Statistical Analysis and Modelling of German and British Stock Return Processes from 1998 to 2007

Master Thesis submitted to

Prof. Dr. Wolfgang Härdle

Institute for Statistics and Econometrics
CASE - Center for Applied Statistics and Economics

Humboldt-Universität zu Berlin

by

Andrija Mihoci
(512584)

in partial fulfillment of the requirements
for the degree of
Master of Economics and Management Science

5 August 2008

Abstract

A comprehensive statistical analysis of return processes on the German and British stock market was carried out. Empirically, data for 40 selected companies and two market performance indices were collected for the period of ten years. The analysis shows that in the period under review the distributional characteristics were in line with stylized facts that hold on stock markets.

In addition, the efficiency market hypothesis was tested. The empirical findings support the fact that prices are walking in an unpredictable manner. Since all return processes show existence of first two moments, one find all return processes suitable for further modelling.

Keywords:
return process, stylized facts, efficiency market hypothesis, existence of moments

Dedication

This thesis is dedicated to my beloved wife Željka and my dear son Gabriel. Their support greatly motivated me in the course of doing this piece of work. For such reasons, I consider this piece of work as a symbol of their support to my academic effort!

Acknowledgement

An essential aspect of life is to appreciate those who sacrifice their time and energy in making contributions to the wellbeing of others. It is therefore no mistake that people usually say "give thanks to whom it is due".

Upon reflection of my entire experience in the MEMS programme, especially with writing this thesis, I recount instances of challenges and in such situations there have always been people who made meaningful interventions that brought me thus far. Today, I am at the ultimate end of this programme because of such people.

First, I am highly thankful to God, who gave me the strength to sail through the waters of the programme.

On the human point of view, I am indebted to my supervisor Prof. Dr. Wolfgang Härdle who has been, and continues to be, a source of inspiration to me in academic matters. There is no doubt whatsoever, that he is my academic mentor. His advice and recommendations have been essential to the start and completion of this thesis. For these reasons I appreciate him greatly.

Furthermore, I recognize and appreciate Professors, Lectures and colleges at School of Economics and Business in Zagreb, especially those at Department of Statistics. Their support has been a constant source of motivation through my academic experience at Humbold University.

The relentless support of my darling wife continues to bear significance to the success story of my academic endeavors. She and our son Gabriel continue to be a strong pillar within and without academia.

Let me take this opportunity of thanking my parents Ivan and Anđela, and my wife's parents Stjepan and Jelena, whose individual roles to my success cannot be downplayed. Sister Ivana, my wife's sister Slađana, nephews David and Dino and all our relatives are well appreciated for their respective roles.

In closing, I recognize and appreciate my numerous friends in Croatia and Germany, especially Dražen, Valentina, Marko, Josip, Marijan, Vladimir, Razvan, Tomáš and Augustine! To them I am thankful!

Contents

1 Introduction **1**

2 Data Description and Certain Definitions **3**

2.1 Data Description 3

2.2 Return Processes 6

2.3 Univariate Time Series of Analyzed Indices 7

3 Statistical Analysis of Return Processes **9**

3.1 Empirical Characteristics of Return Processes 9

3.2 Distribution of Return Processes: Statistical Concepts 11

3.2.1 Stationarity 13

3.2.2 Heteroscedasticity 20

3.2.3 Asymmetry 22

3.2.4 Overkurtosis 23

3.2.5 Non-normality 24

3.3 Linear Time Dependence of Return Processes 25

3.3.1 Empirical Autocorrelation Functions 26

3.3.2 Portmanteau Tests 27

4 Stock Market Efficiency **31**

5 Behaviour of Extreme Returns **37**

5.1 Tail Index Regression 37

5.2 Least Square and Hill Estimators 39

6 Conclusion **42**

List of Figures

2.1 DAX and FTSE 100 index values from 1 Jan 1998 to 31 Dec 2007 7

3.1 DAX and FTSE 100 index returns from 1 Jan 1998 to 31 Dec 2007 15

3.2 DAX and FTSE 100 estimated standard deviations of returns (mooving window 100 days) from 1 Jan 1998 to 31 Dec 2007 . 21

3.3 DAX and FTSE 100 index, boxplots of returns 22

3.4 DAX and FTSE 100 index, histograms of returns 24

3.5 DAX and FTSE 100 index, Normal-QQ plots of returns 25

3.6 Empirical autocorrelation functions of DAX and FTSE 100 plain, squared and absolute returns from 1 Jan 1998 to 31 Dec 2007 27

5.1 The right side of the logged empirical distribution of the DAX and FTSE 100 returns from 1 Jan 1998 to 31 Dec 2007 38

List of Tables

2.1 Stock prices and market capitalization of selected companies on the FSE on 31 Dec 2007 4

2.2 Stock prices and market capitalization of selected companies on the LSE on 31 Dec 2007 5

3.1 Estimated mean and standard deviation of daily returns for DAX index and selected stocks on FSE, and extreme returns from 1 Jan 1998 to 31 Dec 2007, in percentages 11

3.2 Estimated mean and standard deviation of daily returns for FTSE 100 index and selected stocks on LSE, and extreme returns from 1 Jan 1998 to 31 Dec 2007, in percentages 12

3.3 Estimated skewness, kurtosis and Bera-Jarque test statistic of daily returns for DAX index and selected stocks on FSE from 1 Jan 1998 to 31 Dec 2007 13

3.4 Estimated skewness, kurtosis and Bera-Jarque test statistic of daily returns for FTSE 100 index and selected stocks on LSE from 1 Jan 1998 to 31 Dec 2007 14

3.5 ADF test statistics for returns in the FSE from 1 Jan 1998 to 31 Dec 2007 16

3.6 ADF test statistics for returns in the LSE from 1 Jan 1998 to 31 Dec 2007 17

3.7 KPSS test statistics for returns in the FSE from 1 Jan 1998 to 31 Dec 2007 18

3.8 KPSS test statistics for returns in the LSE from 1 Jan 1998 to 31 Dec 2007 19

3.9 Ljung-Box test statistics for plain, squared and absolute returns for DAX index and selected stocks on FSE from 1 Jan 1998 to 31 Dec 2007, $m = 8$ 28

3.10 Ljung-Box test statistics for plain, squared and absolute returns for FTSE 100 index and selected stocks on LSE from 1 Jan 1998 to 31 Dec 2007, $m = 8$ 29

4.1 ADF test statistics for logged index values and stock prices in the FSE from 1 Jan 1998 to 31 Dec 2007 32

4.2 ADF test statistics for logged index values and stock prices in the LSE from 1 Jan 1998 to 31 Dec 2007 33

4.3 KPSS test statistics for logged index values and stock prices in the FSE from 1 Jan 1998 to 31 Dec 2007 34

4.4 KPSS test statistics for logged index values and stock prices in the LSE from 1 Jan 1998 to 31 Dec 2007 35

5.1 Least Square (LS) and Hill estimators of the tail index α for returns on DAX index and selected stocks on FSE with m observations used for the estimation 39

5.2 Least Square (LS) and Hill estimators of the tail index α for returns on FTSE 100 index and selected stocks on FSE with m observations used for the estimation 40

Chapter 1

Introduction

In order to study the empirical characteristics of return processes, many statistical concepts were introduced and statistical test procedures were carried out. These concepts and procedures appear to be important for proving the existence of these analyzed empirical characteristics. Here, the emphasis is on empirical provision of results, which can serve as a good starting point for further modelling of stock market return processes, empirical return distributions, volatility processes, etc.

The second chapter describes the data set and provides important definitions that are used in the succeeding chapters. In particular, two stock markets were analyzed, namely the German and the British stock market. The data was collected only for the largest companies that were traded on the Frankfurt Stock Exchange (FSE) and the London Stock Exchange (LSE). On each stock market 20 largest companies and one performance index were selected for further analysis. Furthermore, daily index values and closing prices were collected for the period of ten years, starting from 1 January 1998 until 31 December 2007.

A comprehensive statistical analysis of return processes was presented in chapter three. There one can find statistical test procedures explained in more detail and also a careful study of stylized facts that one observes on a particular stock market. The chapter starts with the empirical investigation of characteristics of the return processes and continues with a comprehensive analysis of the return distributions. Finally, the linear time (in)dependence of return processes was investigated more closely.

Analysis of the weak form efficiency of stock markets was considered in chapter four for both markets. In particular, the Efficiency Market Hypothesis was tested. The weak form efficiency is based upon the fact that the information set of an investor at some day is given by the set of historical stock prices and index values until the selected day on a particular stock market. Additionally, it is assumed that stock prices and index values perfectly reflect the current market situation, and that there does not exist arbitrage opportunities.

