the NeuroBayes® algorithm to be interpreted as a probability, this dependency must coincide with a diagonal through the origin with slope one.

Signal-Background Histogram

Signal-Background histograms are frequently applied measures in order to visualize an algorithm's separating capability. This chart-type depicts signal and background events separately, as shown in Figure 7.2. The red curve visualizes the distribution of signal events whereas the black curve stands for events containing background. The more the two curves appear separated, the more pronounced is the algorithm's separating capability.

Lorenz-Curve and Gini-Coefficient

The Lorenz-Curve is constructed in order to measure the signal efficiency $\mathcal{S}$ in dependence of the data efficiency $\mathcal{S}_D$. The data efficiency denotes the fraction of the events under consideration which are sorted in a descending order in accordance to their assigned probability p. To be more precise, the data efficiency $\mathcal{S}_D$ is calculated by

$$\mathcal{S}_D = \frac{\# \text{ events selected}}{\text{total number of events}}, \tag{7.4}$$

where the dataset is sorted such that

$$D = \{E_1, E_2, E_3, ..., E_k | p_1 \geq p_2 \geq ... \geq p_k\}. \tag{7.5}$$

E_i denotes a single event and p_i is the corresponding assigned probability of this event's target to be true. Due to the order in which the events occur in the dataset, it is obvious that the signal efficiency must follow a monotonously increasing function. The curve which depicts this relation is indicated as a blue line in Figure 7.3.
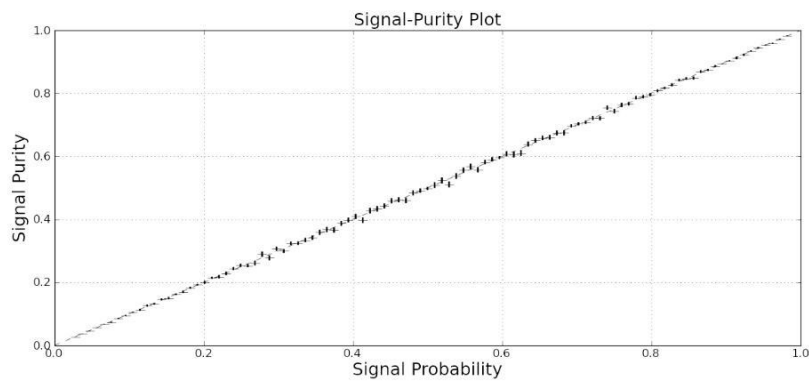
Figure 7.1: The *Signal-Purity plot* is depicted in this figure. It shows the signal purity in dependence of p, the estimated probability resulting from the NeuroBayes® algorithm. In order to interpret this measure as a probability, it has to be ensured that the resulting data points are compatible with a diagonal through the origin as shown in this case.
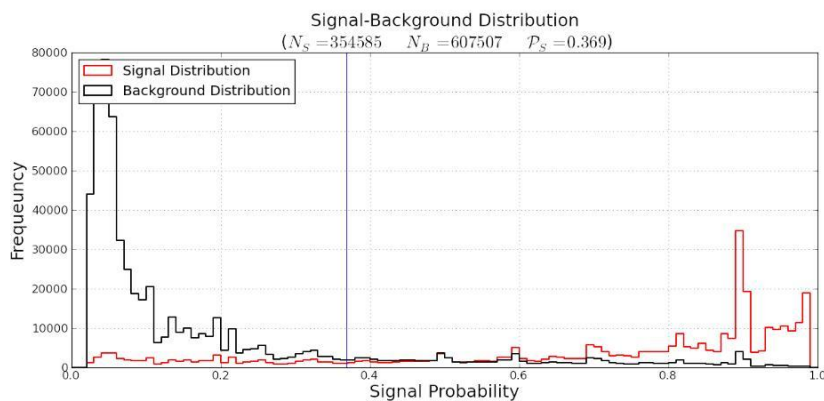
Figure 7.2: The figure shows the *Signal-Background histogram*. The red line indicates the signal distribution whereas the black line shows the background distribution. An indicator of the forecasting capability of the algorithm is how well the background and signal histograms are separated.

The lower black line which confines the white area indicates the lower boundary for any algorithm's performance, i.e. where no distinction between signal and background can be made. In contrast, the upper black line indicates the best performance achievable, that is, a perfect separation between target and background events. The realized Gini-coefficient is constructed by dividing the area which is confined by the blue line and the lower black line by the area given by the lower triangle. The maximum achievable Gini-coefficient is given by dividing the white area by the lower triangle. This corresponds to $1 - P_{uc}$, where P_{uc} denotes the unconditional probability of the target event to occur.

Comparing these two values is an indicator of the algorithm's capability of separating signal and background events. The Gini-coefficient can range from 0 to 1, where 0 zero indicates no capability in separating signal and background events whereas 1 indicates perfect separation capability. In the following, this chart type is referred to as Gini-plot.

Efficiency-Purity Plot

In the Efficiency-Purity plot, the observed signal purity $\mathcal{P}$ is plotted in dependence of the signal efficiency $\mathcal{S}$. This can be achieved by moving through the chart in Figure 7.2 from the right to the left corner while continuously monitoring $\mathcal{P}$ and $\mathcal{S}$ in the overcoated interval. In the right corner, only few signals are selected, thus, the signal efficiency is very low. At the same time, due to the algorithm's separating capability, the signal purity is the highest. When all events are taken into account, the signal efficiency is one and the purity is equal to the unconditional probability of the target to occur. This is indicated by the upper curve in Figure 7.4. The more pronounced the curvature of this dependency is, the better is the algorithm's separating capability. The lower curve indicates the above described process as well, but results from moving from the left to the right corner.
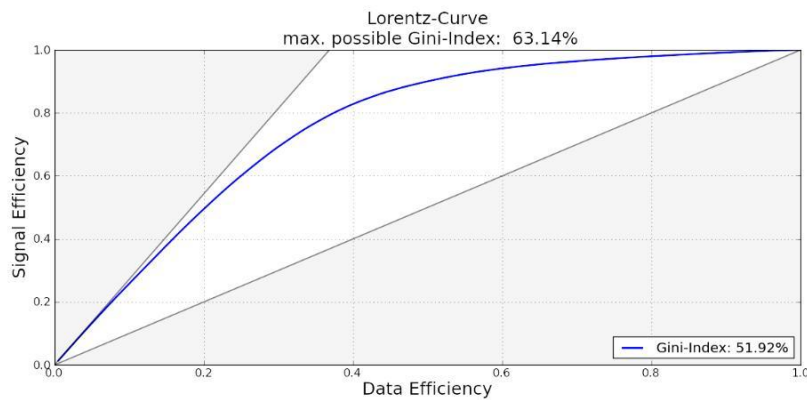
Figure 7.3: The chart illustrates an exemplary Lorenz-Curve. The maximally achievable Gini-coefficient as well as the realized Gini-coefficient are shown above the chart.
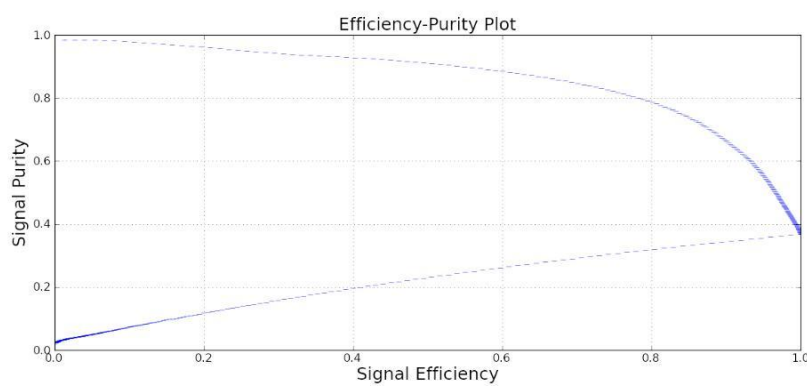

Figure 7.4: An exemplary Effiency-Purity plot. It demonstrates the algorithm's capability of separating signal and background events.

Chapter 8

Forecasting a Limit Order's Execution

As stated in Section 6.1, traders have two fundamental possibilities of participating in the trading process: through limit- and through market(-able) orders. These two order types differ fundamentally in two aspects. First, marketable orders come along with higher costs than passive limit orders. The price difference between a market and a limit order is at least the bid-ask-spread, which can consequently be thought of as an additional cost factor for market orders. Second, market orders are executed immediately at the currently available best price (provided enough liquidity in the order book). As a consequence, there is practically no uncertainty whether the desired amount of shares will change hands.

Limit orders, on the other hand, have lower costs incurred but are afflicted from uncertainty regarding their execution. A trader must choose between the two order types by considering their characteristic advantages and drawbacks.

Bottom line, limit orders are the preferable way of trading shares, in case there is no need for guaranteed execution (Harris and Hasbrouck (1996)).

The uncertainty of execution, which arises when limit orders are used for trading, may be quantified further. The goal of this section is to present an algorithm capable of determining the execution probability of a limit order on a one minute time horizon within a given stock universe.

The algorithm is composed of two parts. First, a model is constructed based on unconditional execution probabilities for each stock's order book level, denoted as *Mean-Matrix-Model*. The second part involves the application of the machine learning algorithm NeuroBayes®, which was presented in Section 7. NeuroBayes® operates in this architecture as a means to find corrections to the Mean-Matrix-Model conditional on a set of variables. Generally, NeuroBayes® could be provided with the order book level and the stock as input variables as well. However, since the execution probability strongly depends on these vari-

ables as will be shown in the next section, smaller effects could potentially be overshadowed as consequence. For this reason, these two variables are treated in an external model.

The final result of the consecutive application of both models is a probability P_e for the execution of a limit order within one minute. This combined technique is presented for two stock exchanges, the French Stock Exchange in Paris and the UK Stock Exchange in London. Both stock exchanges are represented through a selection of their most liquid assets, as given in Appendix B.3. The analysis is carried out on a set of limit orders as described in Section 6.2.2.

The remainder of this section is organized as follows. First, a basic model, the Mean-Matrix-Model, is developed. Subsequently, the application of NeuroBayes® is motivated and presented. Finally, the combined algorithm's forecasting performance is examined in an out-of-sample test. Following on from this, another model based on NeuroBayes® to forecast the full execution of a limit order is presented.

8.1 The Mean-Matrix-Model

When a limit order is inserted into the order book, it will be sorted into the corresponding price level (under consideration of the price-time-priority). As a matter of priority, the execution probability will strongly depend on the according order book level.

In the following, the execution probability of limit orders which are inserted into the lth order book level is estimated from historic data as

$$\hat{P}_{e,l} = \hat{P}_e(t = 1|l) = \frac{n_l(t = 1)}{n_l\,(t \in \{0, 1\})},\tag{8.1}$$

where n_l denotes the number of orders inserted into the the lth level and t is a binary variable indicating whether the order was executed within one minute ($t = 1$) or not ($t = 0$).

The estimated standard error on the quantity $\hat{P}_{e,l}$ can be derived by comparing the order execution process to a Bernoulli experiment with $\hat{P}_{e,l}$ as the probability of success. Thus, the probability of r executions out of n limit orders is given by the binomial distribution:

$$f(r; n, \hat{P}_{e,l}) = P(R = r) = \binom{n}{r}\hat{P}_{e,l}^r(1 - \hat{P}_{e,l})^{n-r},\tag{8.2}$$

where R is the random variable denoting the realized number of executions. The mean of the binomial distribution is given by $n\hat{P}_{e,l}$. The corresponding distribution's variance is $\sigma_r^2 = n\hat{P}_{e,l}(1 - \hat{P}_{e,l})$. This leads to an estimated standard error

Exchange	$\hat{\alpha}$	$\hat{\kappa}$	$\hat{c}$	NDF	χ^2	χ^2/NDF
LSE	2.42	3.35	0.005	2	2.55	1.27
EP	5.86	3.40	0.03	2	4.72	2.36

Table 8.1: The estimated parameters from model 8.4 are shown in this table together with goodness of fit parameters. The model has been applied to data from the London Stock Exchange as well as to data from Euronext Paris.

of the execution probability of:

$$\sigma_{\hat{P}_{e,l}} = \left(\frac{\hat{P}_{e,l}(1 - \hat{P}_{e,l})}{n} \right)^{\frac{1}{2}}. \tag{8.3}$$

In order to empirically investigate the execution probability's dependency on the order book level further, the probability estimation from Equation 8.1 is used. This allows to derive a model which describes the behavior of the execution probability in dependence of the order book level. Figures 8.1a and 8.1b show this dependency for the London Stock Exchange (LSE) and Euronext Paris (EP), respectively.

A power law model is suggested to describe the observed dependency for the order book levels 1-5. The 0th order book level, which denotes those limit orders placed aggressively with a better price than available at the time of insertion, does not fit into this model. This is indicated by a dashed line which is the extrapolation of the model between levels 1-5.

The utilized power law model is given by

$$P_e(l) = \kappa \cdot l^{-\alpha} + c, \tag{8.4}$$

where l denotes the order book level and P_e the unconditional execution probability resulting from this model. The model's estimated parameters are given in Table 8.1.

The model parameter α is a measure for the dependence of the execution probability on the order book level. The higher this value is, the more pronounced is this dependency. In accordance with the estimated parameter $\hat{\alpha}$, the change of the execution probability over the first five order book levels is more intensified at the EP as compared to limit orders at the LSE.

It has to be considered, though, that the meaning of an order book level may fundamentally differ among the stocks. One of the reasons is the stock's relative tick size, given by $TS_r(t) = TS(t)/S(t)$, where $S(t)$ denotes the stock price at time t. The higher this value is, the bigger are the relative price differences between adjacent levels.

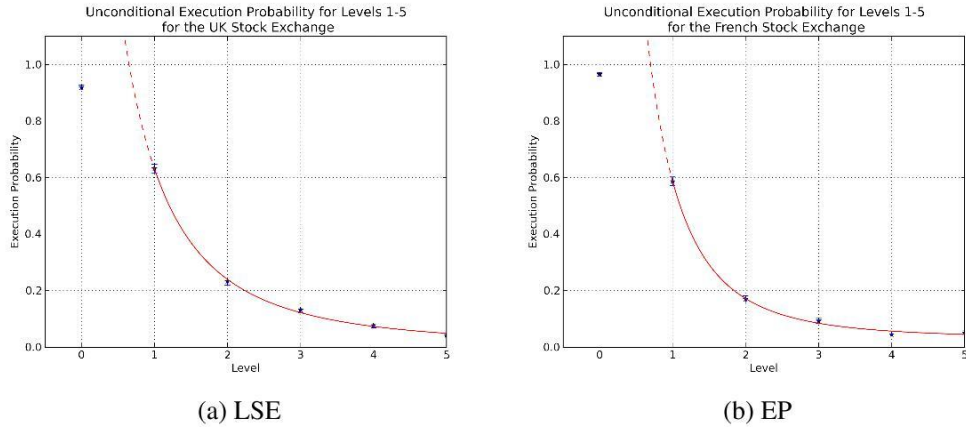(a) LSE   (b) EP

Figure 8.1: Execution Probabilities for the LSE and the EP in dependence of the order book level. The 0th level does not fit into this power law model as indicated by the dashed red line. For higher levels, the execution probability remains only very slightly above zero.