An empirical analysis of extreme returns and risk was carried out in chapter five. In particular, the tail index regression was introduced and the regression procedure was fitted to the collected dataset. The modelling of tails of return distributions leads to interesting conclusions about the existence of moments, in particular the variance. This additionally motivates the selection of models that are dealing with volatility processes.

In the computation and provision of the results various statistical software packages were used, in particular *XploRe*, *MATLAB* and *R*.

Chapter 2

Data Description and Certain Definitions

This chapter seeks to give a description of the dataset which was used in the empirical analysis. Also, important definitions of return processes were given in the sequel. These return processes form the basis for further modelling. Univariate time series of analyzed stock market indices closes this chapter.

2.1 Data Description

For the purposes of the statistical analysis daily data were considered. The dataset was collected mainly from *Datasteam* and *EcoWin*. In addition, information published on the web pages of the analyzed stock markets was used (www.deutsche-boerse.de, www.londonstockexchange.com).

Besides the two performance indices, DAX and FTSE 100, on each analyzed stock market the largest 20 companies were selected and included in the statistical analysis. For selecting the companies, the following criteria were used:

- Companies are primarily listed on the stock market,
- Largest market capitalization on 31 Dec 2007,
- Companies are listed in the whole period, and

- Dual listed companies are included.

In the period under review, the dataset consists of 2608 observations of daily closing stock prices and index values. These observations were collected from Monday to Friday. For nontrading days, the previous listed prices and index values were used. During this period, on the Frankfurt Stock Exchange (FSE) there were 2543 trading days, whereas on the London Stock Exchange (LSE) one could notice 2524 trading days.

The stock prices and the market capitalizations of the largest 20 companies per stock market on 31 Dec 2007 are presented in Tables 2.1 and 2.2 for FSE and LSE, respectively. These tables contain also the corresponding rank of the companies with respect to the market capitalization.

	Price	Capitalization	
	in EUR	Rank	Bill. of EUR
ADIDAS	50.93	18	10.4
ALLIANZ	148.08	3	66.6
BASF	101.26	4	49.2
BAYER	62.37	5	47.8
BMW	42.73	15	13.6
COMMERZBANK	26.29	12	15.8
CONTINENTAL	88.74	14	14.4
DEUTSCHE BANK	89.48	6	44.8
DEUTSCHE TELEKOM	15.03	8	41.9
EON	145.61	1	80.1
HENKEL	38.35	20	6.8
LINDE	90.49	16	11.6
LUFTHANSA	18.19	19	8.3
MAN	114.16	17	11.2
MUENCHENER RUECK	132.85	10	29.0
RWE	96.45	7	44.7
SAP	35.62	9	31.5
SIEMENS	108.42	2	79.0
THYSSENKRUPP	38.36	13	14.8
VOLKSWAGEN	156.43	11	22.1

Table 2.1: Stock prices and market capitalization of selected companies on the FSE on 31 Dec 2007

	Price	Capitalization	
	in £	Rank	Bill. of £
ASTRAZENECA	2164.00	13	3.2
BARCLAYS	504.00	12	3.3
BG GROUP	1150.00	9	3.9
BHP BILLITON	1546.00	11	3.5
BP	615.00	2	11.7
BRITISH AMERICAN TOBACCO	1965.00	8	4.0
BT GROUP	272.75	19	2.2
DIAGEO	1080.00	14	2.8
GLAXO SMITH KLINE	1279.00	5	7.1
HBOS	735.00	15	2.8
HSBC	842.00	3	10.0
LLOYDS TSB GROUP	472.00	16	2.7
NATIONAL GRID	834.00	20	2.2
RIO TINTO	5317.00	6	5.3
ROYAL BANK OF SCOT. GROUP	444.00	7	4.5
ROYAL DUTCH SHELL A	28.75	1	13.4
STANDARD CHARTERED	1844.00	17	2.6
TESCO	477.25	10	3.8
UNILEVER	1890.00	18	2.5
VODAFONE GROUP	187.80	4	9.9

Table 2.2: Stock prices and market capitalization of selected companies on the LSE on 31 Dec 2007

On FSE four large companies were not included in the analysis, namely DAIMLER, DEUTSCHE BOERSE, DEUTSCHE POST and HYPO REAL ESTATE with market capitalizations of 6.5, 27.2, 19.7 and 7.3 Bill. EUR, respectively. The reason is that for those companies there were no trade in the whole period under review. With the same reasoning as above, two large companies on LSE were excluded from the analysis, namely ANGLO-AMERICAN and XSTRATA, with market capitalizations of 4.1 and 3.4 Bill. £, respectively.

As one can calculate, the overall market value of all 20 selected companies was 643.6 Bill. EUR on FSE, and 101.2 Bill. £ on LSE. Additionally, one can find that the largest five selected companies on each stock market count for slightly more than 50% (50.1% on FSE, 51.5% on LSE) of the market capitalization of all selected companies per stock market.

2.2 Return Processes

The basic process examined in our analysis is the return process. In the following, various return processes were theoretically defined, in particular, the daily simple gross return process without dividends, the daily simple net return process and the continuously compounded return process, also known as log return process.

For given individual daily stock prices or index values at time point t, denoted by S_{it}, we can define the daily simple gross return without dividends, $1+R_{it}$, as

$$1 + R_{it} = \frac{S_{it}}{S_{i,t-1}}. \tag{2.1}$$

After subtracting one from both sides, one would get the daily simple net return in time period t for individual stock or stock market index i, R_{it},

$$R_{it} = \frac{S_{it}}{S_{i,t-1}} - 1. \tag{2.2}$$

Continuously compounded returns or daily log returns for individual stocks and indices at time t, r_{it}, approximate the net returns. This approximation is done mathematically by taking the natural logarithm of the daily simple gross returns. This definition of daily log returns is often used in statistical application in finance, and is suitable for modelling of the return processes. Thus, the log returns are defined by

$$r_{it} = \log\left(1 + R_{it}\right) = \log \frac{S_{it}}{S_{i,t-1}}. \tag{2.3}$$

The continuously compounded returns r_{it} enjoy some advantages over the simple net returns R_{it}. For example, the continuously compounded multiperiod return is simply the sum of continuously compounded one-period returns. Also, statistical properties of log returns are more tractable, see [2].

2.3 Univariate Time Series of Analyzed Indices

For the analyzed stock market indices, DAX and FTSE 100, two discrete univariate time series of closing values $\{S_{it}\}_{t=1,2,...,2608;\, i=1,2}$ are shown in Figure 2.1. One can compare both time series since both graphs are plotted with the same scale.

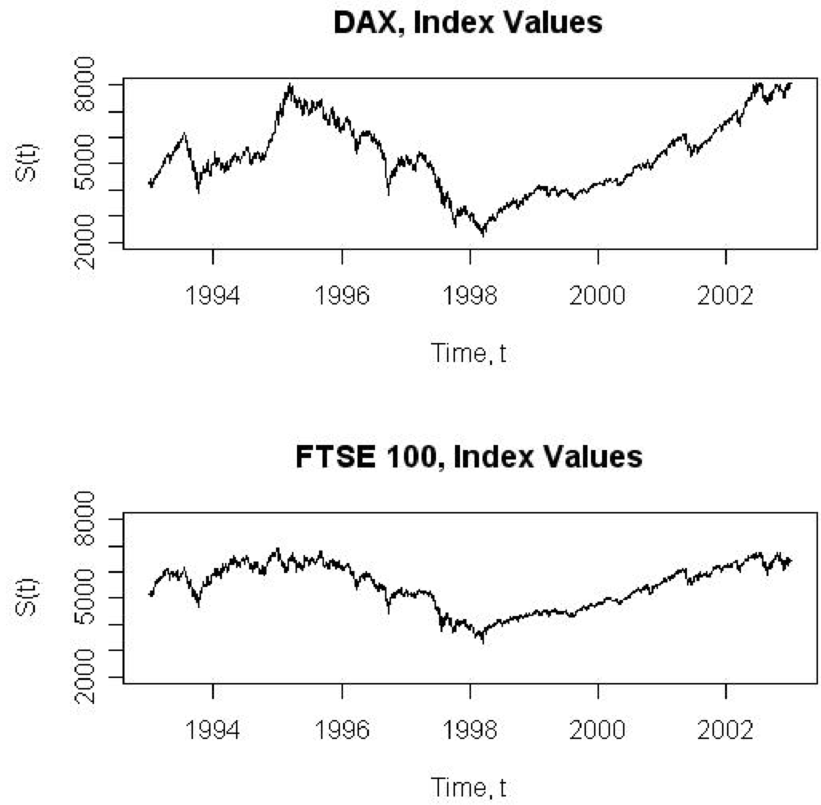

Figure 2.1: DAX and FTSE 100 index values from 1 Jan 1998 to 31 Dec 2007

It is obvious that both time series clearly show a nonstationary behaviour in the period under review. When considering the spread of the index values, one can conclude that the variation of the closing index level is evidently higher for DAX than for FTSE 100.

In order to give a more detailed explanation of the time series of analyzed index values, the sample period of ten years was divided in three subperiods

- the first representing years 1998 and 1999, second subperiod covering the years 2000, 2001 and 2002, and the last period representing the last five years.

In the first subperiod one can observe either a slightly upward trend with low stochastic component or a stationary process with relatively large absolute disturbances. After 1999 both markets entered a period of low index values. In this second subperiod, one observe the lowest index values in the whole sample period, especially at end of 2002 and on begin of 2003.

After the recovering phase in 2003, the index values clearly indicate an upward trend in the last subperiod. However, at end of 2007 one can observe a short stationary period. This short period may indicate a possible change in the time series behaviour of the index levels.

Chapter 3

Statistical Analysis of Return Processes

On a particular stock market there exists several stylized facts. To empirically support these findings, the return processes were analyzed in more detail. In the succeeding paragraphs, various statistical concepts were introduced, all supported by statistical test procedures.

This chapter starts with the analysis of empirical characteristics of the return processes. It continues with a comprehensive analysis of the empirical distribution of returns. At the end of this chapter the linear time dependence of the return processes was analyzed more carefully.

3.1 Empirical Characteristics of Return Processes

The important empirical characteristics of return processes (mean, standard deviation, skewness and kurtosis) were analyzed more detailed in the succeeding paragraphs.

In order to estimate the mean of the return processes for all individual stocks and indices, μ_i, one uses a sample of data. In our case, 2607 return values were obtained for all individually analyzed stocks and indices. If we denote the sample size by n, the unbiased estimator for the mean for stock (index)

i for a determined time period is given by

$$\hat{\mu}_i = \frac{\sum_{t=1}^{n} r_{it}}{n}. \tag{3.1}$$

In order to estimate the unknown standard deviation of the return processes for the whole period for individual stocks and indices, σ_i, one uses the estimator

$$\hat{\sigma}_i = \sqrt{\frac{\sum_{t=1}^{n} (r_{it} - \hat{\mu}_i)^2}{n-1}}. \tag{3.2}$$

The mean and standard deviation were estimated for return processes of individual stocks and indices on FSE and LSE. The estimated values are summarized in Tables 3.1 and 3.2 for FSE and LSE, respectively. These tables also include the observed extreme daily returns in the period under review. All values are given in percentages.

The parameters skewness, S_i, and kurtosis, K_i, as well as the Bera-Jarque test statistic, BJ_i, can be estimated for the return processes of individual stocks and indices with the following estimators (given in [1])

$$\hat{S}_i = (n-1)^{-1} \sum_{t=1}^{n} (r_{it} - \hat{\mu}_i)^3 / \hat{\sigma}_i^3 \tag{3.3}$$

$$\hat{K}_i = (n-1)^{-1} \sum_{t=1}^{n} (r_{it} - \hat{\mu}_i)^4 / \hat{\sigma}_i^4 \tag{3.4}$$

$$\widehat{BJ}_i = n \left(\frac{\hat{S}_i^2}{6} + \frac{\left(\hat{K}_i - 3\right)^2}{24} \right). \tag{3.5}$$

In Tables 3.3 and 3.4 one can find the estimated skewness, kurtosis and Bera-Jarque test statistic for analyzed indices and selected stocks on FSE and LSE, respectively. This parameter estimates will be used later for testing purposes.

	$\hat{\mu}_i$	$\hat{\sigma}_i$	$r_{i,min}$	$r_{i,max}$
DAX	0.02	1.56	-8.87	7.55
ADIDAS	0.02	1.97	-11.30	10.28
ALLIANZ	-0.01	2.23	-15.68	13.81
BASF	0.04	1.69	-8.04	10.74
BAYER	0.03	2.09	-18.44	32.25
BMW	0.03	2.14	-12.38	11.19
COMMERZBANK	-0.01	2.23	-13.23	18.03
CONTINENTAL	0.06	2.01	-11.69	12.83
DEUTSCHE BANK	0.01	2.08	-12.35	12.77
DEUTSCHE TELEKOM	-0.01	2.45	-14.06	12.65
EON	0.03	1.80	-9.16	9.94
HENKEL	0.03	1.75	-8.41	10.79
LINDE	0.02	1.80	-9.77	10.50
LUFTHANSA	0.00	2.19	-15.19	16.41
MAN	0.06	2.23	-11.64	11.40
MUENCHENER RUECK	-0.01	2.30	-17.18	16.53
RWE	0.03	1.80	-8.15	13.22
SAP	0.02	2.97	-18.39	23.52
SIEMENS	0.04	2.27	-10.48	15.66
THYSSENKRUPP	0.03	2.15	-16.59	11.36
VOLKSWAGEN	0.04	2.11	-9.65	9.37

Table 3.1: Estimated mean and standard deviation of daily returns for DAX index and selected stocks on FSE, and extreme returns from 1 Jan 1998 to 31 Dec 2007, in percentages

3.2 Distribution of Return Processes: Statistical Concepts

Various statistical concepts are strong in explaining the behaviour of return processes and their distributions on a particular market in some time period. They serve as a starting point in investigating the movements of a stock market in certain time and its empirical phenomena that may exist. In this sense, these concepts can reflect the existence of stylized facts on stock markets. With help of established test procedures one can verify this claims.

Thus, various statistical concepts were introduced. Among others, the con-

	$\hat{\mu}_i$	$\hat{\sigma}_i$	$r_{i,min}$	$r_{i,max}$
FTSE 100	0.01	1.14	-5.89	5.90
ASTRAZENECA	0.00	1.83	-12.57	12.36
BARCLAYS	0.01	2.13	-8.98	9.37
BG GROUP	0.07	1.91	-9.50	9.90
BHP BILLITON	0.09	2.40	-11.72	16.42
BP	0.02	1.70	-8.18	9.44
BRITISH AMERICAN TOBACCO	0.07	2.05	-12.20	32.26
BT GROUP	-0.01	2.33	-19.82	11.60
DIAGEO	0.02	1.72	-9.22	10.79
GLAXO SMITH KLINE	-1.00	1.79	-13.89	18.81
HBOS	-1.00	2.04	-9.83	19.77
HSBC	0.02	1.73	-14.33	11.27
LLOYDS TSB GROUP	-0.02	2.07	-9.23	11.87
NATIONAL GRID	0.04	1.53	-15.05	12.56
RIO TINTO	0.08	2.18	-7.34	19.68
ROYAL BANK OF SCOT. GROUP	0.02	2.13	-15.26	10.63
ROYAL DUTCH SHELL A	0.01	1.62	-10.32	8.58
STANDARD CHARTERED	0.04	2.32	-16.65	15.15
TESCO	0.04	1.71	-8.43	10.40
UNILEVER	0.02	1.72	-14.82	8.84
VODAFONE GROUP	0.03	2.47	-11.70	13.71

Table 3.2: Estimated mean and standard deviation of daily returns for FTSE 100 index and selected stocks on LSE, and extreme returns from 1 Jan 1998 to 31 Dec 2007, in percentages

cepts of stationarity, heteroscedasticity, asymmetry, overkurtosis and non-normality were carried out in more detail.

All concepts are characterized by a short explanation and a graphical illustration, the latter serving as a motivation for the test procedure. Thereafter, well known statistical test procedures were carried out, and finally the results were interpreted in the concrete case and conclusions about the return processes and their distribution were made.

	$\hat{S}_i$	$\hat{K}_i$	$\widehat{BJ}_i$
DAX	-0.99	5.99	981
ADIDAS	0.49	6.49	1325
ALLIANZ	-0.11	8.11	2841
BASF	0.76	5.76	831
BAYER	0.96	28.96	73596
BMW	0.21	6.21	1118
COMMERZBANK	0.16	8.16	2899
CONTINENTAL	0.50	6.50	1332
DEUTSCHE BANK	-0.32	6.32	1199
DEUTSCHE TELEKOM	0.19	6.19	1112
EON	0.57	5.57	718
HENKEL	0.27	6.27	1171
LINDE	0.82	5.82	868
LUFTHANSA	-0.31	7.31	2020
MAN	-0.02	5.02	443
MUENCHENER RUECK	-0.25	9.25	4246
RWE	0.41	6.41	1306
SAP	0.60	9.60	4792
SIEMENS	0.42	5.42	648
THYSSENKRUPP	-0.13	6.13	1066
VOLKSWAGEN	-0.11	5.11	488

Table 3.3: Estimated skewness, kurtosis and Bera-Jarque test statistic of daily returns for DAX index and selected stocks on FSE from 1 Jan 1998 to 31 Dec 2007

3.2.1 Stationarity

The concept of stationarity plays an important role in time series analysis and modelling. Let us observe the time series of returns on DAX and FTSE 100 index (Figure 3.1).