Additionally, each stock has a characteristic average order book level population with limit orders. As a consequence, order book levels are not straightforwardly comparable between different stocks.

Hence, it is to assume that the execution probabilities for one distinct order book level compared between different stocks will show some degree of dispersion. This is indicated in Figure 8.2. The observed unconditional execution probabilities within one level are shown in differently colored histograms.

Given the execution probability's strong dependency on the order book level and the dispersion of this probability among different stocks, a basic approach is chosen which treats each stock-level pair individually, denoted as Mean-Matrix-Model. The execution probabilities in the Mean-Matrix-Model are derived in the same way as given in Equation 8.1, but for each stock-level pair individually. Thus, this model is given by

$$P_e^{mm}(\text{s},\text{l}) = \hat{P}_{e,l}^{(s)}, \tag{8.5}$$

where $\hat{P}_{e,l}^{(s)}$ denotes the estimated execution probability of a stock-level pair, s is used as an index to refer to a stock and l is used to index the order book level.

A significant difference in the unconditional execution probabilities between bid- and ask limit orders is not detected in the utilized data sample. As a consequence, the Mean-Matrix-Model is used as a basic model without distinguishing the order type.
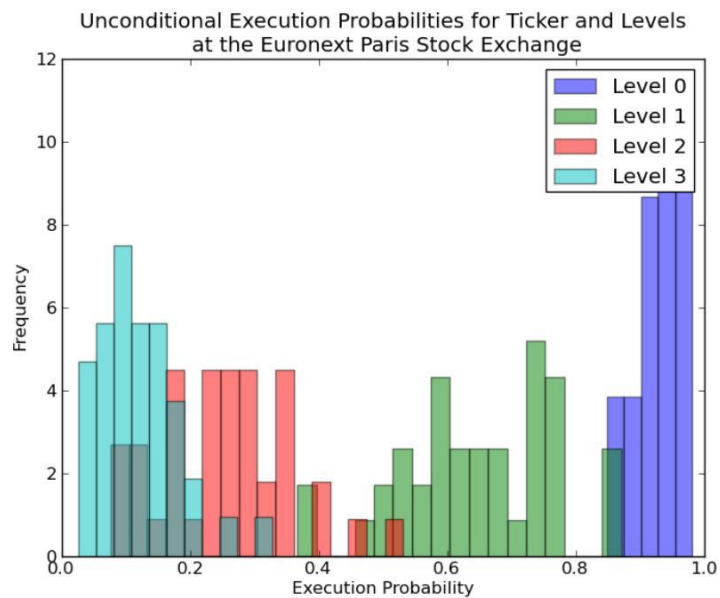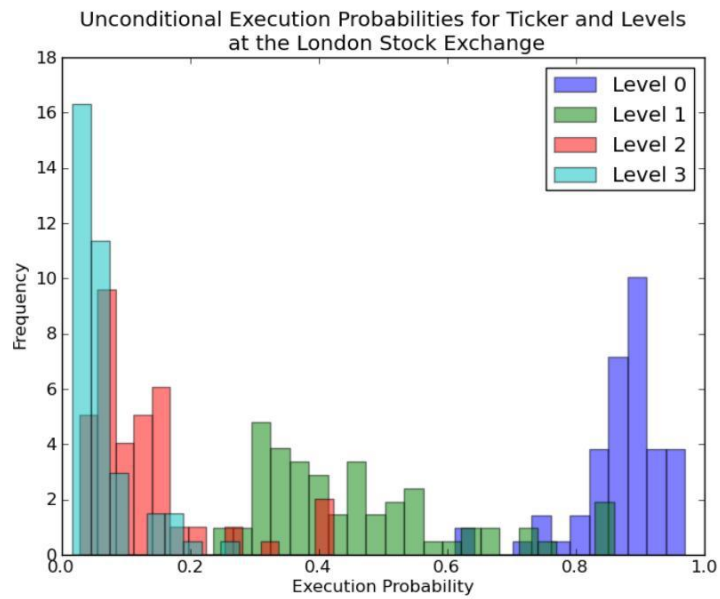
(a) LSE

(b) EP

Figure 8.2: The dispersion of the execution probabilities in the stock universe within one order book level is shown in the above figures. Each order book level is shown in a different color. The widths of the respective distributions indicate that the execution probabilities differ within one level among the stocks.

8.2   Application of NeuroBayes®

The Mean-Matrix-Model is taken as a starting point for the development of a more sophisticated model based on the machine learning algorithm NeuroBayes®. The Mean-Matrix-Model contains the mean stock- and level related execution probabilities. It is obvious, though, that the execution probability of a limit order will in general depend on even more quantities than just those two. For example, order book asymmetries, intraday volatility, trading volume, order book thickness, etc. may play important roles as descriptive quantities as well. Based on these assumptions, NeuroBayes® is provided with a collection of potentially descriptive variables, as outlined in Section 6.2.

In order to account for asymmetries in the order book, under which bid- and ask-type limit orders will potentially exhibit a different behavior, NeuroBayes® is applied on either of the order-types separately.

More precisely, NeuroBayes® is employed as a means to find corrections on the execution probabilities from the Mean-Matrix-Model conditioned on a selected set of such input variables. In a wider sense, probabilities are constructed in a sample space, in which the new information is orthogonal to the Mean-Matrix-Model. The resulting quantity of the NeuroBayes® algorithm is a probability P_i^c which is to be interpreted as a correction term. The meaning of a value of $P_i^c = 0.5$ is that no correction has to be carried out.[1] That is, the execution of a limit order is well estimated by the Mean-Matrix-Model. Thus, for the Mean-Matrix-Model to be correct in average, the resulting probabilities should be centered around 0.5.

8.2.1   NeuroBayes® Training Results

After applying NeuroBayes® as a means to find corrections on the Mean-Matrix-Model, the resulting probability corrections are investigated in this section. The resulting additional classification capability of signal and background events on the back of the Mean-Matrix-Model is shown in Figure 8.3 for both order types separately as well as for the LSE and the EP.

These charts are an indicator for the learning capability of the NeuroBayes® algorithm. The red histogram illustrates the probability correction for those limit orders which were executed by an arriving marketable order. Those limit orders which were not executed within one minute and thus remained in the order book are depicted by the black histogram. The red distribution is rather centered around higher probability values whereas the black curve appears attracted towards low probability values. This is an indicator for the separation capability of the NeuroBayes® algorithm.

---

[1] A prove of this statement is given in Section 8.3.

(a) LSE Bid-Limit-Orders

(b) LSE Ask-Limit-Orders

(c) EP Bid-Limit-Orders

(d) EP Ask-Limit-Orders
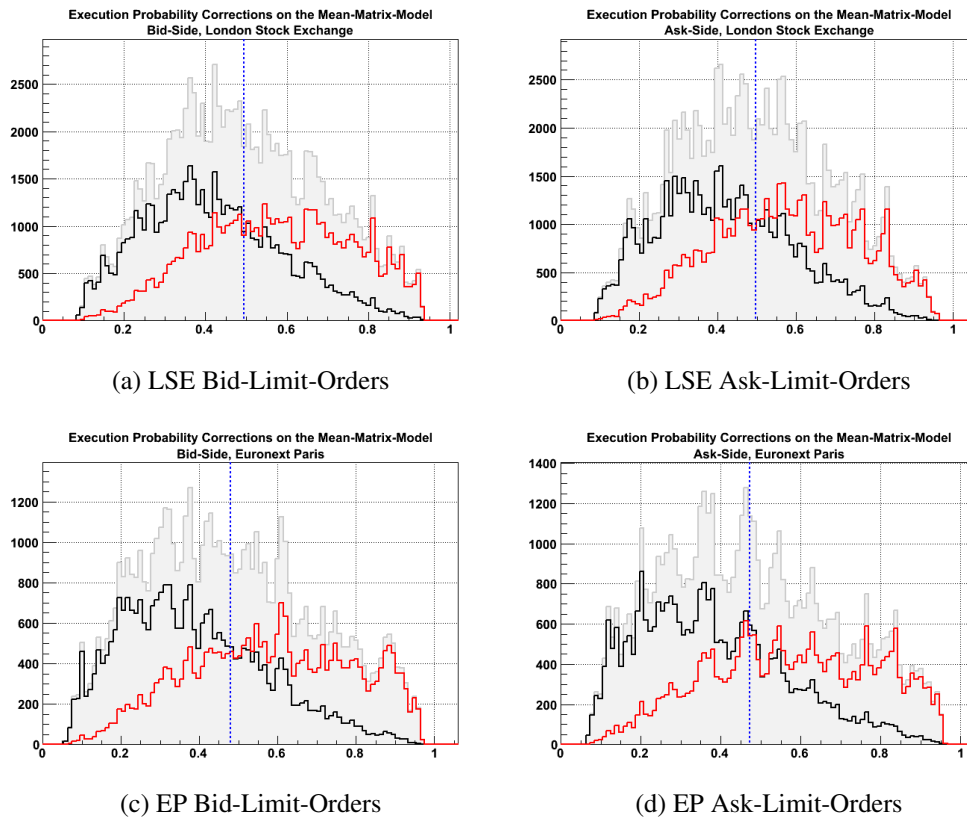
Figure 8.3: The above charts show the distribution of the limit order execution probability corrections on the Mean-Matrix-Model, resulting from NeuroBayes®. The gray histogram depicts the distribution of all resulting probabilities, which is roughly centered around 0.5 in all charts. The actual mean value of the gray distribution is indicated by the dashed blue line. The red histogram depicts only those limit orders for which the target variable was true whereas the black curve illustrates the distribution of those probability corrections for which the target variable was false. The red curve is more drawn towards high probabilities whereas the black curve is rather centered around low probability values. This indicates the separation capability of the NeuroBayes® algorithm conditional on the given set of input variables.

The gray histogram in these charts depicts the joint distribution of executed and un-executed limit orders. It has to be centered around 0.5, indicating that the Mean-Matrix-Model is correct on average. The actually observed mean value of the joint distribution is indicated by a dashed blue line. Hence, this condition is almost exactly fulfilled for limit orders at the LSE. For limit orders from the EP, there appears to be a slight bias in the Mean-Matrix-Model towards lower

execution probabilities. This effect is removed by applying the NeuroBayes®
algorithm.

8.2.2 Discussion of Descriptive Variables

After having introduced the Mean-Matrix-Model in Section 8.1, NeuroBayes®
was applied in order to find corrections to this model conditional on a large set of
input variables. A selection of those variables that exhibited strong dependencies
on the execution of a limit order is presented in this section. These dependencies
are to be understood as corrections on the Mean-Matrix-Model. A probability cor-
rection with a value of 0.5 indicates that no correction is to be carried out. Thus,
the observed dependencies on the execution probability are centered around 0.5,
indicating that the Mean-Matrix-Model is correct in average. Without the Mean-
Matrix-Model, the dependencies would potentially show a changed structure, but
not disproving the conclusions which are drawn from the observed patterns.

In order to have the examined quantities comparable between different stocks,
they are further normalized. The normalization which is carried out for each vari-
able individually is explained in the respective section.

Recall that NeuroBayes® was applied to bid- and ask-type orders individu-
ally in order to detect impacts of asymmetries in the order book to the execution
probability. Hence, the dependencies listed below are shown for bid- and ask-type
limit orders individually.

Dependency on Price Asymmetries in the Order Book

The variable *Volume Weighted Mid-price* ($VWMP_s$), as introduced in Sec-
tion 6.2, exhibits a significant correlation to the execution probability at both the
London Stock Exchange as well as the Euronext Paris Stock Exchange. This quan-
tity is further normalized in that the difference to the mid-price is constructed and
the resulting value is divided by a volatility measure defined on a one minute time
horizon. By this, the $VWMP_s$ becomes comparable between different stocks.

The dependency of the quantity $VWMP_s$ to the execution probability is de-
picted in Figure 8.4 for both the LSE and the EP as well as for both order types,
bid-type limit orders and ask-type limit orders.

A first observation is that the two order types exhibit a strong antisymmetric
behavior. This indicates that the two order types do not exhibit significant struc-
tural differences in their dependencies to the execution probability. In a model
where both order types are equally present, this and similar effects would not be
identifiable without particularly considering the antisymmetric behavior in a pre-
ceding step.

(a) LSE Bid-Limit-Orders

(b) LSE Ask-Limit-Orders
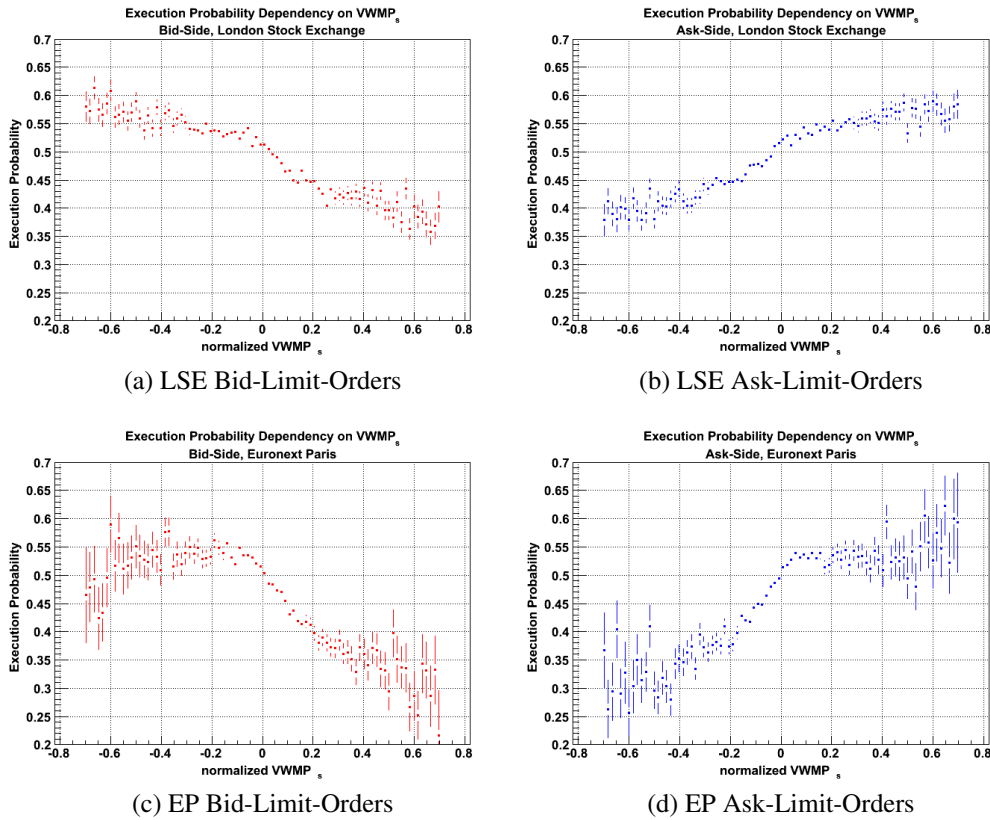
(c) EP Bid-Limit-Orders

(d) EP Ask-Limit-Orders

Figure 8.4: The above figures illustrate the dependency of the execution probability on the quantity $VWMP_s$ for limit-orders of either bid- or ask-type. It is clearly evident that the two order types show antisymmetric behavior. These dependencies are depicted for both limit orders at the LSE as well as at the EP.