The spread is larger for DAX index returns than for FTSE 100 index returns implying that the volatility of the return process is higher for the DAX index. Thus, one can expect higher risk when investing in a portfolio build from blue chips on FSE.

A stochastic process is called covariance, or weakly, stationary, if its mean and autocovariance function are time invariant. In order to show whether a

	$\hat{S}_i$	$\hat{K}_i$	$\widehat{BJ}_i$
FTSE 100	-0.66	5.66	783
ASTRAZENECA	-0.55	7.55	2246
BARCLAYS	0.15	5.15	505
BG GROUP	0.16	5.16	515
BHP BILLITON	0.21	6.21	1180
BP	0.42	5.42	635
BRITISH AMERICAN TOBACCO	0.47	30.47	83374
BT GROUP	-0.19	7.19	1904
DIAGEO	0.35	7.35	2058
GLAXO SMITH KLINE	0.49	11.49	7861
HBOS	0.02	9.02	4003
HSBC	0.36	8.36	3121
LLOYDS TSB GROUP	0.07	6.07	1045
NATIONAL GRID	-0.88	10.88	6747
RIO TINTO	0.64	6.64	1518
ROYAL BANK OF SCOT. GROUP	-0.84	6.84	1599
ROYAL DUTCH SHELL A	-0.24	6.24	1176
STANDARD CHARTERED	0.52	7.52	2218
TESCO	0.84	5.84	880
UNILEVER	-0.33	9.33	4449
VODAFONE GROUP	0.55	5.55	721

Table 3.4: Estimated skewness, kurtosis and Bera-Jarque test statistic of daily returns for FTSE 100 index and selected stocks on LSE from 1 Jan 1998 to 31 Dec 2007

process is covariance stationary or not, various statistical test procedures are available. Here, the Augmented Dickey Fuller (ADF) tests and the KPSS tests from Kwiatkowski, Phillips, Schmidt and Shin were used. Thus, the following test procedures were considered:

- ADF test for returns without time trend,
- ADF test for returns with time trend,
- KPSS test for returns without time trend,
- KPSS test for returns with time trend.

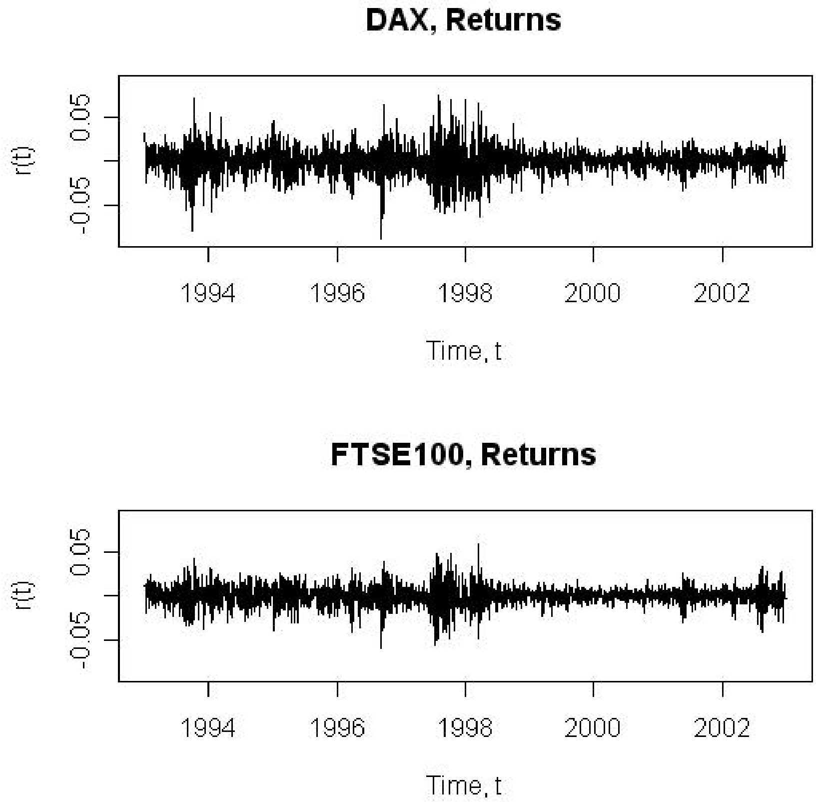

Figure 3.1: DAX and FTSE 100 index returns from 1 Jan 1998 to 31 Dec 2007

These procedures are explained more carefully in the sequel. After their application to returns on DAX and FTSE 100 indices and selected stocks, one gets the ADF test statistics, summarized in Tables 3.5 and 3.6 for the FSE and LSE, respectively, and the KPSS test statistics, shown in Tables 3.7 and 3.8 for FSE and LSE, respectively.

ADF test for returns without time trend

Consider the following regression model:

$$\Delta r_t = c + (\alpha - 1)\, r_{t-1} + \sum_{i=1}^{p} \Delta r_{t-i} + \epsilon_t. \tag{3.6}$$

p	Without time trend 0	4	With time trend 0	4
DAX	-52.24	-23.37	-52.24	-23.36
ADIDAS	-47.29	-24.15	-47.30	-24.15
ALLIANZ	-47.57	-24.10	-47.58	-24.10
BASF	-49.89	-24.10	-49.88	-24.06
BAYER	-50.32	-23.51	-50.33	-23.51
BMW	-48.86	-23.20	-48.87	-23.20
COMMERZBANK	-47.86	-21.79	-47.87	-21.79
CONTINENTAL	-50.84	-23.31	-50.83	-23.27
DEUTSCHE BANK	-47.62	-23.97	-47.63	-23.97
DEUTSCHE TELEKOM	-49.77	-23.41	-49.78	-23.41
EON	-50.98	-24.07	-50.98	-24.05
HENKEL	-50.42	-23.96	-50.43	-23.95
LINDE	-53.54	-23.80	-53.55	-23.79
LUFTHANSA	-49.65	-22.72	-49.66	-22.72
MAN	-50.13	-23.07	-50.12	-23.04
MUENCHENER RUECK	-46.11	-24.08	-46.12	-24.09
RWE	-52.14	-22.33	-52.15	-22.32
SAP	-48.26	-22.93	-48.27	-22.94
SIEMENS	-47.26	-23.44	-47.26	-23.42
THYSSENKRUPP	-48.66	-21.76	-48.67	-21.76
VOLKSWAGEN	-46.68	-23.43	-46.68	-23.41

Table 3.5: ADF test statistics for returns in the FSE from 1 Jan 1998 to 31 Dec 2007

If we specify the following hypotheses:

$$H_0 : \textit{unit root}\,(\alpha = 1) \text{ vs. } H_1 : \textit{stationarity},$$

we know that the corresponding test statistic (critical value -1.95 for significance level 5%) is obtained in the following way (according to [1])

$$\hat{t}_n = \frac{1-\hat{\alpha}}{\sqrt{\hat{\sigma}^2\left(\sum_{t=2}^{n} r_{t-1}^2\right)^{-1}}}. \tag{3.7}$$

p	Without time trend 0	4	With time trend 0	4
FTSE 100	-52.19	-25.52	-52.19	-25.53
ASTRAZENECA	-48.22	-24.57	-48.23	-24.58
BARCLAYS	-47.28	-26.20	-47.29	-26.21
BG GROUP	-55.87	-25.13	-55.84	-25.03
BHP BILLITON	-50.30	-25.34	-50.27	-25.25
BP	-51.77	-25.16	-51.78	-25.16
BRITISH AMERICAN TOBACCO	-51.75	-25.04	-51.73	-24.97
BT GROUP	-49.75	-25.71	-49.76	-25.71
DIAGEO	-51.20	-25.83	-51.20	-25.82
GLAXO SMITH KLINE	-51.09	-25.05	-51.10	-25.06
HBOS	-50.20	-25.24	-50.21	-25.24
HSBC	-50.27	-25.46	-50.27	-25.45
LLOYDS TSB GROUP	-49.67	-25.46	-49.68	-25.46
NATIONAL GRID	-53.13	-25.43	-53.12	-25.39
RIO TINTO	-49.12	-24.66	-49.10	-24.59
ROYAL BANK OF SCOT. GROUP	-48.89	-25.37	-48.89	-25.36
ROYAL DUTCH SHELL A	-50.94	-23.99	-50.95	-23.99
STANDARD CHARTERED	-49.60	-24.76	-49.60	-24.74
TESCO	-55.86	-26.26	-55.86	-26.22
UNILEVER	-49.84	-23.88	-49.85	-23.88
VODAFONE GROUP	-51.27	-24.62	-51.28	-24.62

Table 3.6: ADF test statistics for returns in the LSE from 1 Jan 1998 to 31 Dec 2007

Under the null hypothesis, the test statistic is not normally distributed, but the critical values are provided in tables. In our case, the empirical range of the test statistic was $[-55.87, -46.11]$ and $[-26.26, -21.76]$, for cases $p = 0$ and $p = 4$, respectively. One can conclude that all return processes are stationary at significance level of 5%, for the cases $p = 0$ and $p = 4$.