Differences are recognizable between the exchanges as well. The descriptive quality of $VWMP_s$ appears to be stronger at the Euronext Paris stock exchange, indicated by a more pronounced slope of the dependency and a wider dispersion in the according execution probabilities. This is especially true for the lower boundary which is at roughly 30% at the EP, whereas, at the LSE, the lower boundary is at roughly 40%. The upper boundary is in both cases at approximately 55%.

Negative values of the normalized $VWMP_s$ indicate that the observed value of the original quantity was below the mid-price. This, in turn, implies that more liquidity was provided close to the mid-price on the bid-side than on the ask-side. As a result, the execution probability of ask-type limit orders is increased when more liquidity is available on the ask-side whereas the execution probability is lower when more liquidity is on the bid-side. The contrary interpretation holds for bid-type orders. This antisymmetry is detected on both the London and the

Paris stock exchange. This phenomenon may be explained if the $VWMP_s$ is assumed to be a more realistic market-price measure than the mid-price. If positive differences between the $VWMP_s$ and the mid-price imply that the stock's price is going to be rising, ask-type limit orders would have a higher execution probability as a result. That this effects actually exists is shown in Figure 8.5 for both, the LSE and the EP.

(a) LSE	(b) EP
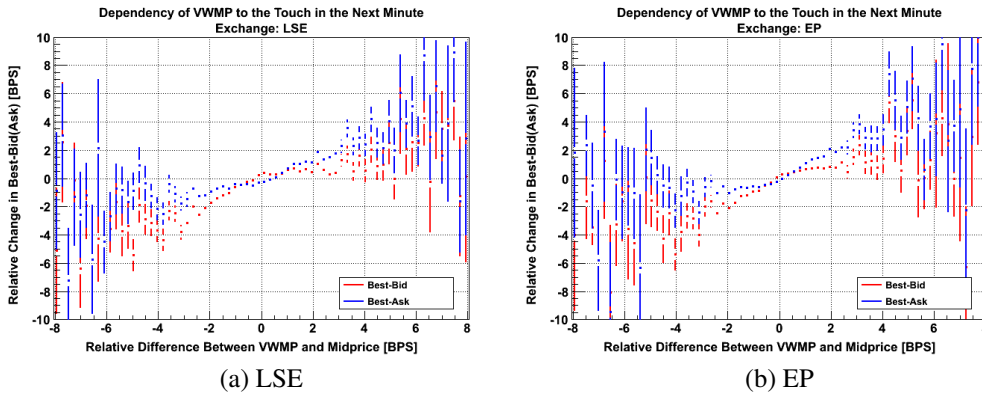
Figure 8.5: The above charts show that the difference between the quantity VWMP and the mid-price contains information about the future development of the best-ask as well as the best-bid price. This is shown for both the LSE and the EP.

Dependency on Liquidity Differences in the Order Book

A further descriptive variable is constructed which takes asymmetries in the number of pending shares in the first ten order book levels into account. This input variable is constructed as

$$
\begin{aligned}
N_A &= \text{\# of shares pending in the first 10 order book ask levels} \\
N_B &= \text{\# of shares pending in the first 10 order book bid levels} \\
\Delta_{AB} &= \frac{N_A - N_B}{N_A + N_B}
\end{aligned}
\tag{8.6}
$$

This variable is a relative measure of how many shares are pending in the order book levels on the ask-side as compared to the bid-side. It is therefore a relative measure of the provided liquidity compared between the ask- and the bid-side. Negative values of this quantity imply more pending shares on the bid-side whereas positive values imply more shares on the ask-side. The dependency of

(a) LSE Bid-Limit-Orders

(b) LSE Ask-Limit-Orders

(c) EP Bid-Limit-Orders

(d) EP Ask-Limit-Orders
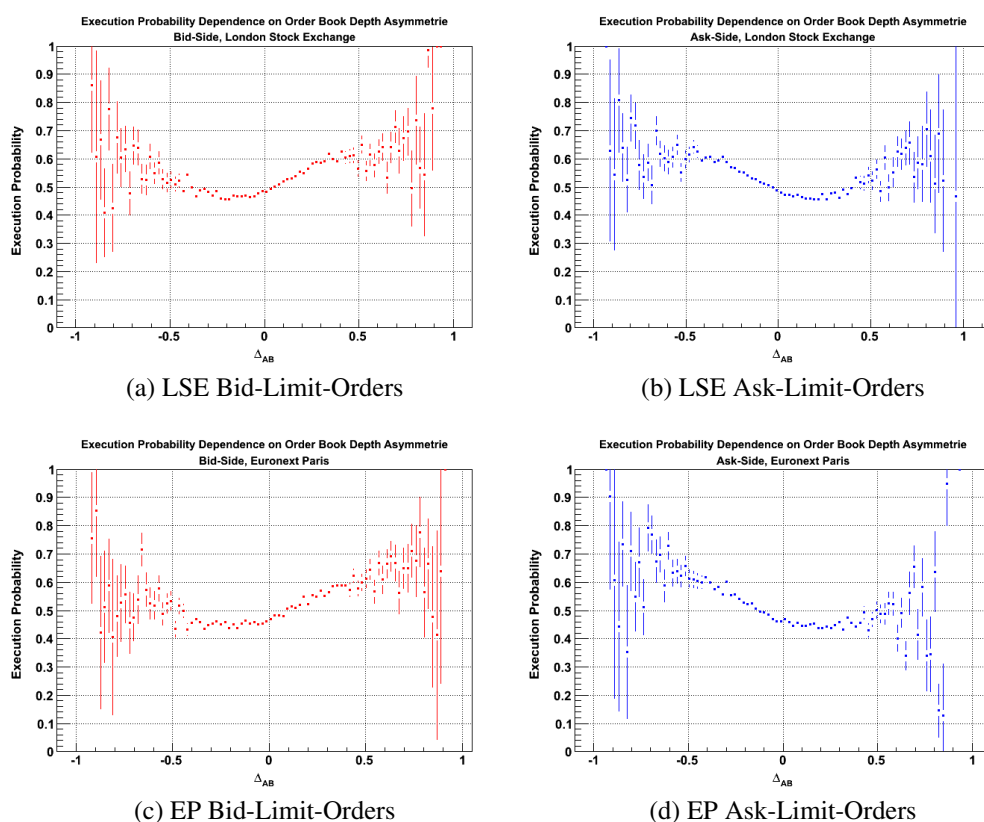
Figure 8.6: The dependency of the execution probability on the descriptive variable Δ_{AB} is depicted in the above figures for bid- and ask-type orders separately. The two order types exhibit a significant antisymmetric behavior. This is shown for both the LSE and the EP.

the execution probability on this variable is depicted in Figure 8.6 for both the LSE and the EP as well as for bid- and ask-type limit orders separately.

This variable is distinguished from the quantity $VWMP_s$ in that it does not take the level-price distribution into account. It is a pure measure of the relative differences in the provided liquidity. The quantity $VWMP_s$, however, is constructed by first determining the weighted *center-of-mass* prices of each order book side individually. Following this, the mean of these prices is computed. By that, asymmetries in the overall provided liquidity between the bid- and ask side are not contained in this measure. The $VWMP_s$ reflects the difference in the price and volume distributions of the limit orders between the two order book sides.

The interpretation of this variable is somewhat opposed to the interpretation of the quantity $VWMP_s$. With increasing positive values, that is, more pending shares on the ask-side than on the bid-side, the execution probability of ask-type

limit orders is decreasing. The opposite is true for bid-type limit orders. This relationship is prevalent on both the LSE and the EP.

This behavior is in line with the intuitive assumption that, the more shares are already pending, the less likely it is for a newly arriving limit order to be executed within a given time horizon. Note that the above charts consider limit orders on all order book levels. Therefore, the effect might differ in strength between different order book levels.

The interpretation of this variable is in line with the assumption that the short term stock price development depends on the quantity Δ_{AB} as well. This is empirically shown in Figure 8.7. The ask- and bid-price log-returns of the following minute are shown in dependence of Δ_{AB}. Here, the stock price decreases in average with more liquidity on the ask-side. This statement is valid for both analyzed stock exchanges.

(a) LSE   (b) EP
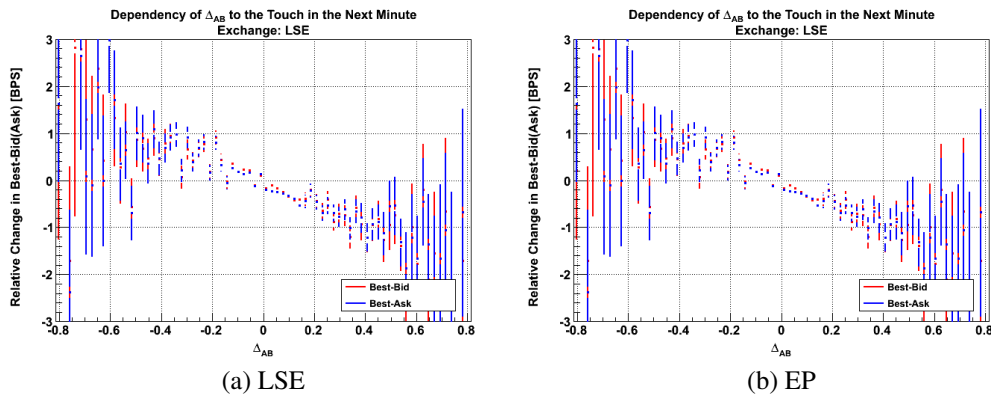
Figure 8.7: The dependency of the one minute ahead touch (the best-bid and best-ask price) development in dependence on the quantity Δ_{AB} for both the LSE and the EP.

Note that, if $\Delta_{AB} = 0$, the execution probability is roughly centered at 0.5, that is, no correction to the Mean-Matrix-Model is carried out in this case. The same statement holds for the quantity $VWMP_s$. This implies that the Mean-Matrix-Model is correct in average.

Dependency on the Trading Activity

The trading activity in this analysis is measured by the number of shares traded within one minute. This quantity is normalized by dividing the traded quantity of shares by the number of shares which are in the order book at the time when a limit order under consideration is inserted. By this, the quantity becomes a measure of how much of the provided liquidity was traded in the past minute. Since this

(a) LSE Bid-Limit-Orders

(b) LSE Ask-Limit-Orders

(c) EP Bid-Limit-Orders

(d) EP Ask-Limit-Orders
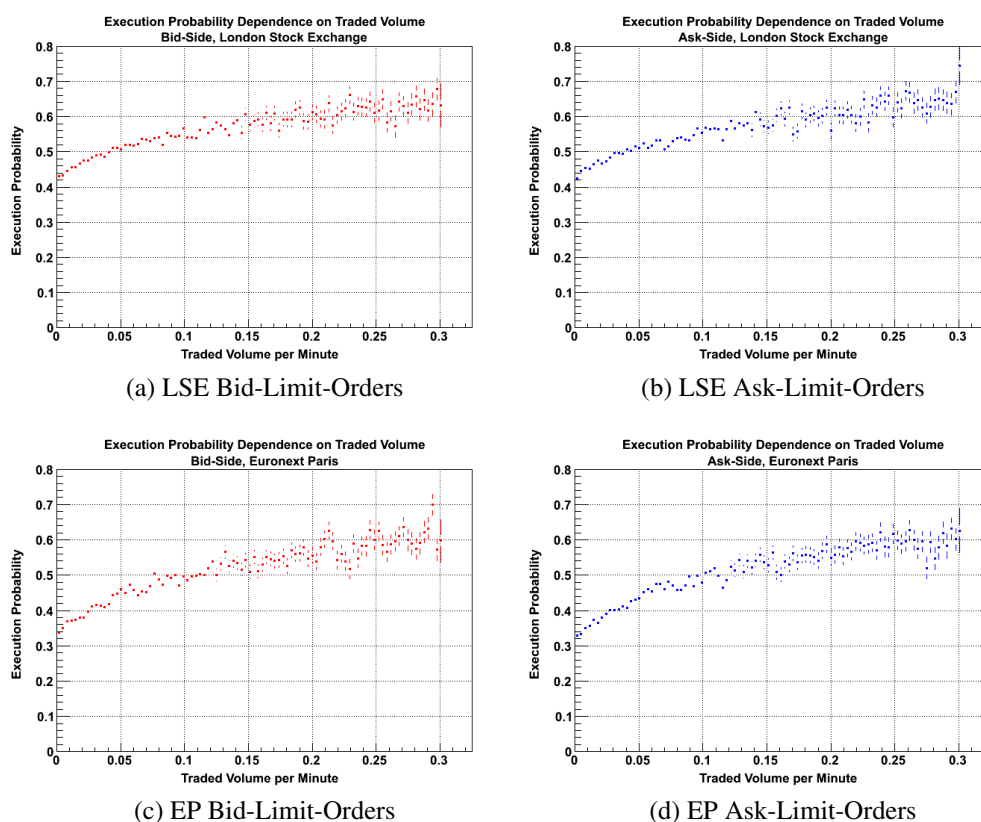
Figure 8.8: The dependency of the trading activity is depicted in the above figures for bid- and ask-type orders separately. This quantity contains bid- as well as ask-side relevant information. Thus, both order types exhibit the same dependency. This is shown for both the LSE and the EP.

variable is a quantity which is not sensitive to a distinct order book side, both order types should exhibit the same dependency structure. Figure 8.8 illustrates the relationship between the normalized trading activity and the execution probability of both order types and for both exchanges.

It appears that no significant difference in the relation between the execution probability and the order type is observable. The execution probability increases with increasing trading activity for both order types equally. Additionally, both exchanges show a very similar structure in the relation between the execution probability and the trading activity. This implies that, in phases of strong activity, the execution of a newly inserted limit order is more likely than in phases of weak activity as one would intuitively assume.

8.3   Investigation of the Combined Model

After the application of NeuroBayes® as a correction algorithm to the Mean-Matrix-Model, the resulting probabilities, P_e^{mm} and P_i^c have to be combined to receive the final execution probability P_i^e, conditional on the variables employed in the NeuroBayes® algorithm. In the space of log-likelihood ratios, this corresponds to an addition of the corresponding log-likelihood ratios. A likelihood ratio is defined by $LR(p) = \ln(p/(1-p))$. Thus, the final probability is constructed by

$$LR(P_i^e) = \ln(P_i^e/(1 - P_i^e)) = LR(P_e^{mm}) + LR(P_i^c). \tag{8.7}$$

Transforming this equation back into the probability space and solving the result for P_i^e leads to:

$$P_i^e = \frac{P_e^{mm} P_i^c}{P_e^{mm} P_i^c + (1 - P_e^{mm})(1 - P_i^c)}. \tag{8.8}$$

A derivation of the above formula is given in Appendix B.2.

By inserting 0.5 for P_i^c in Equation 8.8, P_i^e equals P_e^{mm} as a result. This is indicating that a NeuroBayes® result of $P_i^c = 0.5$ does not change the Mean-Matrix-Model, as qualitatively explained above.

8.3.1   Forecasting Performance of the Combined Model

After the application of NeuroBayes® on a training-sample and combining the expertise with the Mean-Matrix-Model, the resulting algorithm is applied to forecast the execution of limit orders in an out-of-sample test. Here, the relevant question is if the forecasting capability of the combined algorithm is comparable to the forecasting capability in an in-sample test. By examining this question, it can be proven that the algorithm is actually capable of forecasting the target variable and is not just learning in-sample idiosyncracies.

This issue can be investigated by considering the charts which were introduced in Section 7. Consequently, the resulting in-sample and out-of-sample forecasts are compared by considering the Signal-Background histogram, the Signal-Purity plot, the Efficiency-Purity plot as well as the Lorenz-Curve and the corresponding Gini-coefficient. This collection of charts is presented for limit orders at the LSE in Figure 8.9 for the in-sample test and in Figure 8.10 on out-of-sample data. The respective charts for limit orders at the EP are shown in Figures 8.11 and 8.12. Note that these charts consider both bid- and ask-type limit orders. They are constructed after merging the NeuroBayes® results from both order types.

One first observation is that all Signal-Background histograms show that signal and background events are considerably separated. The resulting combined model probability P^e for a limit order to be executed within one minute is clearly

well coinciding with a diagonal through the origin with slope one in all cases as shown by the respective Signal-Purity plots. Considering these charts is especially important in the out-of-sample test in order to prove the interpretability of the combined model's output as a probability.

The blue line in the Gini-plots indicates the Lorenz-Curve of the combined model. The maximally achievable Gini-coefficients are approximately equal among the in-sample and out-of-sample data sources, indicating that the same unconditional distributions of the target variable were present in the respective samples. The observed Gini-coefficients for the combined models, provided in Table 8.2, appear to be approximately equal as well.

	In-Sample		Out-Of-Sample	
Exch.	Gini-Index	Max. Possible	Gini-Index	Max. Possible
LSE	59.62	72.19	59.35	71.86
EP	54.04	63.27	53.38	62.51

Table 8.2: The Gini-Coefficients for the applied combined model are shown in the above table for in-sample and out-of-sample tests for both the London Stock Exchange and the Euronext Paris Stock Exchange.

Additionally, the added forecasting capability of the NeuroBayes® algorithm on the execution probability is measured in order to investigate if there could be achieved a significant improvement to the Mean-Matrix-Model. This examination can be carried out by comparing the respective Lorenz-Curves and Gini-coefficients. The probability range in which applying the NeuroBayes® algorithm leads to significant improvements can be detected by comparing the spread between the two curves. The Lorenz-Curve of the Mean-Matrix-Model is visualized by a green line in the respective Gini-plots in all Figures 8.9 to 8.12. The curve depicting the combined model is in all cases above the green Mean-Matrix-Model curve. This implies that applying the NeuroBayes® algorithm improves the forecasting capability on the Mean-Matrix-Model. The execution probabilities received from the Mean-Matrix-Model have a high descriptive quality already. This is due to the fact that it contains the two strongest effects, the order book level dependency as well as the stock characteristics. A comparison of the respective out-of-sample Gini-coefficients is provided in Table 8.3.

8.4 Forecasting a Limit Order's Execution Fraction

After having addressed the question whether a limit order is at least going to be partly executed, this issue is further investigated by considering the limit or-

(a) Signal-Background Distributions

(b) Signal-Purity Diagonal

(c) Efficiency-Purity Plot

(d) Gini-Plot
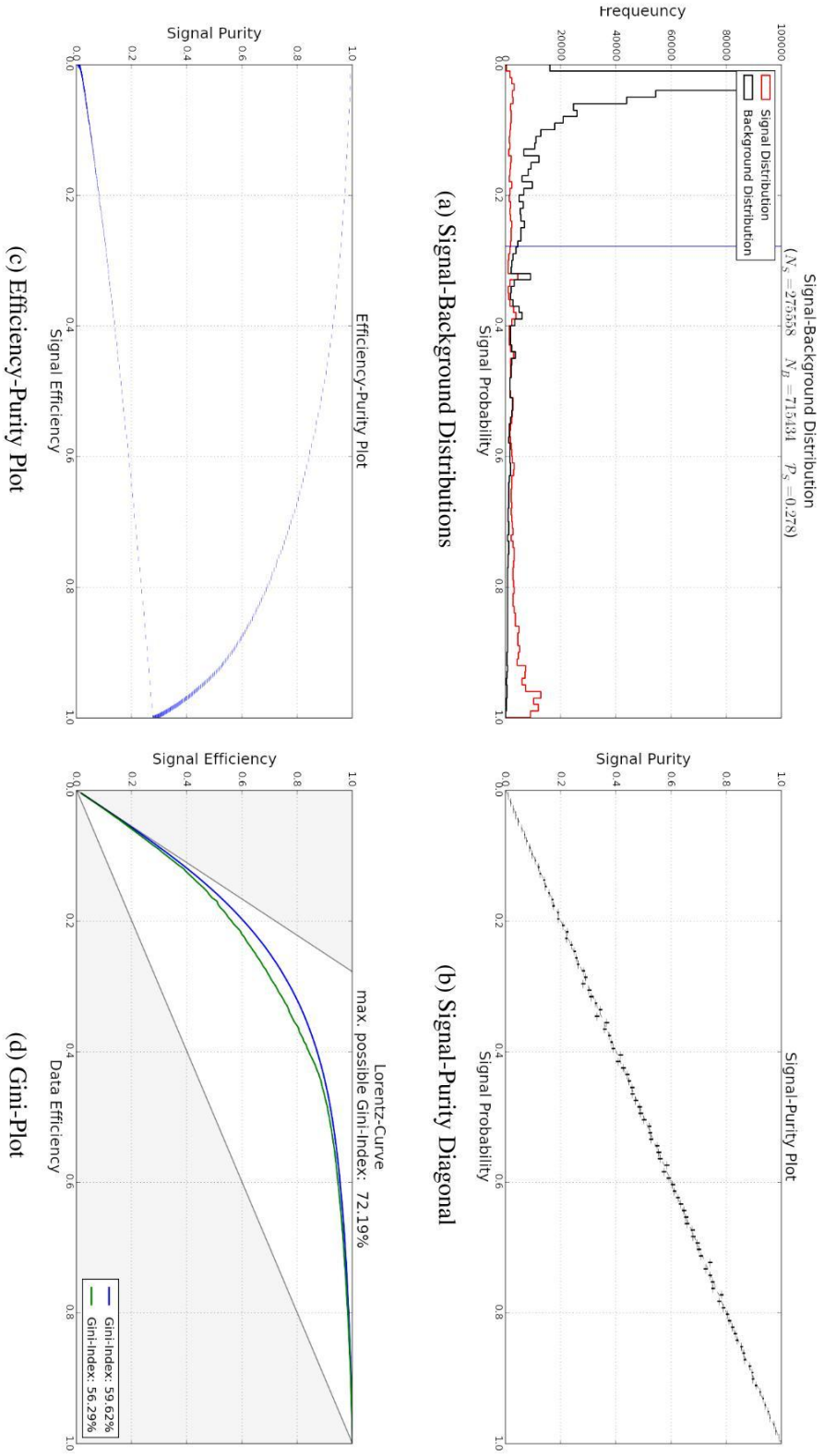
Figure 8.9: LSE — in-sample investigation of the forecasting capability of the combined model.

(a) Signal-Background Distributions

(b) Signal-Purity Diagonal

(c) Efficiency-Purity Plot

(d) Gini-Plot
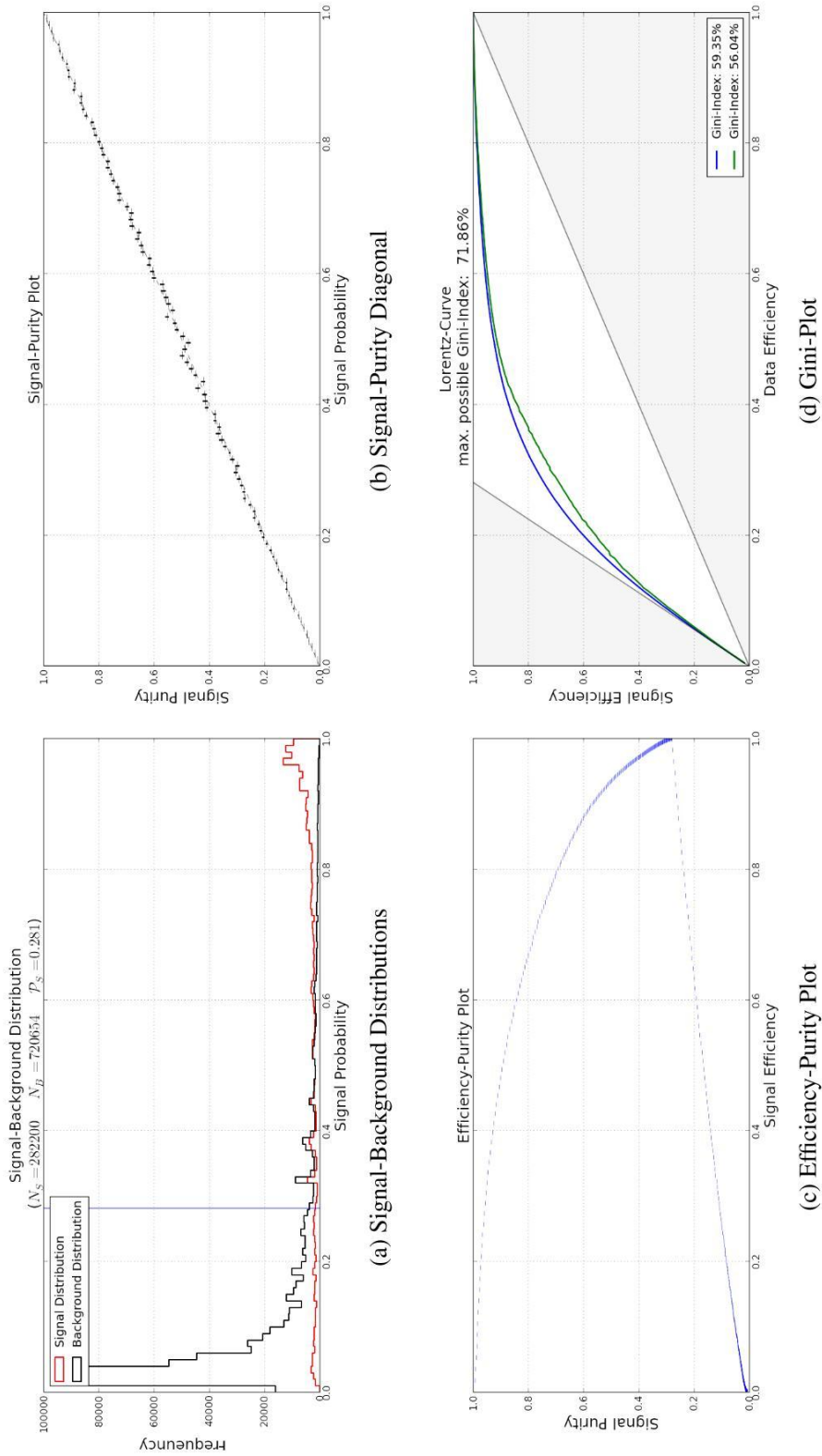
Figure 8.10: LSE — out-of-sample investigation of the forecasting capability of the combined model.

(a) Signal-Background Distributions

(b) Signal-Purity Diagonal

(c) Efficiency-Purity Plot

(d) Gini-Plot
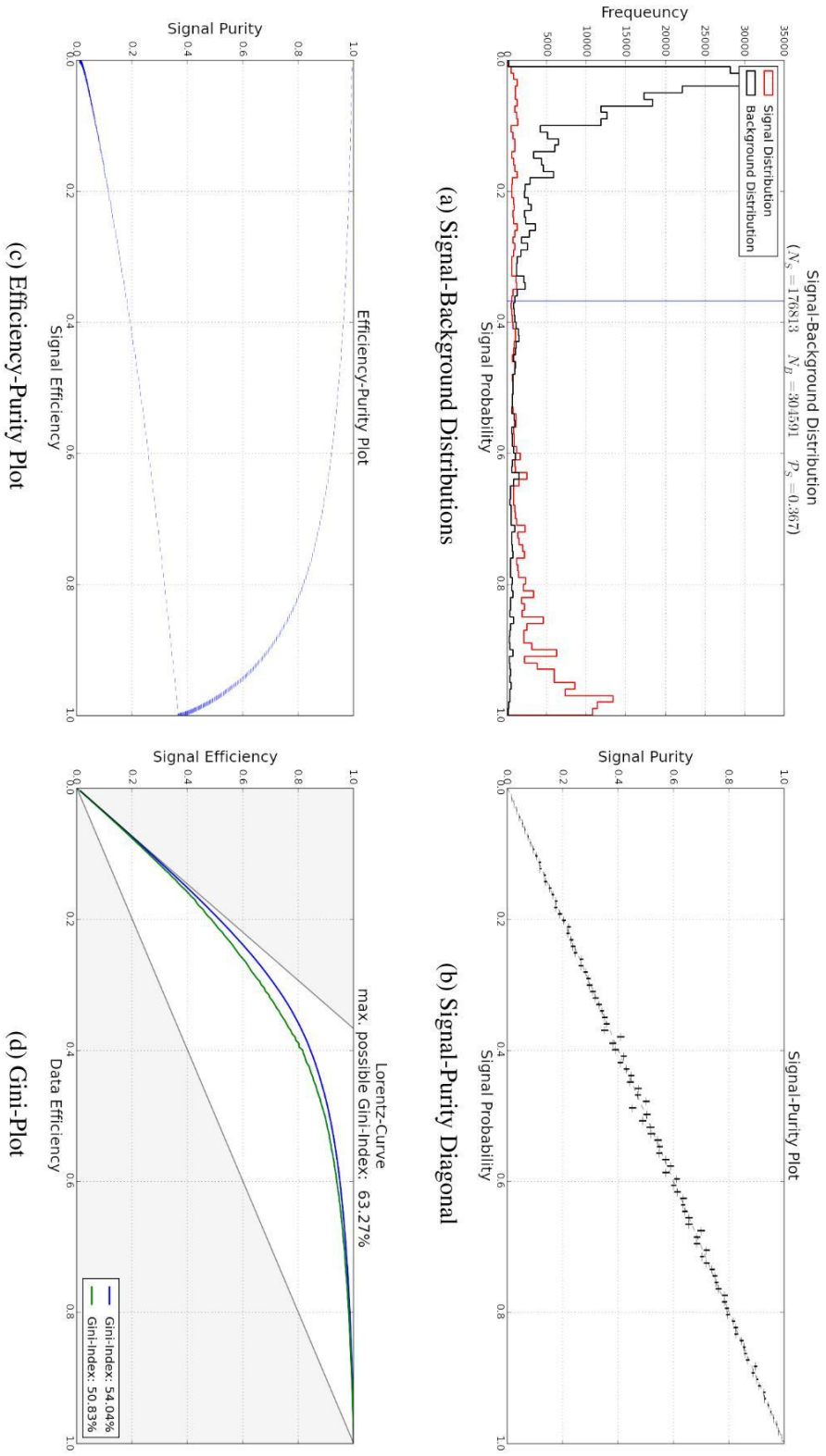
Figure 8.11: Euronext Paris — in-sample investigation of the forecasting capability of the combined model.