ADF tests for returns with trend

If we insert a trend component in the regression model considered above, we end with the following regression model:

T	Without time trend 8	Without time trend 12	With time trend 8	With time trend 12
DAX	0.22	0.16	0.22	0.16
ADIDAS	0.31	0.04	0.32	0.04
ALLIANZ	0.17	0.16	0.17	0.16
BASF	0.11	0.05	0.12	0.05
BAYER	0.26	0.11	0.25	0.11
BMW	0.04	0.03	0.05	0.03
COMMERZBANK	0.20	0.09	0.20	0.09
CONTINENTAL	0.28	0.11	0.29	0.12
DEUTSCHE BANK	0.05	0.05	0.06	0.05
DEUTSCHE TELEKOM	0.20	0.19	0.19	0.18
EON	0.43	0.03	0.47	0.04
HENKEL	0.08	0.04	0.09	0.05
LINDE	0.22	0.04	0.22	0.04
LUFTHANSA	0.09	0.07	0.10	0.07
MAN	0.36	0.07	0.37	0.07
MUENCHENER RUECK	0.15	0.15	0.16	0.16
RWE	0.38	0.05	0.42	0.06
SAP	0.04	0.04	0.04	0.04
SIEMENS	0.08	0.08	0.09	0.08
THYSSENKRUPP	0.22	0.05	0.22	0.05
VOLKSWAGEN	0.38	0.10	0.37	0.10

Table 3.7: KPSS test statistics for returns in the FSE from 1 Jan 1998 to 31 Dec 2007

$$\Delta r_t = c + \mu t + (\alpha - 1)\, r_{t-1} + \sum_{i=1}^{p} \Delta r_{t-i} + \epsilon_t. \tag{3.8}$$

The stated hypotheses and the test statistic are same as above. Now, the critical value changes to -3.41 for significance level 5%

In this case, the empirical range of the test statistic was $[-55.86, -46.12]$ and $[-26.22, -21.76]$, for cases $p = 0$ and $p = 4$, respectively. One can conclude that all return processes are trend stationary at significance level of 5%, for the case of $p = 0$ and $p = 4$.

T	Without time trend 8	12	With time trend 8	12
FTSE 100	0.16	0.12	0.17	0.13
ASTRAZENECA	0.09	0.04	0.10	0.04
BARCLAYS	0.04	0.03	0.04	0.04
BG GROUP	0.28	0.07	0.30	0.08
BHP BILLITON	0.05	0.03	0.05	0.03
BP	0.06	0.05	0.06	0.06
BRITISH AMERICAN TOBACCO	0.06	0.04	0.06	0.04
BT GROUP	0.21	0.22	0.22	0.23
DIAGEO	0.04	0.03	0.04	0.03
GLAXO SMITH KLINE	0.07	0.06	0.08	0.07
HBOS	0.04	0.04	0.04	0.04
HSBC	0.07	0.03	0.07	0.03
LLOYDS TSB GROUP	0.04	0.04	0.05	0.04
NATIONAL GRID	0.13	0.14	0.14	0.14
RIO TINTO	0.15	0.04	0.16	0.05
ROYAL BANK OF SCOT. GROUP	0.20	0.03	0.20	0.03
ROYAL DUTCH SHELL A	0.07	0.06	0.08	0.07
STANDARD CHARTERED	0.05	0.03	0.05	0.03
TESCO	0.07	0.04	0.08	0.04
UNILEVER	0.09	0.03	0.10	0.04
VODAFONE GROUP	0.25	0.22	0.25	0.23

Table 3.8: KPSS test statistics for returns in the LSE from 1 Jan 1998 to 31 Dec 2007

KPSS tests for returns without trend

For the KPSS test the following regression model is used,

$$\Delta r_t = c + k\sum_{i=1}^{t} \xi_i + \eta_t, \tag{3.9}$$

where η_t is stationary and ξ_t is independent and identically distributed standardized random variable, see [1].

The hypotheses and the test statisic (critical value 0.46 for significance level

5%) of the KPSS test are given as follows:

$$H_0 : \textit{stationarity}\,(k = 0) \text{ vs. } H_1 : \textit{unit root}$$

$$KPSS_T = \frac{\sum_{t=1}^{n} \left(\sum_{i=1}^{t} \hat{\eta}_i \right)^2}{n^2 \hat{\omega}_T^2}, \tag{3.10}$$

where $\hat{\omega}_T^2$ is an estimator of the spectral density at a frequency of zero, as stated in [1].

Under the null hypothesis, the test statistic is not normally distributed, but the critical values are given in tables.

In our case, the empirical range of the test statistic was $[0.04, 0.43]$ and $[0.03, 0.22]$, for cases $T = 8$ and $T = 12$, respectively. One can conclude that all return processes showed a stationary pattern at significance level of 5%, for the cases $T = 8$ and $T = 12$.

KPSS tests for returns with trend

Similarly as before, we shall include a trend component in the regression model,

$$\Delta r_t = c + \mu t + k \sum_{i=1}^{t} \xi_i + \eta_t. \tag{3.11}$$

The critical value is now 0.15 for significance level 5%. In the case of German and British stock market the empirical range of the test statistic for returns was $[0.04, 0.47]$ and $[0.03, 0.23]$, for cases $T = 8$ and $T = 12$, respectively. Majority of the return processes are stationary at significance level of 5%, in both cases ($T = 8$ and $T = 12$).

3.2.2 Heteroscedasticity

As a matter of fact, the empirical return series exhibit heteroscedasticity. In Figure 3.2, the estimated standard deviations of index returns were plotted against time. This standard deviations were estimated for 100 consecutive days in the period under review.

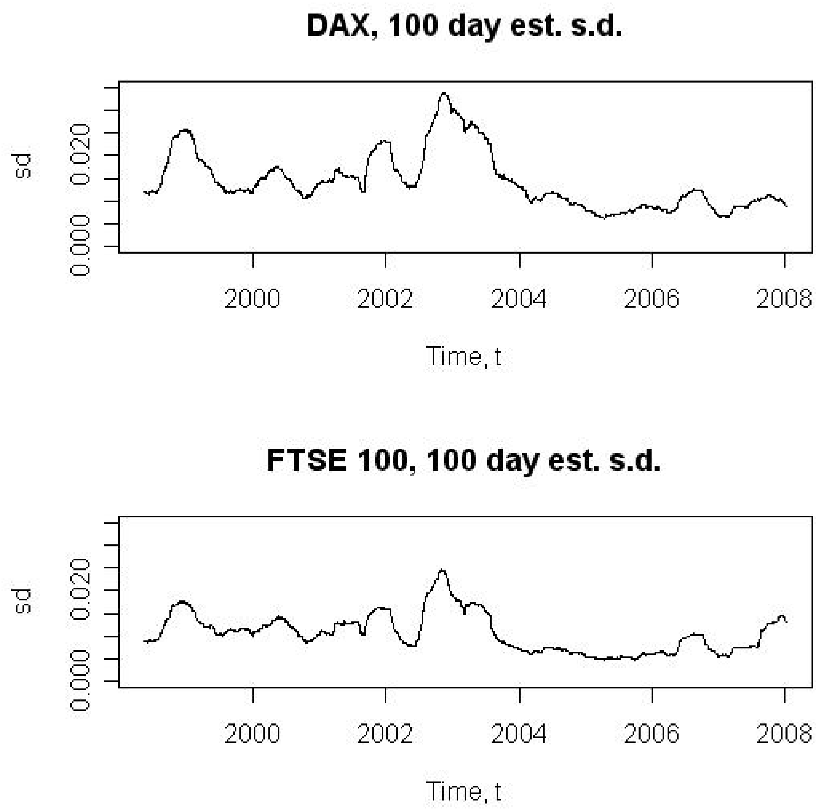

Figure 3.2: DAX and FTSE 100 estimated standard deviations of returns (mooving window 100 days) from 1 Jan 1998 to 31 Dec 2007

As one can observe, the volatility of the return processes is not constant in the analyzed period, as claimed above. Moreover, the volatility of DAX index returns appear to be higher than the volatility of FTSE 100 index returns. Under these circumstances, a portfolio constructed of blue chip stocks on the FSE bears a higher risk than a portfolio of blue chips on the LSE.