(a) Signal-Background Distributions

(b) Signal-Purity Diagonal

(c) Efficiency-Purity Plot

(d) Gini-Plot
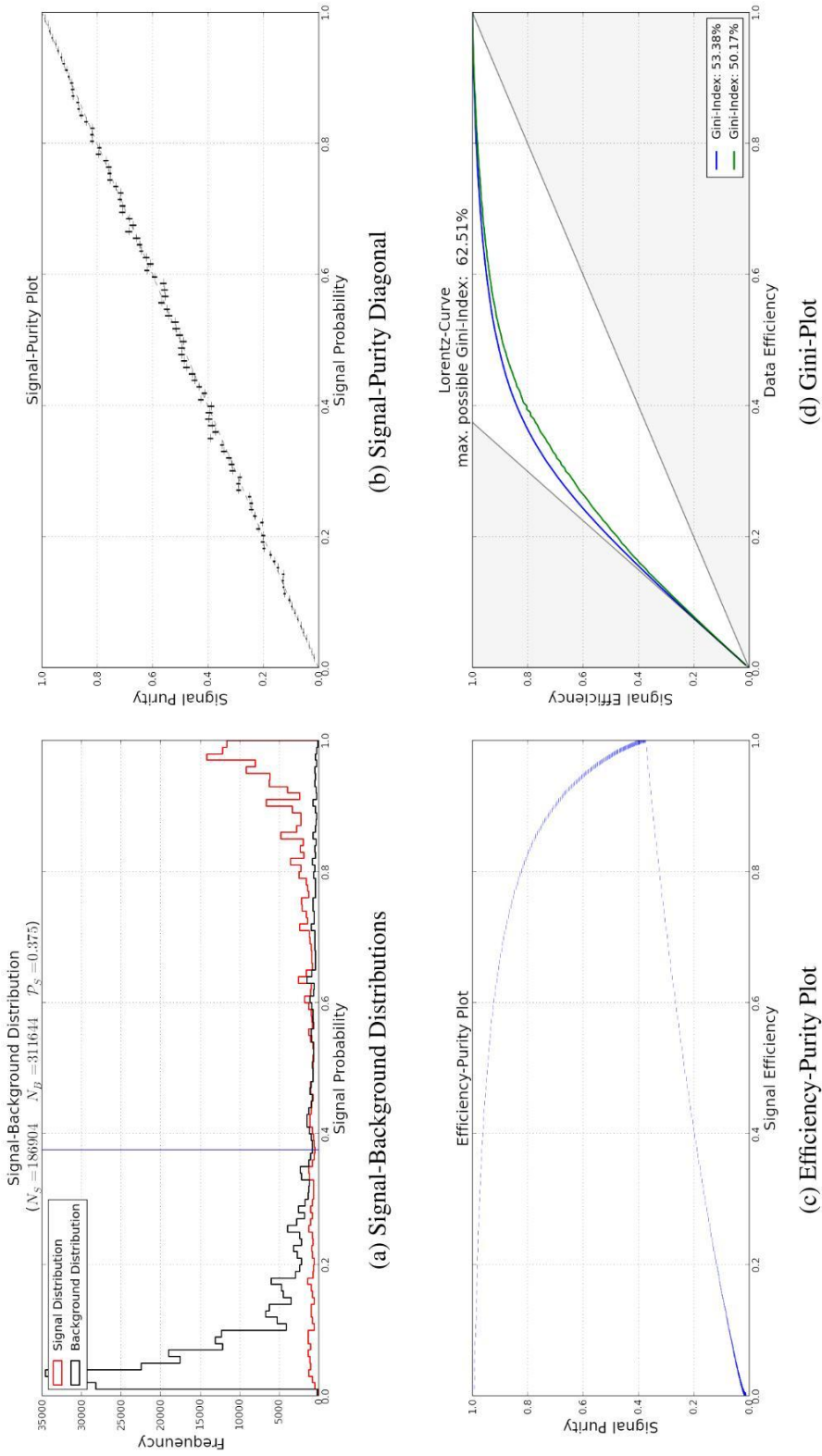
Figure 8.12: Euronext Paris — out-of-sample investigation of the forecasting capability of the combined model.

| | Gini-Coefficient | | |
Exch.	Combined Model	Mean-Matrix-Model	Max. Possible
LSE	59.35	56.04	71.86
EP	53.38	50.17	62.51

Table 8.3: The achieved out-of-sample Gini-coefficients are compared between the combined model and the Mean-Matrix-Model. The respective values are shown in the above table.

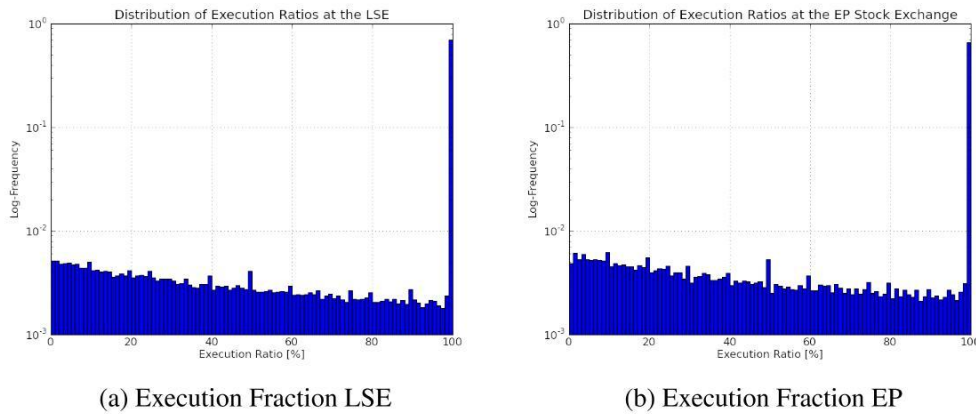(a) Execution Fraction LSE   (b) Execution Fraction EP

Figure 8.13: The above figures illustrate the unconditional execution fractions for limit orders at the London Stock Exchange (LSE) and the Euronext Paris Stock Exchange (EP). An execution fraction of 100% occurs most frequently.

der's execution fraction. As stated in Section 6.1, limit orders may be only partly executed by arriving marketable orders. Figures 8.13a and 8.13b depict the distribution of these fractions for the London Stock Exchange (LSE) and the Euronext Paris Stock Exchange (EP). By examining these charts it becomes clearly evident that a full execution (Execution Ratio of 100%) occurs in most cases. This is indicated by the strongly increased frequency of execution fractions at 100%. The average execution ratios are $\approx 82\%$ for limit orders at the LSE and $\approx 80\%$ at the EP. The unconditional probabilities for full execution are $\approx 70\%$ at the LSE and $\approx 65\%$ at the EP. In order to forecast whether a limit order is going to be fully executed conditional on a set of market-descriptive variables, NeuroBayes® can be applied as in the previous section on limit order execution probabilities.

The unconditional probabilities for full execution do not show a strong dependency on the order book level, as indicated by Table 8.5. Thus, the development of a basic model, as carried out in the previous section, is not necessary. As a conse-
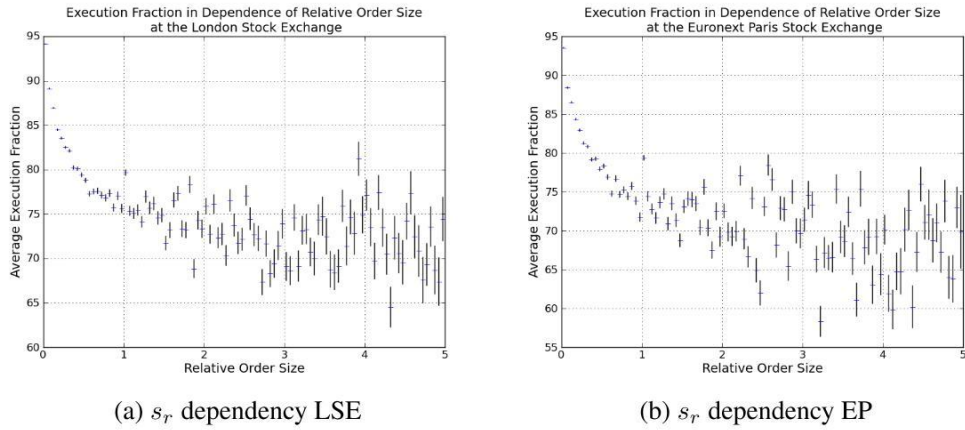
(a) s_r dependency LSE
(b) s_r dependency EP

Figure 8.14: The above charts show the dependency of the relative order size r_s to the probability of full execution.

quence, NeuroBayes® can be applied without taking unconditional probabilities as a basis.

By applying NeuroBayes®, the question of full execution of a limit order can be forecasted conditional on a set of descriptive variables (here, the same set of input variables is used as in the analysis on forecasting a limit order's execution).

The by far most important explanatory variable was found to be the relative order size, constructed as

$$s_r = \frac{\# \text{ of shares in the limit order}}{\# \text{ of shares pending in the level of insertion}}. \tag{8.9}$$

For orders which are inserted into the 0th order book level, the number of pending shares in the first order book level is taken as a reference.

Generally, distinct stocks have different characteristic average limit order sizes and level sizes. Thus, using the plain size of a limit order may potentially result in a training on stock characteristics, and not, as desired, a training on order sizes. By weighting the order size with the number of shares which are pending in the level in which the limit order is inserted, this input variable becomes comparable among different stocks. The dependency of full execution to the relative order size is shown in Figure 8.14 for both the LSE and EP.

The declining probability of the full execution with increasing relative order size is in line with the intuitive assumption that the bigger a limit order is in relative terms, the more likely it is to be executed by several marketable orders. If the size of an inserted limit order coincides with the already existing quantity of pending shares in the level of insertion, the unconditional probability of full

execution shows a somewhat distinct behavior, indicated by the significant outlier in both Figures 8.14a and 8.14b, where the relative order size is one.

The NeuroBayes® machine learning algorithm is applied to a sample containing executed limit orders and their corresponding execution fraction. The target variable in this training is whether a limit order is going to be fully executed or whether only a fraction of it is going to be traded. As an indicator of the resulting separation capability of the algorithm on the training sample, the corresponding Gini-coefficients are provided in Table 8.4 for an in-sample as well as an out-of-sample application of the obtained expertise.

Exch.	In-Sample		Out-Of-Sample	
	Gini-Index	Max. Possible	Gini-Index	Max. Possible
LSE	10.0	30.2	9.47	30.14
EP	10.5	34.7	9.55	34.46

Table 8.4: The realized Gini-coefficients as well as the corresponding maximally achievable Gini-coefficients are provided in this table for in-sample and out-of-sample data.

The out-of-sample test is conducted in order to examine the algorithms forecasting capability for both the LSE and the EP. Here, it achieves a Gini-coefficient of 9.47% (where 30.14%is maximally achievable) for limit orders at the LSE whereas, for limit orders at the EP, it achieves 9.55% (where 34.46% is maximally achievable). These figures are in line with those of the training sample. Since there is no severe loss in the forecasting capability in the out-of-sample test, the algorithm turns out to present reliable forecasts.

The corresponding out-of-sample indicators for the algorithm's forecasting capability are provided in Figures 8.15 and 8.16 for the LSE and the EP, respectively. The Signal-Purity coincides in both cases with a diagonal through the origin, thus indicating that the forecasted probability is reliable.

(a) Signal-Background Distributions

(b) Signal-Purity Diagonal

(c) Efficiency-Purity Plot

(d) Gini-Plot

Figure 8.15: The above charts give an overview of the NeuroBayes® algorithm's forecasting capability whether a limit order is going to be fully executed. The charts result from an out-of-sample test on LSE data.

(a) Signal-Background Distributions

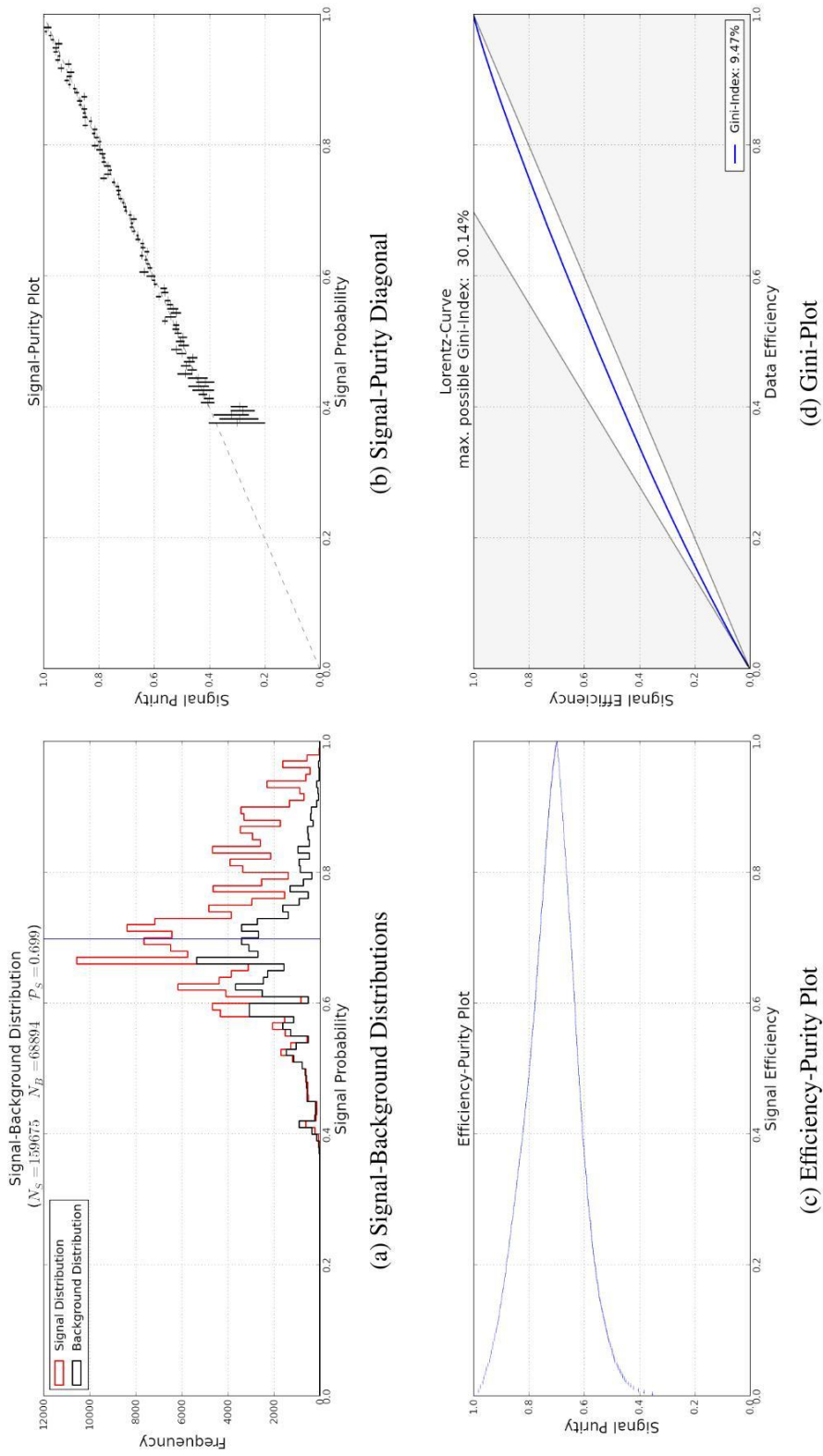(b) Signal-Purity Diagonal

(c) Efficiency-Purity Plot

(d) Gini-Plot
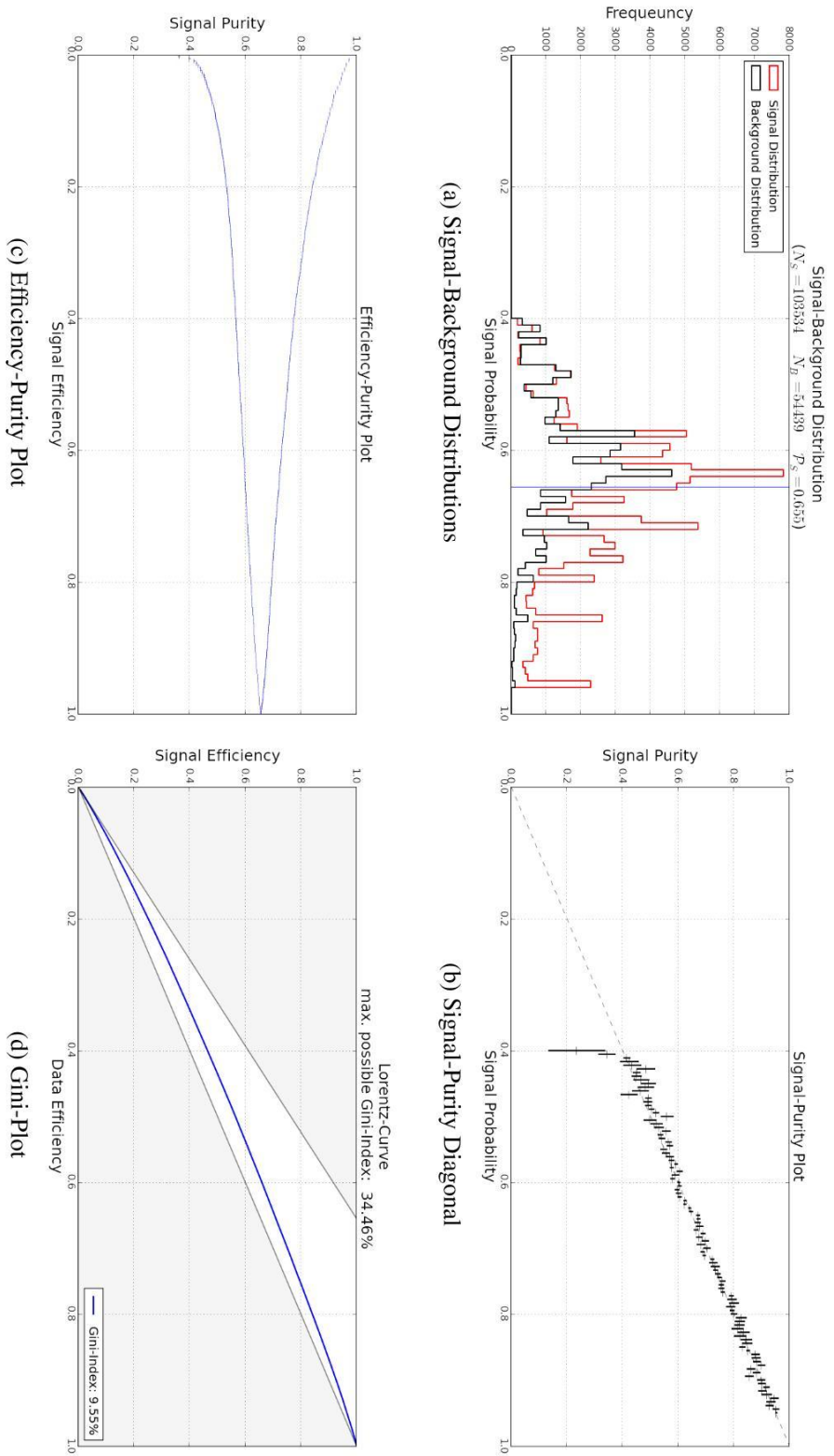
Figure 8.16: The above charts give an overview of the NeuroBayes® algorithm's forecasting capability whether a limit order is going to be fully executed. The charts result from an out-of-sample test on EP data.

Exch.	Side	L0	L1	L2	L3	L4	L5
EP	Ask-Side	61.91(0.30)	67.60(0.24)	65.91(0.55)	68.89(0.79)	69.00(1.23)	69.23(1.55)
	Bid-Side	61.83(0.30)	67.19(0.24)	67.52(0.54)	66.37(0.79)	67.27(1.19)	70.76(1.42)
LSE	Ask-Side	66.15(0.25)	72.20(0.18)	67.00(0.48)	66.23(0.75)	68.99(1.09)	67.94(1.33)
	Bid-Side	66.02(0.25)	73.35(0.18)	65.84(0.50)	66.06(0.75)	67.02(1.08)	66.72(1.31)

Table 8.5: The mean probabilities (in %) for the full execution of a limit order. These figures do not exhibit strong dispersion, thus, a significant dependency on the order book level is not prevalent.

Chapter 9

Conclusion

The subject of the second part of this thesis was the quantitative assessment of execution probabilities and execution fractions of passive limit orders.

First, a technique to forecast the execution of a limit order on a one minute time horizon was developed. One important prerequisite for this development was the construction of a software capable of constructing and analyzing order books from recorded historic data. This software was developed as a part of this work.

Two of the strongest descriptive quantities, the order book level in which a limit order is inserted and the stock which was intended to trade, were modeled separately. This led to the development of the Mean-Matrix-Model. Based on this approach, the machine learning algorithm NeuroBayes® was applied as a means to construct more sensitive probabilities for the execution of a limit order. This second step is carried out with the aim of finding corrections to the Mean-Matrix-Model conditional on observed market variables, order intrinsic properties as well as order book parameters. In particular, variables that are sensitive to asymmetries in the order book were found to contain strong dependencies on the execution of passive limit orders. Strong antisymmetric behavior was observed for the dependency of the execution probability between bid- and ask-type limit orders on those variables. Other important descriptive quantities were found to be variables sensitive to the trading activity shortly before a limit order is inserted.

This combined model was developed and tested on a collection of limit orders, selected from an assortment of the most liquid stocks at the London Stock Exchange as well as the Euronext Paris Stock Exchange. Altogether, 108 stocks and roughly 3 million limit orders were used in order to derive and test this approach.

The constructed model was tested on out-of-sample data in order to prove its capability of forecasting the execution of a limit order on data other than those which were used for the training of the NeuroBayes® algorithm. The resulting

forecasts were found to be consistent with the forecasting capability observed on the training sample. Hence, the combined model was proven to be applicable in practice.

The impact of the conditional corrections to the Mean-Matrix-Model was also examined in an in-sample and an out-of-sample analysis. In all cases, the application of the NeuroBayes$^®$ machine learning algorithm lead to significant improvements of the forecasting capability.

Finally, the possibility of a fractional execution of limit orders was further investigated. Significant stock and order book level dependencies were not observed in this case. Hence, the NeuroBayes$^®$ algorithm was directly applied in order to forecast the full execution of limit orders. It could be shown that the order size of a limit order relative to the already pending volume in the level of insertion is the most explanatory variable. An out-of-sample test was provided in order to prove the applicability of this model to practical applications.

The algorithm developed may be used in order to develop intraday trading strategies that rely on an accurate forecasting on limit order executions. Other time horizons may be straightforwardly implemented by following the steps carried out here but on a customized data set.

Bibliography

Abramowitz, M. and Stegun, I. (1965). *Handbook of Mathematical Functions: with Formulas, Graphs and Mathematical Tables*. Dover Publications.

Acerbi, C. and Tasche, D. (2001). Expected Shortfall: a natural coherent alternative to Value at Risk. *arXiv.org*.

Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, *4*(2-3), 115 – 158. High Frequency Data in Finance, Part 1.

Anderson, T. W. and Darling, D. A. (1954). A Test of Goodness-of-Fit. *Journal of the American Statistical Association*, *49*, 765–769.

Artzner, P., Delbaen, E. J.-M., F., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*.

Bai, X., Russell, J. R., and Tiao, G. C. (2003). Kurtosis of garch and stochastic volatility models with non-normal innovations. *Journal of Econometrics*, *114*(2), 349 – 360.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*(1), 3–30.

Basel Committee on Banking Supervision (2006). International convergence of capital measurements and capital standards.

Beck, A., Kim, Y., Feindt, M., and Rachev, S. T. (2011). Empirical analysis of arma-garch models in market risk estimation on high-frequency u.s. data. Paper in preparation, Karlsruhe Institute of Technology (KIT), Department of Economics and Business Engineering.

Beltran-Lopez, H. and Frey, S. (2006). Auction design in order book markets. Tech. rep.

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business and Economic Statistics*, *19*(4), 465–474.

Biais, B., Hillion, P., and Spatt, C. (1995). An empirical analysis of the limit order book and the order flow in the paris bourse. *The Journal of Finance*, *50*(5), 1655–1689.

Bianchi, M. L. (2009). *Tempered stable models in finance: theory and applications*. Ph.D. thesis, University of Bergamo.

Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, *81*(3).

Blobel, V. and Lohrmann, E. (1998). *Statistische und numerische Methoden der Datenanalyse*. Teubner Studienbuecher.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, *69*(3).

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). Arch modeling in finance : A review of the theory and empirical evidence. *Journal of Econometrics*, *52*(1-2), 5 – 59.

Bollerslev, T., Litvinova, J., and Tauchen, G. E. (2006). Leverage and Volatility Feedback Effects in High-Frequency Data. *Journal of Financial Econmetrics*, *4*(3), 354–384.

Bollerslev, T. and Woolridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Tech. rep.

Bouchaud, J.-P. (2002). An introduction to statistical finance. Tech. rep.

Bougerol, P. (1992). Stationarity of Garch processes and of some nonnegative time series. *Journal of Econometrics*, *52*(1-2), 115–127.

Boyarchenko, S. I. and Levendorskiĭ, S. Z. (2000). Option pricing for truncated lÉvy processes. *International Journal of Theoretical and Applied Finance*, *3*(3).

Breusch, T. S. and Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, *47*(5), 1287+.

Campbell, S. D. (2005). A Review of Backtesting and Backtesting Procedures. Tech. rep., Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2000). The Fine Structure of Asset Returns: An Empirical Investigation.

Cho, J. W. and Nelling, E. (2000). The probability of limit-order execution. *Financial Analysts Journal*, *56*(5).

Choirat, C. and Seri, R. (2009). Econometrics with Python. *J. Appl. Econ.*, *24*(4), 698–704.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, *39*(4), 841–62.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, *2*, 223 – 236.

Cont, R. (2009). Volatility Clustering in Financial Markets: Empirical Facts and Agent-Based Models. *Social Science Research Network Working Paper Series*.

Cont, R. (2011). Long range dependence in financial markets. In *Fractals in Engineering*. Springer, 159–180.

Cont, R., Stoikov, S., and Talreja, R. (2008). A Stochastic Model for Order Book Dynamics. *Social Science Research Network Working Paper Series*.

Deutsche Boerse Group (2010). Market model equities.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, *74*(366), 427–431.

Engle, R. (2001). The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives*, *15*(4), 157–168.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, *50*(4), 987–1007.

Engle, R. F. (2000). The econometrics of ultra-high-frequency data. *Econometrica*, *68*(1), 1–22.

Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1–50.

Engle, R. F. and Russell, J. R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica*, *66*(5), 1127–1162.

Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, *38*(1), 34–105.

Feindt, M. (2004). A neural bayesian estimator for conditional probability densities.

Foucault, T. (1999). Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets*, *2*(2), 99–134.

Fraenkle, J. (2010). *Theoretical and Practical Aspects of Algorithmic Trading*. Ph.D. thesis.

Glosten, L. R. (1994). Is the electronic open limit order book inevitable? *The Journal of Finance*, *49*(4), 1127–1161.

Godfrey, L. G. (1978). Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables. *Econometrica*, *46*(6).

Goodhart, C. A. E. and O'Hara, M. (1997). High frequency data in financial markets: Issues and applications. *Journal of Empirical Finance*, *4*(2-3), 73 – 114. High Frequency Data in Finance, Part 1.

Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, *1*(1), 15–29.

Handa, P. and Schwartz, R. A. (1996). Limit Order Trading. *The Journal of Finance*, *51*(5).

Hansen, B. E. (1994). Autoregressive Conditional Density Estimation. *International Economic Review*, *35*(3).

Harris, L. and Hasbrouck, J. (1996). Market vs. limit orders: The superdot evidence on order submission strategy. *The Journal of Financial and Quantitative Analysis*, *31*(2), 213–231.

Hosking, J. R. M. (1981). Fractional Differencing. *Biometrika*, *68*(1).

Hull, J., Hull, J., White, A., and White, A. (1998). Incorporating volatility updating into the historical simulation method for value at risk. *Journal of Risk*, *1*, 5–19.

Hull, J. C. (2007). *Risk Management and Financial Institutions*. Pearson.

Hurst, H. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, *116*, 770–799.

Itai, L. and Sussmann, A. (2009). US equity high frequency trading:Strategies, sizing, and market structure.

Kim, Y., Rachev, S., Bianchi, M., and Fabozzi, F. (2008a). Financial market models with Lévy processes and time-varying volatility. *Journal of Banking & Finance*, *32*(7), 1363–1378.

Kim, Y., Rachev, S., Chung, D., and Bianchi, M. (2008b). A modified tempered stable distribution with volatility clustering. Paper in preparation, Karlsruhe Institute of Technology (KIT), Department of Economics and Business Engineering.

Kim, Y., Rachev, S., Chung, D. M., and L., B. M. (2008c). The modified tempered stable distribution, GARCH-models and option pricing. *Probability and Mathematical Statistics*. To appear.

Kim, Y. S., Rachev, S., Bianchi, M., and Fabozzi, F. J. (2010a). Tempered stable and tempered infinitely divisible garch models. *Journal of Banking and Finance*, *34*(9), 2096 – 2109.

Kim, Y. S., Rachev, S., Bianchi, M. L., and Fabozzi, F. J. (2009). Computing var and avar in infinitely divisible distributions. *Yale ICF Working Paper No. 09-07*.

Kim, Y. S., Rachev, S. T., Bianchi, M. L., Mitov, I., and Fabozzi, F. J. (2010b). Time series analysis for financial market meltdowns. *Journal of Banking & Finance*.

Kim, Y. S., Rachev, S. T., Bianchi, M. L., Mitov, I., and Fabozzi, F. J. (2010c). Time series analysis for financial market meltdowns. *Journal of Banking & Finance*, *In Press, Corrected Proof*, –.

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated lévy flights towards the gaussian stochastic process. *Phys. Rev. E*, *52*(1), 1197–1199.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. Tech. rep.