To additionally support this claim, boxplots of returns were shown in Figure 3.3. It is obvious that the range and interquartile range of returns on DAX is higher than for FTSE 100. Thus, in the analyzed period of ten years, it holds that the volatility is much higher for the blue chips on the FSE than the volatility of blue chips on the LSE.

3.2.3 Asymmetry

The empirical distribution of returns needs to be tested for symmetry. As a motivation for the statistical test, let us consider boxplots of daily returns on DAX and FTSE 100 index (Figure 3.3). It is difficult to distinguish whether the empirical distribution of returns for both indices is symmetric or not. Additionally, one can observe that there are many outliers in the tails of the empirical distributions. In this case it is recommended to construct a statistical test to formally prove the claims about the symmetry of distributions of returns.

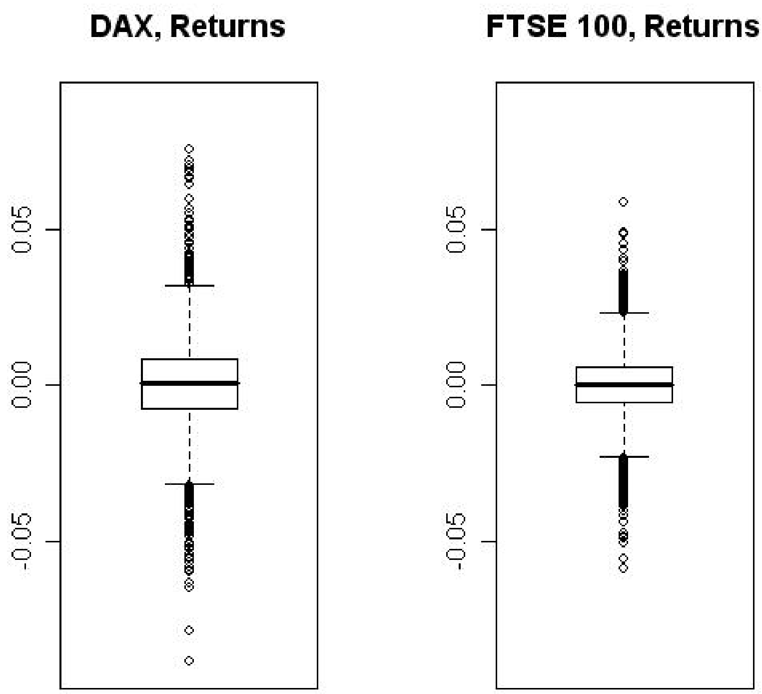

Figure 3.3: DAX and FTSE 100 index, boxplots of returns

The hypotheses for the test are as follows:

$$H_0 : symmetry\,(S = 0) \text{ vs. } H_1 : asymmetry.$$

Since it holds that

$$\sqrt{n}\hat{S} \xrightarrow{\ell} N(0,6), \tag{3.12}$$

according to [1], one can obtain the critical values for the two-sided test. In our case, the estimated skewness, $\hat{S}$, should be between -0.09 and 0.09 to accept the null hypothesis with 5% significance (symmetry of the return process). By considering the Tables 3.3 and 3.4 one can conclude that the empirical return distributions for both indices and 37 stocks are asymmetric at significance level of 5%. Only 3 stocks had an symmetric distribution of returns in the analyzed period at same significance level (MAN, HBOS and LLOYDS TSB GROUP).

3.2.4 Overkurtosis

To check for overkurtosis of the empirical return processes, histograms for returns on DAX and FTSE 100 are shown in Figure 3.4. The bandwidth was set to 0.01 which corresponds to 1% in the return scale. It holds that both distributions have more probability mass in the centre of the distribution and that they also have fat tails, which indicates the presence of outliers, i.e. extreme low and high returns.

The formal hypotheses are:

$$H_0 : K = 3 \text{ vs. } H_1 : overkurtosis.$$

The estimated kurtosis, $\hat{K}$, is distributed as follows (see [1]),

$$\sqrt{n}\left(\hat{K} - 3\right) \xrightarrow{\ell} N(0,24). \tag{3.13}$$

In our case, by considering the one-tailed test, the estimated excess kurtosis, $\left(\hat{K} - 3\right)$, should be lower than 0.16, and the estimated kurtosis, $\hat{K}$, lower than 3.16, to accept the null hypothesis with significance level of 5% (no overkurtosis). From the Tables 3.3 and 3.4 one can see that the lowest estimated kurtosis is by MAN (5.02) and the largest by British American Tobacco (30.47). Therefore, all empirical distributions of returns have kurtosis higher than 3 in the period under review, at significance level of 5%.

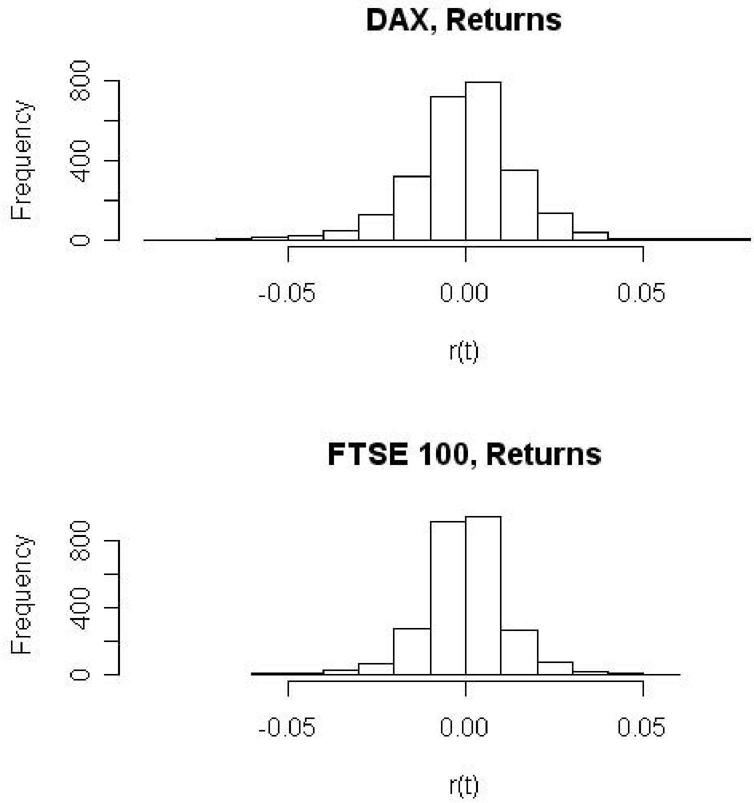

Figure 3.4: DAX and FTSE 100 index, histograms of returns

3.2.5 Non-normality

For high frequency data the normality assumption for distribution of returns is often violated. By considering our data, consisting from daily observations, one can clearly show the asserted claim. Indeed, by comparing the quantiles of the normal distribution and the quantiles of the empirical distribution of returns (Figure 3.5) one finds clear violations of the normality assumption.

The hypotheses are given as follows:

$$H_0 : r_t \sim N \text{ vs. } H_1 : \; non - normality.$$

The critical value for the Bera-Jarque test statistic is 5.99 for significance

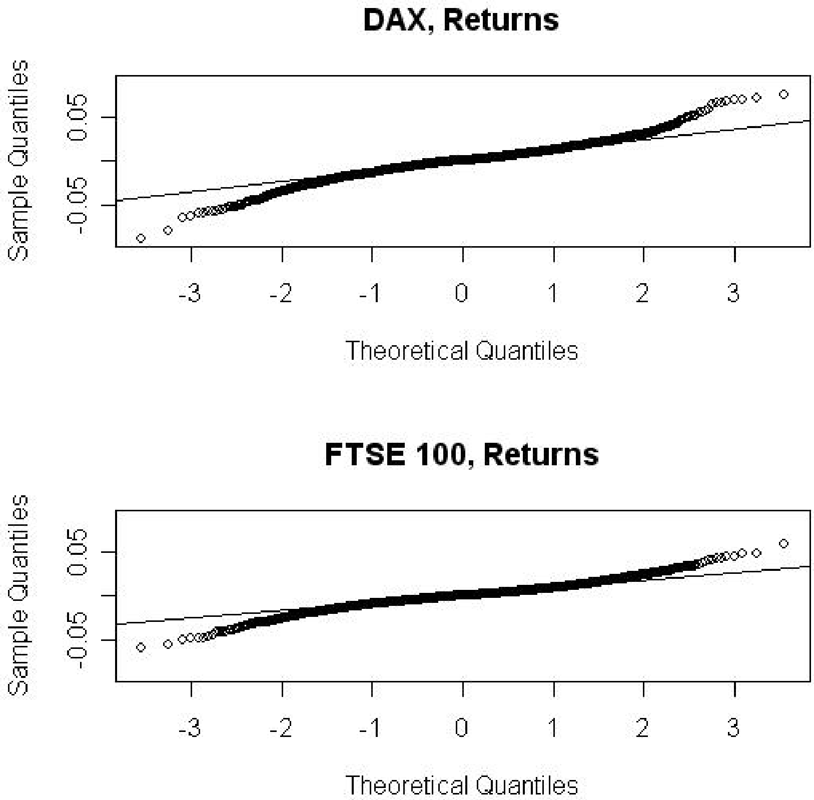

Figure 3.5: DAX and FTSE 100 index, Normal-QQ plots of returns

level of 5%. In our case, the lowest BJ test statistic was estimated by MAN (443), the largest by British American Tobacco (83374). Thus, all return distributions reject the normality hypothesis at significance level of 5%.