Lee, J., Kim, T. S., and Lee, H. K. (2011). Return-volatility relationship in high frequency data: Multiscale horizon dependency. *Studies in Nonlinear Dynamics & Econometrics*, *15*(1).

Lindenberg, J. (1921). Exponentialgesetz in der wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, *15*.

Lo, A. W., MacKinlay, A. C., and Zhang, J. (2000). Econometric models of limit-order executions. *Social Science Research Network Working Paper Series*.

Longerstaey, J. and More, L. (1995). Introduction to RiskMetrics.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, *36*, 394.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91.

Marsaglia, G. and Marsaglia, J. (2004). Evaluating the Anderson-Darling Distribution. *Journal of Statistical Software*, *9*(2), 1–5.

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, *46*(253), 68–78.

Menn, C. (2005). A GARCH option pricing model with α-stable innovations. *European Journal of Operational Research*, *163*(1), 201–209.

Miller, L. H. (1956). Table of Percentage Points of Kolmogorov Statistics. *Journal of the American Statistical Association*, *51*(273).

Moody's Analytics (2011). Basel III New Capital and Liquidity Standards. Tech. rep., Moody's Analytics.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, *59*(2), 347–370.

NYSE Euronext (2011). Euronext instruction.

Omura, K., Tanigawa, Y., and Uno, J. (2000). Execution probability of limit orders on the tokyo stock exchange. *Social Science Research Network Working Paper Series*.

Rachev, S., Menn, C., and Fabozzi, F. (2005). *Fat-Tailed and Skewed Asset Return Distributions*. Wiley Finance.

Rachev, S. and Mittnik, S. (2000). *Stable Paretian Models in Finance*. Wiley.

Rachev, S., Mittnik, S., Fabozzi, F., Focardi, S. M., and Jasic, T. (2007). *Financial Econometrics*. Wiley Finance.

Racheva-Iotova, B. and Samorodnitsky, G. (2011). Long range dependence in heavy tailed stochastic processes.

Roll, R. (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance*, *39*(4), 1127–1139.

Rosenblatt, M. (1952). Remarks on a Multivariate Transformation. *The Annals of Mathematical Statistics*, *23*(3), 470–472.

Scherer, M., Rachev, S., Kim, Y., and Fabozzi, F. (2010). A fft based approximation of tempered stable and tempered infinitely divisible distributions. *Working Paper*.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464.

Smith, E., Farmer, J. D., Gillemot, L., and Krishnamurthy, S. (2003). Statistical theory of the continuous double auction. *Quantitative Finance*, *3*, 481–514.

So, M. and Yu, P. (2006). Empirical analysis of garch models in value at risk estimation. *Journal of International Financial Markets, Institutions and Money*, *16*(2), 180–197.

Stephens, M. A. (1974). Edf statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, *69*(347), 730–737.

Sun, W., Rachev, S. Z., and Fabozzi, F. (2008). Long-range dependence, fractal processes, and intra-daily data. In P. Bernus, J. Błażewics, G. Schmidt, M. Shaw, D. Seese, C. Weinhardt, and F. Schlottmann (Eds.), *Handbook on Information Technology in Finance*, Springer Berlin Heidelberg, International Handbooks on Information Systems. 543–585.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 817–838.

Zhang, S. S. and Riordan, R. (2011). Technology and Market Quality: The case of High Frequency Trading. In *Proceedings of the 19th European Conference on Information Systems*. Helsinki, Finland.

Appendix A

A.1   Profile Plots

Profile plots are a class of chart, which allow to graphically analyze two dimensional data $S = \{(x, y) | x, y \in \mathbb{R}\}$ in a very practical fashion. They are extremely useful tools in order to detect and visualize dependencies between two variables in a cloud of scattered data points.[1] Profile plots are comparable to scatter plots with the additional property that the x-axis is divided into bins, that is, equally spaced regions on the x-axis. The mean values and mean-errors of the y-components of the data points within one bin are determined and visualized.

To be more precise, assume that the ith bin has the left and right boundaries $x = l_i$ and $x = r_i$, ($l_i, r_i \in \mathbb{R}$). The data points which belong to this bin are given by $Y_i = \{y(x) | l_i \leq x < r_i\}$.

From this definition, the mean $\bar{y}_i$ and the mean-error $\hat{\sigma}_{\bar{y}_i}$ for the ith bin with N_i data points are constructed as

$$\bar{y}_i = \mathbb{E}[Y_i] = \frac{1}{N_i} \sum_{k=1}^{N_i} y_k \tag{A.1}$$

$$\hat{\sigma}_{\bar{y}_i} = \sqrt{\frac{Var[Y_i]}{N_i}} = \frac{1}{N_i} \sqrt{\sum_{k=1}^{N_i} (y_k - \bar{y}_i)^2}. \tag{A.2}$$

This procedure allows to extract the most relevant information from the data. An exemplary profile plot is shown in Figure A.1b. The profile plot is constructed from the cloud of data points shown in Figure A.1a. It is clearly evident that the dependency between x and y values can be easily observed in the right chart whereas the left chart provides almost no obvious evidence for a relation between the two quantities.

---

[1]In fact, there are situations, where it is impossible to visualize dependencies between x and y without the use of profile plots. For example, assume that $y \in \{0, 1\}$ and that x is a discrete variable, $x \in \mathbb{N}$. All one could observe in a scatter plot are data points at 0 and 1 for each x-value. No visual information would be available about how many data points with y = 1 (y = 0) lie on top of each other.
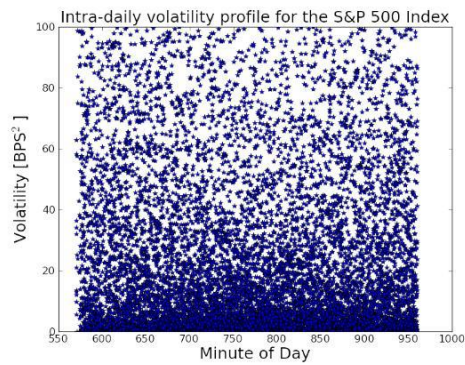
(a) Data displayed in a scatter plot.
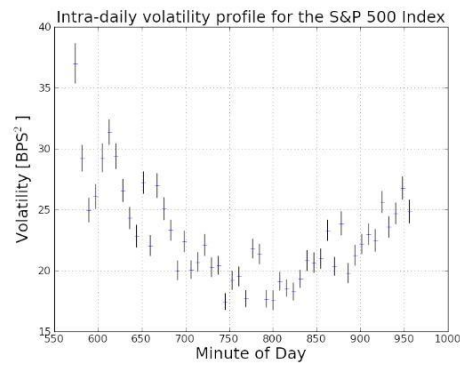
(b) Data displayed in a profile plot.

Figure A.1: Figure A.1a shows a cloud of data points with no obvious correlation. Figure A.1b shows the same data but aggregated in the fashion as described above. This type of chart is known as *profile plot*. In the right picture, it is easy to visually observe the dependency of y as a function of x.

A.2   Fat-Tails and Excess Kurtosis

Fat-tails are a property of probability distributions whose tails decline slower than exponentially (for example, by decaying hyperbolically). Fat-tail behavior results in a significant excess kurtosis, i.e. a kurtosis which is significantly larger than the normal distribution's kurtosis.
An exact definition for fat-tail behavior is given by

$$\lim_{X \to \infty} e^{\lambda x} Pr(X > x) \to \infty \quad \forall \lambda > 0.$$

Especially in the field of financial markets modeling, one common definition is:[2]

> The extreme events are contained in the tail region and if they appear with higher probability compared to that implied by the Gaussian distribution, we say that such a distribution possesses "fat tails".

Fat-tails are observed in many fields, such as economics, physics, computer sciences, etc. Their important implication in the field of financial markets is that extreme events, such as large negative price movements, have a higher probability to occur than suggested by the normal distribution assumption.

One way for observing fat-tailed behavior is by monitoring a distribution's excess kurtosis. The excess kurtosis, also known as *Fisher Kurtosis*, denotes the fourth moment of a probability density function, normalized with its squared variance, minus 3 (the normal distribution's kurtosis).

The sample kurtosis is constructed as follows:

$$\hat{\beta}_4 = \frac{\hat{\mu}_4}{\hat{\sigma}^4} - 3 = \frac{\frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^4}{\left(\frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^2\right)^2} \tag{A.3}$$

Various non-normal distributions can be applied in order to model fat-tailed distributions, such as, for example, Student-t, Stable Paretian non-Gaussian and tempered stable distributions. Figure A.2 compares the tail behavior for three probability density functions, which are frequently applied in financial time series modeling, namely, the normal distribution, the Student-t distribution and the standardized CTS distribution. These distributions, except for the normal distribution, are capable of modeling fat-tail behavior.

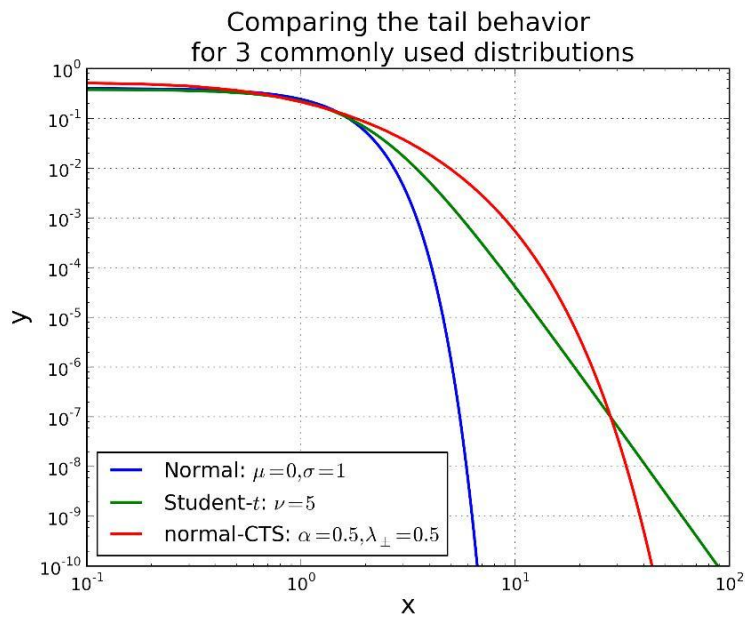[2]cited from Rachev *et al.* (2007))

Figure A.2: The above figure compares the tail behavior for the normal, the Student-t and the standardized CTS distribution. The normal distribution is rapidly declining whereas the Student-t distribution's tail decays hyperbolically and decreases rather slowly. The standardized CTS distribution's tail declines slower than the Student-t distribution for small x values. It then continues to decline more rapidly towards zero for higher x values and eventually crosses the Student-t distribution's tail function.

A.3   Kupiec and Christofferson Likelihood Ratio Test

The Kupiec Likelihood Ratio Test Kupiec (1995) is used in order to determine if an observed number of violations in a Value-at-Risk backtest is in conformity with an allowed number of violations. This is commonly referred to as a test for unconditional coverage of interval forecasts. A violation occurs if the observed return exceeds a pre-defined value, usually the $\eta = 1\%$ or the $\eta = 0.1\%$ quantile of the underlying return distribution. In the context of Value-at-Risk analysis, this number is denoted as $\text{VaR}(1\%)$ or $\text{VaR}(0.1\%)$ (Value-at-Risk was introduced in Section 2.6.2).

In a perfect world (or with an infinite number of trials), the underlying risk model would be appropriate, if out of N observations, N_η would be counted as violations while satisfying the following condition:

$$\frac{N_\eta}{N} = \text{VaR}(\eta),\qquad\text{(A.4)}$$

where $\text{VaR}(\eta)$ denotes the Value-at-Risk at the $\eta\%$ quantile.

In real world applications, every measurement and observation contains statistical fluctuations. Hence, depending on a confidence interval (usually 95% or 99%), a range of allowed violations is used in order to accept or reject a risk model.

Applying the idea of *Bernoulli's Trial*, the range of the allowed number of violations can be constructed as shown in the following. In a Bernoulli trial, there are two possible outcomes of an experiment, denoted as *success* and *failure*. If the probability of a success is p, the probability q for a failure is consequently $1 - p$. The probability of observing k successes in an experiment with n trials is given by the binomial distribution:

$$P(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}.\qquad\text{(A.5)}$$

Consequently, the empirically observed ratio $\hat{p} = \frac{N_\eta}{N}$ is tested against p in a likelihood ratio test, with the following test statistic:

$$LR_{uc} = -2\ln \frac{p^{N_\eta}(1-p)^{N-N_\eta}}{\left(\frac{N_\eta}{N}\right)^{N_\eta}\left(1 - \frac{N_\eta}{N}\right)^{N-N_\eta}}.\qquad\text{(A.6)}$$

LR_{uc} is asymptotically χ^2 distributed with one degree of freedom. If the probability of observing LR_{uc} is higher or equal to the pre-defined confidence interval, the Kupiec test accepts the risk model.

	$I_{t-1} = 0$	$I_{t-1} = 1$
$I_t = 0$	n_{00}	n_{01}
$I_t = 1$	n_{10}	n_{11}

Table A.1: Overview of the counting variables which are employed for testing the independence of violations in a Value-at-Risk backtest.

It has to be considered though, that the Kupiec test has two severe shortcomings: First, if the sample size on which a VaR backtest is conducted is small, the Kupiec test will not lead to reliable results.[3] This problem can obviously be circumvented by employing backtests with a longer history. Second, the Kupiec test only counts the number of violations but does not test their independence over time. That is, the Kupiec test is not sensitive for a clustered occurrence of violations.

To overcome this second drawback, Christoffersen (1998) introduces a second test statistic in order to cover the conditional occurrence of violations. To this end, the variable I_t is introduced as

$$I_t = \begin{cases} 1 & \text{if violation in } t \\ 0 & \text{if no violation in } t. \end{cases} \tag{A.7}$$

Subsequently, I_t and I_{t-1} are employed in order construct secondary auxiliary variables in a way as described in Table A.1. n_{xx} are counting variables which get increased by one whenever the respective conditions regarding I_t and I_{t-1} are fulfilled.

Using the definitions

$$\pi_0 = \frac{n_{01}}{n_{00} + n_{01}} \tag{A.8}$$

$$\pi_1 = \frac{n_{11}}{n_{10} + n_{11}} \tag{A.9}$$

$$\pi = \frac{n_{01} + n_{11}}{n_{00} + n_{01} + n_{10} + n_{11}}, \tag{A.10}$$

the test statistic in order to evaluate the conditional occurrences of violations is constructed as

$$LR_{ind} = -2\ln\left(\frac{(1-\pi)^{n_{00}+n_{10}}\pi^{n_{01}+n_{11}}}{(1-\pi_0)^{n_{00}}\pi_0^{n_{01}}(1-\pi_1)^{n_{10}}\pi_1^{n_{11}}}\right). \tag{A.11}$$

---

[3]This is the case within the current Basel II regulatory framework. It requires VaR backtests over a backtesting period of one year of daily data, leading to a sample size of just roughly 250, see Basel Committee on Banking Supervision (2006)(Part 2.II.3.ii.§178).

This quantity is asymptotically χ^2 distributed with one degree of freedom.

The joint likelihood ratio test statistic is constructed as

$$LR_{cc} = LR_{uc} + LR_{ind}. \tag{A.12}$$

It is asymptotically χ^2 distributed with two degrees of freedom. It has to be noted that in the case of investigating the combined probability, a low likelihood ratio in one test can conceal a high likelihood ratio in the other test. As a consequence, both requirements, unconditional coverage and independence, should be satisfied individually.

A comprehensive overview of backtesting methods is given in Campbell (2005).

A.4 Berkowitz Likelihood Ratio Test

The Berkowitz Likelihood Ratio (BLR) test focuses not only on the accurate forecasting of quantiles, such as the CLR tests. It rather tests a whole forecasted distribution function for its accuracy. It can therefore be used in order to determine if a probability distribution, which is applied in a financial model, will describe the actual underlying return distribution properly, based on historic realizations. This is especially of interest when considering the risk measure AVaR, since this risk measure is defined as the expectation value of the tail region. Thus, it is extremely sensitive to an appropriate modeling of the tail's shape.

Using the Rosenblatt transform (Rosenblatt (1952)), all realizations are transformed into an IID sequence of random variables. Rosenblatt's transformation is defined as

$$x_t = \int_{-\infty}^{y_t} \hat{f}(u)du = \hat{F}(y_t), \tag{A.13}$$

where y_t is the realized portfolio return and $\hat{f}(\cdot)$ is the forecasted probability density function at time $t-1$. Following Rosenblatt, the distribution of x_t is IID and uniform on the interval (0,1).

Berkowitz (2001) proposes the transformation of x_t into a $\mathcal{N}(0,1)$ distributed variable:

$$z_t = \Phi^{-1}\left[\int_{-\infty}^{y_t} \hat{f}(u)du\right] = \Phi^{-1}\left(\hat{F}(y_t)\right). \tag{A.14}$$

By testing z_t for normality (for example by employing likelihood ratio tests), Berkowitz shows that this test will discover inaccurate modeling of the actual return series. In other words, correct density forecasts automatically lead to a sequence of $\mathcal{N}(0,1)$ distributed z_t's.

A.4.1 Focusing on Large Losses

It has to be noted that the vast majority of realized returns, especially on short time scales, occurs in the center region of the underlying return distribution. On very short time scales, this typically results in return distributions which are narrowly peaked around the center region while exhibiting fat tails at the same time. A model which is not capable of describing these effects but still models the tail behavior accurately, should not be automatically rejected. It may still be extremely useful in risk estimation application such as, for example, in the Average Value-at-Risk framework.

For such cases, Berkowitz (2001) suggests a likelihood ratio test based on the above transformation, but which is only sensitive to the tail behavior of the return distribution. Consequently, all observations which do not fall in the tail region

which is confined by a pre-defined VaR value will be cut off. The adjusted test statistic is given by

$$z_t^* = \begin{cases} z_t & \text{if} \quad z_t \leq \text{VaR} \\ \text{VaR} & \text{if} \quad z_t > \text{VaR}. \end{cases}$$

z_t^* can be tested for normality in a likelihood ratio test framework, as prosed by Berkowitz.

Appendix B

B.1   Determining the Execution Status of a Limit Order

Once a plain limit order (without additional attributes such as validity time etc.) has been inserted into the order book it can either be canceled by the owner's intention or through an execution, that is, the limit order is matched against an arriving marketable order. Especially the latter case is of interest if the conditional execution probabilities of limit orders are investigated. A robust method is needed in order to tag limit orders in an ex-post analysis to determine whether they have been executed through incoming marketable orders. One approach which has been utilized in recent research projects is to monitor the bid- and ask-prices over a pre-defined time period. If the bid (ask) price hits or exceeds the limit price of a bid (ask) limit order under investigation, there is a certain chance that this order has been executed. This method is especially inaccurate when order book levels are populated with more than one limit order. If no exact information about the position of a limit within its level is available, such as within MBL data, this method may be applied. Similar methods to determine order executions can be found in, for example, Omura *et al.* (2000).

In contrast to the MBL data format, the MBO data format, as explained in Section 6.2, allows for a very robust method to determine the cause of the order removal. This method utilizes the microscopic order flow process and reported trades. Since each order can be tracked individually, the exact time, when the order under investigation is removed, can be extracted from the data. The time at which a limit order is removed from the order book is compared to trades occurring within a pre-defined time span $\Delta\tau$. This procedure is necessary since stock exchanges usually do not synchronize the data feed over which trades are reported with the data feed that updates the order book as a consequence of a trade.[1] If a limit order under consideration has been canceled from the first position in the order book and if it matches in size and price with a reported trade within $\Delta\tau$, the limit order has almost certainly been executed through this trade.[2] Note that this method is also very robust in order to determine if a trade was buyer- or seller-initiated. Figure B.1 illustrates the mechanism for tagging limit orders whether they have been executed or canceled.

---

[1]This phenomenon has been observed on London Stock Exchange and Euronext Paris data provided from Morningstar, an investment research company, as of June 2011, when this thesis was written.

[2]In fact, it is possible that two adjacent limit orders with exactly the same size are canceled subsequently within a very short time period. If one of these orders has been executed, it is not possible in this case to clearly distinguish which of one of them has been executed through the arriving marketable order.
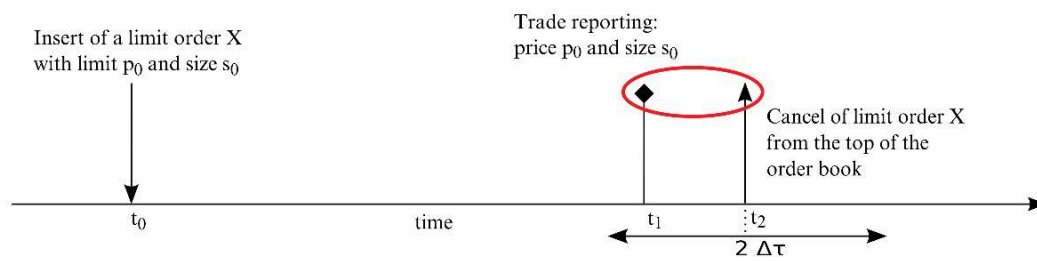
Figure B.1: Illustration of the process of how limit orders are tagged whether they have been executed or canceled by purpose.

B.2   Combining two Models with Bayes' Theorem

Bayes' theorem states that, given two random variables a and b, the conditional probability $P(a|b)$ is given by

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)}. \tag{B.1}$$

This relationship is often applied in data analysis for $P(t|\mathbf{x})$, where $\mathbf{x}$ denotes a vector containing a set of observations and t denotes a target event. The application of the theorem is then based on a *Prior-Model* which is reflected in $P(t)$.

In Section 8.3, two models are to be combined by adding up the corresponding likelihood ratios. This method is derived here by applying Bayes' theorem.

The combined probability may be expressed as

$$P^c(t|\mathbf{x}) = P(t|\mathbf{x}) = \frac{P(\mathbf{x}|t)P(t)}{P(\mathbf{x})}. \tag{B.2}$$

Here, $P(t)$ describes the probability of the target to occur as a result from the Mean-Matrix-Model. More generally, this could be any Prior-Model that is correct in average. Assume that P^b denotes the probability resulting from corrections on $P(t)$. $P^b(\mathbf{x}|t)$ itself can be transformed by applying Bayes' Theorem, resulting in

$$P^b(\mathbf{x}|t) = \frac{P^b(t|\mathbf{x})}{P^b(t)}P^b(\mathbf{x}). \tag{B.3}$$

Since $P^b(t|\mathbf{x})$ is a conditional correction on the Prior-Model, the unconditional target distribution must contain no information, that is, $P^b(t) = 0.5$. Otherwise, the Prior-Model would be systematically wrong.

The equations for $P^c(\bar{t}|\mathbf{x})$ and $P^b(\bar{t}|\mathbf{x})$ are constructed by simply replacing t with $\bar{t}$ in the above formulas. Substituting $P^b(\mathbf{x}|t)$ from Equation B.3 into Equation B.2, leads to

$$P^c(t|\mathbf{x}) = \frac{P^b(t|\mathbf{x})}{P^b(\mathbf{x})}P^b(t)\frac{P(t)}{P(\mathbf{x})}. \tag{B.4}$$

Dividing $P^c(t|\mathbf{x})$ from Equation B.4 with $P^c(\bar{t}|\mathbf{x})$ and considering that $P^b(t) = P^b(\bar{t}) = 1 - P^b(t) = 0.5$, leads to

$$\frac{P^c(t|\mathbf{x})}{P^c(\bar{t}|\mathbf{x})} = \frac{P^b(t|\mathbf{x})}{P^b(\bar{t}|\mathbf{x})}\frac{P(t)}{P(\bar{t})}. \tag{B.5}$$

By applying the logarithmic function to both sides of the above equation, and, thus constructing the likelihood-ratios, the above product decays into a sum, leading to:

$$\ln\frac{P^c(t|\mathbf{x})}{P^c(\bar{t}|\mathbf{x})} = \ln\frac{P^b(t|\mathbf{x})}{P^b(\bar{t}|\mathbf{x})} + \ln\frac{P(t)}{P(\bar{t})}. \tag{B.6}$$

This can now be expressed in likelihood-ratios, leading to

$$LR(P^c) = LR(P^b) + LR(P).$$

<div align="right">(B.7)</div>

B.3   Stock Universe - Company Overview

Table B.1: Overview of the stocks used in the analysis on order execution prob-
abilities. The exchanges are abbreviated as follows: EP = France, Paris; LSE =
UK, London.

CompanyName	Exchange	Sector
AIR LIQUIDE	EP	Basic Materials
RHODIA SA	EP	Basic Materials
ALCATEL LUCENT	EP	Communications
VIVENDI (EX-SOFIEE)	EP	Communications
FRANCE TELECOM SA	EP	Communications
RENAULT	EP	Consumer
PEUGEOT	EP	Consumer
PPR	EP	Consumer
MICHELIN	EP	Consumer
AIR FRANCE-KLM	EP	Consumer
ACCOR SA	EP	Consumer
PERNOD RICARD	EP	Consumer
DANONE	EP	Consumer
SANOFI	EP	Consumer
CARREFOUR	EP	Consumer
LOREAL	EP	Consumer
LVMH MOET	EP	Diversified
TECHNIP COMMON STOCKPAR VALUE	EP	Energy
TOTAL SA	EP	Energy
COMPAGNIE GENERALE DE GEOPHYSI	EP	Energy
NATIXIS	EP	Financial
AXA	EP	Financial
CREDIT AGRICOLE	EP	Financial
UNIBAIL - RODAMCO	EP	Financial
BNP PARIBAS	EP	Financial
SOC GENERALE	EP	Financial
LAFARGE	EP	Industrial
EUROPEAN AERONAUTIC DEFENSE	EP	Industrial
SCHNEIDER ELECT	EP	Industrial
ST.GOBAIN EUR16	EP	Industrial
VALLOUREC	EP	Industrial
BOUYGUES S	EP	Industrial
ALSTOM	EP	Industrial
VINCI	EP	Industrial
CAP GEMINI	EP	Technology
VEOLIA ENVIRONNEMENT	EP	Utilities
GDF SUEZ	EP	Utilities
ELECTRICITE DE FRANCE	EP	Utilities
LONMIN PLC	LSE	Basic Materials
BHP BILLITON PLC	LSE	Basic Materials
RIO TINTO PLC	LSE	Basic Materials

VEDANTA RESOURCES PLC	LSE	Basic Materials
KAZAKHMYS PLC	LSE	Basic Materials
XSTRATA PLC	LSE	Basic Materials
ANGLO AMERICAN PLC	LSE	Basic Materials
ANTOFAGASTA PLC	LSE	Basic Materials
YELL GROUP PLC	LSE	Communications
PEARSON PLC	LSE	Communications
BT GROUP	LSE	Communications
BRIT SKY BROAD	LSE	Communications
REED ELSEVIER PLCORD	LSE	Communications
WPP GROUP PLC	LSE	Communications
CABLE WIRELESS	LSE	Communications
VODAFONE GROUP	LSE	Communications
KINGFISHER PLC	LSE	Consumer
WOLSELEY PLC	LSE	Consumer
MITCHELLS BUTLERS PLC	LSE	Consumer
MARKS AND SPENCER GROUP PLC	LSE	Consumer
PERSIMMON	LSE	Consumer
PUNCH TAVERNS PLC ORD SHS	LSE	Consumer
INTERNATIONAL CONSOLIDATED AIR	LSE	Consumer
HOME RETAIL GROUP	LSE	Consumer
BARRATT DEV	LSE	Consumer
CARNIVAL PLCORD	LSE	Consumer
COMPASS GROUP PLC	LSE	Consumer
NEXT GROUP	LSE	Consumer
IMPERIAL TOBACCO	LSE	Consumer
RECKITT BENCKISER PLC	LSE	Consumer
SAB MILLER PLC	LSE	Consumer
SAINSBURY J PLC	LSE	Consumer
SHIRE PLC	LSE	Consumer
BRITISH AMER TOBAC	LSE	Consumer
ASTRAZENECA PLC	LSE	Consumer
MORRISON W SUPMARTORD	LSE	Consumer
DIAGEO PLC	LSE	Consumer
SMITH NEPHEW PLC	LSE	Consumer
ASSOC BRIT FOODS	LSE	Consumer
TESCO PLC	LSE	Consumer
GLAXOSMITHKLINE PLC ORD	LSE	Consumer
UNILEVER PLC	LSE	Consumer
TULLOW OIL ORD	LSE	Energy
BG GROUP	LSE	Energy
CAIRN ENERGY PLC	LSE	Energy
BP PLC	LSE	Energy
BARCLAYS	LSE	Financial
STANDARD CHARTERED	LSE	Financial
BRITISH LAND CO PLC	LSE	Financial
RBS	LSE	Financial
ICAP PLC	LSE	Financial
3I GROUP PLC LONDON	LSE	Financial

HAMMERSON PLC	LSE	Financial
LAND SECURITIES GROUP	LSE	Financial
LEGAL GENERAL GROUP	LSE	Financial
LLOYDS BANKING GROUP PLC	LSE	Financial
MAN GROUP PLC	LSE	Financial
CAPITAL SHOPPING CENTERS GROUP PLC	LSE	Financial
OLD MUTUAL PLC	LSE	Financial
PRUDENTIAL	LSE	Financial
AVIVA PLC	LSE	Financial
HSBC HOLDINGS	LSE	Financial
BAE SYSTEMS PLC	LSE	Industrial
ROLLS ROYCE GROUP	LSE	Industrial
SMITHS GROUP PLC	LSE	Industrial
UNITED UTILITIESGROUP PLC	LSE	Utilities
SCOTTISH SOUTHERN ENERGY	LSE	Utilities
CENTRICA PLC	LSE	Utilities
NATIONAL GRID PLC	LSE	Utilities
INTERNATIONAL POWER PLC	LSE	Utilities

Statement of Authorship / Selbstständigkeitserklärung

Ich versichere wahrheitsgemäß, die Dissertation bis auf die in der Abhandlung angegebene Hilfe selbstständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und genau kenntlich gemacht zu haben, was aus Arbeiten anderer und aus eigenen Veröffentlichungen unverändert oder mit Änderungen entnommen wurde.

Alexander Beck (Dipl. Phys.)

Acknowledgment