3.3 Linear Time Dependence of Return Processes

To empirically support the fact that return processes are serially uncorrelated or with some minor serial correlation, and that some transformated processes of the original return processes are serially correlated, the empirical autocorrelation functions as well as the portmanteau tests were introduced and explained in the sequel.

3.3.1 Empirical Autocorrelation Functions

An estimator for the autocorrelation function, given in [1], at lag τ for individual index or stock, i, is given by

$$\hat{\rho}_{\tau,i} = \frac{\hat{\gamma}_{\tau,i}}{\hat{\gamma}_{0,i}}, \tag{3.14}$$

where $\hat{\gamma}_{\tau,i}$ denotes the estimator of the autocovariance function at lag τ for individual index or stock i.

The empirical autocorrelation functions for the return processes and their transformations of DAX and FTSE 100 indices are given in Figure 3.6. The plotted graphs represent the empirical autocorrelation functions for the plain, squared and absolute return processes of DAX and FTSE 100 index, for different lags τ.

Under the null hypothesis $\sqrt{n}\hat{\rho}_{\tau,i}$ is asymptotically standard normally distributed, see [1]. Therefore, on each graph there are two dashed lines showing whether an estimated autocorrelation coefficient is statistically significant or not.

The left pair of graphs shows the empirical autocorrelation of the plain returns. One can observe that the returns on DAX are stationary, i.e. there is no serial correlation among consecutive returns. At the same time, the return process of FTSE 100 is not stationary, since there exist autocorrelation coefficients that are statistically different from zero, in particular on lags two, three and five.

For the empirical autorcorrelation functions for the squared and absolute returns there exist strong serial dependence. This clearly shows that there exists volatility clustering. This means that on days with low volatility one expect lower volatility on upcoming days. It also holds that periods of high volatitility are followed by periods with higher volatility.

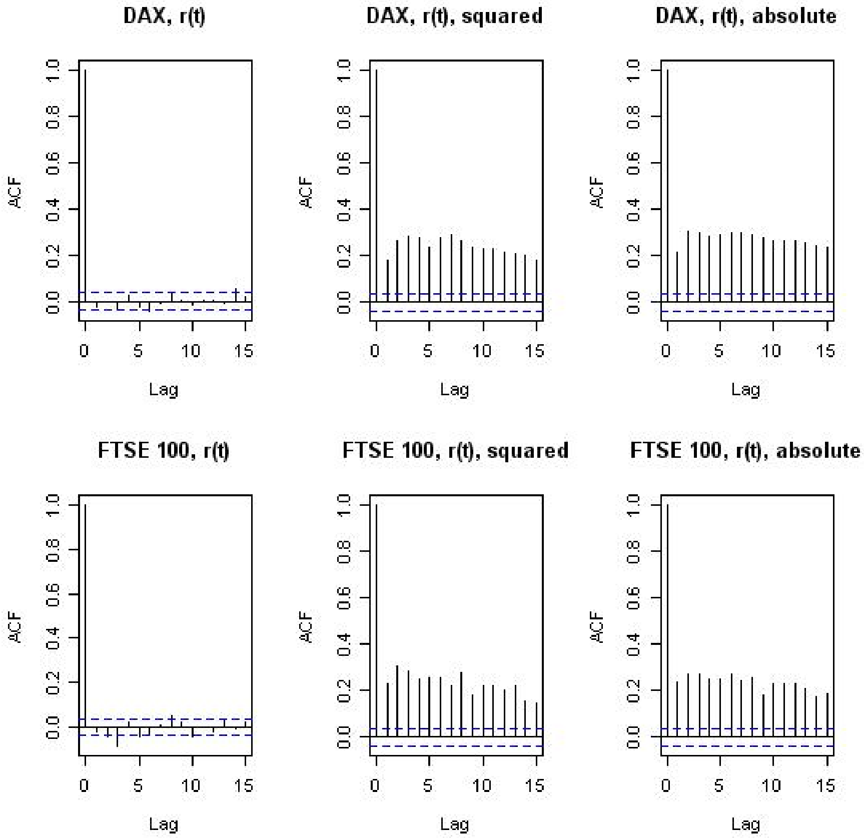

Figure 3.6: Empirical autocorrelation functions of DAX and FTSE 100 plain, squared and absolute returns from 1 Jan 1998 to 31 Dec 2007

3.3.2 Portmanteau Tests

To statistically prove whether the original or transformed return processes exhibit serial autocorrelation, portmanteau tests were constructed. For analyzed indices and stocks the Ljung-Box test statistics are given in tables 3.9 and 3.10 for FSE and LSE, respectively.

The hypotheses for testing for absence of autocorrelation are given as follows:

$$H_0 : \rho_1 = \ldots \rho_m = 0 \text{ vs. } H_1 : \exists \rho_i \neq 0.$$

Basically, there exist two important test statistics, namely the Box-Pierce

	$Q_8^*(r_i)$	$Q_8^*(r_i^2)$	$Q_8^*(\|r_i\|)$
DAX	15.1	1445.0	1728.8
ADIDAS	32.3	476.4	569.0
ALLIANZ	28.2	678.4	1361.0
BASF	17.3	637.7	582.0
BAYER	7.4	37.5	412.8
BMW	22.0	668.4	967.7
COMMERZBANK	15.4	731.0	1008.3
CONTINENTAL	14.7	311.3	432.0
DEUTSCHE BANK	26.7	842.4	1186.2
DEUTSCHE TELEKOM	19.7	836.9	1665.5
EON	35.4	404.0	696.4
HENKEL	12.2	577.2	710.3
LINDE	11.9	214.8	373.5
LUFTHANSA	10.1	447.0	685.6
MAN	7.5	340.2	392.8
MUENCHENER RUECK	57.0	783.1	1441.5
RWE	13.0	369.5	540.9
SAP	17.7	317.1	1141.2
SIEMENS	32.3	403.5	1003.2
THYSSENKRUPP	17.6	106.2	334.1
VOLKSWAGEN	49.8	544.3	557.1

Table 3.9: Ljung-Box test statistics for plain, squared and absolute returns for DAX index and selected stocks on FSE from 1 Jan 1998 to 31 Dec 2007, $m = 8$

and the Ljung-Box test statistics, for testing the stated hypotheses. Both test statistics are given in [1].

Box-Pierce test statistic (critical value 15.5 for significance level 5%) is given as:

$$Q_m = n\sum_{j=1}^{m} \hat{\rho}_j^2 \xrightarrow{\ell} \chi_m^2. \tag{3.15}$$

Ljung-Box test statistic (critical value 15.5 for significance level 5%) counts additionally for the bias of the estimator $\hat{\rho}_\tau$. It is given by

	$Q_8^*(r_i)$	$Q_8^*(r_i^2)$	$Q_8^*(\|r_i\|)$
FTSE 100	41.9	1426.3	1393.7
ASTRAZENECA	28.8	246.6	493.9
BARCLAYS	60.6	1050.1	1266.3
BG GROUP	49.7	340.8	475.1
BHP BILLITON	28.5	239.0	331.4
BP	33.3	448.5	471.4
BRITISH AMERICAN TOBACCO	32.7	32.0	834.8
BT GROUP	40.6	341.0	944.2
DIAGEO	37.6	580.2	1013.8
GLAXO SMITH KLINE	25.3	67.2	459.2
HBOS	34.1	734.9	1054.5
HSBC	24.9	429.1	1185.3
LLOYDS TSB GROUP	33.6	711.0	1288.8
NATIONAL GRID	18.2	117.5	472.2
RIO TINTO	25.9	103.3	198.7
ROYAL BANK OF SCOT. GROUP	30.4	749.8	1532.1
ROYAL DUTCH SHELL A	15.9	811.0	963.8
STANDARD CHARTERED	17.0	339.1	941.4
TESCO	46.0	438.9	652.2
UNILEVER	8.2	161.8	547.6
VODAFONE GROUP	28.8	444.8	949.0

Table 3.10: Ljung-Box test statistics for plain, squared and absolute returns for FTSE 100 index and selected stocks on LSE from 1 Jan 1998 to 31 Dec 2007, $m = 8$

$$Q_m^* = n\,(n+2)\sum_{j=1}^{m}\frac{1}{n-j}\hat{\rho}_j^2 \xrightarrow{\ell} \chi_m^2. \tag{3.16}$$

Since both test statistics perform equally well in large samples, only the Ljung-Box test statistic was used in our case. Actually, both test statistics in our empirical investigation resulted with almost equal values. Since it is advisible to take the number of lags m close to $\log(n)$, according to [2], we choose 8 autocorrelation coefficients in our analysis.

The returns on DAX index show no serial dependence at significance level of 5%. This is not suprisingly, since the lower order empirical autocorrelation coefficients showed already that they are not individually statistically different from 0. On the other hand, the returns on FTSE 100 index are serially autorcorrelated at same significance level.

For selected stocks, 9 return processes show linear time independency (only one on LSE, namely UNILEVER), whereas 31 return processes rejected the time independency hypothesis at significance level of 5%.

All squared and absolute return processes clearly show linear time dependence at significance level of 5%. Thus the stylized fact that return processes are dependent white noise processes was clearly proven for both markets, FSE and LSE, in the analyzed period of ten years. Therefore, ARCH and GARCH models should be used for modelling of the volatility processes.

Chapter 4

Stock Market Efficiency

Efficiency of financial markets, in particular stock markets, plays an important role when analyzing possible existence of arbitrage opportunities or market imperfections. Therefore, one should test whether a market supports a given efficiency market hypothesis, i.e. it should be tested whether stock prices walk in a unpredictable (random) manner.

This chapter deals with the analysis of stock market efficiency in the weak form - the analysis and testing procedures refer only to historical prices. By assumption, this historical prices are included in every private information set and they reflect all information at some given point on a particular stock market. Therefore, the information set in time t is given by the sequence of stock prices $F_t = (\{S_t\})$.

For a suitable model for testing the weak form efficiency hypothesis, the random walk with a drift for logged stock prices and logged index values was used. It relies on the assumption that the stock price processes are governed by geometric Brownian motion (W_t represents the Wiener process) in the following way,

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \tag{4.1}$$

according to [1]. In our case, after applying Ito's Lemma for logged stock prices and logged index values, and after discretization we finally get our desired model, given in [1],

$$\log S_{t\Delta} = \log S_{(t-1)\Delta} + \Delta\left(\frac{\mu - \sigma^2}{2}\right) + \sqrt{\Delta}\sigma\xi_t. \quad (4.2)$$

In the sequel, this model will be used within the ADF and KPSS test procedures, which were already explained before. The difference is now, that the logged index values and logged stock prices are used instead of the return processes itself. After applying the ADF and KPSS test procedures to logged index values and logged stock prices, one gets the ADF test statistics summarized in Tables 4.1 and 4.2 for the FSE and LSE, respectively, and the KPSS test statistics in Tables 4.3 and 4.4 for FSE and LSE, respectively.

	Without time trend		With time trend	
p	0	4	0	4
DAX	-1.03	-0.95	0.78	0.77
ADIDAS	-0.74	-0.76	0.43	0.44
ALLIANZ	-1.13	-1.22	-0.41	-0.48
BASF	-0.73	-0.65	1.25	1.24
BAYER	-0.94	-0.91	0.52	0.48
BMW	-3.28	-3.19	0.46	0.43
COMMERZBANK	-1.44	-1.56	-0.49	-0.51
CONTINENTAL	-0.11	-0.09	1.39	1.32
DEUTSCHE BANK	-2.03	-2.06	0.20	0.22
DEUTSCHE TELEKOM	-1.14	-1.05	-0.31	-0.27
EON	-0.24	-0.07	0.90	1.00
HENKEL	-1.73	-1.58	0.65	0.62
LINDE	-1.01	-0.79	0.57	0.63
LUFTHANSA	-1.73	-1.83	-0.15	-0.19
MAN	0.17	0.09	1.29	1.19
MUENCHENER RUECK	-1.23	-1.35	-0.29	-0.29
RWE	-0.43	-0.48	0.69	0.63
SAP	-3.09	-3.11	0.05	0.05
SIEMENS	-1.76	-1.73	0.80	0.75
THYSSENKRUPP	-0.83	-1.01	0.50	0.41
VOLKSWAGEN	-0.01	-0.10	1.04	0.96

Table 4.1: ADF test statistics for logged index values and stock prices in the FSE from 1 Jan 1998 to 31 Dec 2007

	Without time trend		With time trend	
p	0	4	0	4
FTSE 100	-1.60	-1.34	0.36	0.39
ASTRAZENECA	-3.17	-2.98	-0.02	-0.05
BARCLAYS	-2.71	-2.50	0.11	0.09
BG GROUP	0.12	0.85	1.89	2.42
BHP BILLITON	-0.44	-0.31	1.86	2.06
BP	-3.02	-2.77	0.42	0.53
BRITISH AMERICAN TOBACCO	-0.70	-0.38	1.66	1.97
BT GROUP	-1.17	-1.05	-0.27	-0.30
DIAGEO	-2.12	-1.64	0.68	0.82
GLAXO SMITH KLINE	-2.32	-2.11	-0.19	-0.27
HBOS	-3.00	-2.65	-0.11	-0.12
HSBC	-2.75	-2.75	0.46	0.56
LLOYDS TSB GROUP	-2.14	-1.96	-0.57	-0.62
NATIONAL GRID	-1.96	-1.50	1.29	1.38
RIO TINTO	0.01	0.03	1.76	1.81
ROYAL BANK OF SCOT. GROUP	-2.85	-2.65	0.45	0.43
ROYAL DUTCH SHELL A	-2.27	-2.14	0.03	0.09
STANDARD CHARTERED	-1.65	-1.50	0.81	0.92
TESCO	-1.04	-0.53	1.16	1.48
UNILEVER	-2.39	-2.36	0.52	0.53
VODAFONE GROUP	-2.28	-2.06	0.45	0.55

Table 4.2: ADF test statistics for logged index values and stock prices in the LSE from 1 Jan 1998 to 31 Dec 2007

ADF test for logged index values and logged stock prices without time trend

Consider the following regression model:

$$\Delta \log S_t = c + (\alpha - 1) \log S_{t-1} + \sum_{i=1}^{p} \Delta \log S_{t-i} + \epsilon_t \tag{4.3}$$

In the case of FSE and LSE, the empirical range of the test statistic was $[-3.28, 0.17]$ and $[-3.19, 0.85]$, for cases $p = 0$ and $p = 4$, respectively. Since the theoretical critical value is -1.95 for significance level 5%, one can con-

T	Without time trend 8	12	With time trend 8	12
DAX	2.10	2.16	1.44	1.47
ADIDAS	7.16	2.48	4.89	1.70
ALLIANZ	7.50	1.84	5.11	1.25
BASF	8.90	2.16	6.10	1.49
BAYER	2.34	2.25	1.60	1.54
BMW	6.33	0.60	4.39	0.42
COMMERZBANK	3.08	2.37	2.10	1.62
CONTINENTAL	10.23	3.02	6.96	2.06
DEUTSCHE BANK	2.78	1.30	1.91	0.89
DEUTSCHE TELEKOM	7.78	1.18	5.30	0.81
EON	7.75	2.58	5.30	1.77
HENKEL	6.04	2.42	4.15	1.67
LINDE	4.78	2.57	3.27	1.76
LUFTHANSA	4.09	1.96	2.80	1.35
MAN	5.91	2.85	4.03	1.94
MUENCHENER RUECK	7.48	1.44	5.09	0.98
RWE	5.03	2.65	3.43	1.81
SAP	0.94	0.82	0.65	0.57
SIEMENS	2.30	1.04	1.58	0.71
THYSSENKRUPP	3.32	2.55	2.27	1.74
VOLKSWAGEN	2.48	2.37	1.70	1.63

Table 4.3: KPSS test statistics for logged index values and stock prices in the FSE from 1 Jan 1998 to 31 Dec 2007

clude that logged values of both indices and 28 stocks followed a random walk process, whereas the remaining 14 logged stock price processes did not follow a random walk process, for cases $p = 0$ and $p = 4$. The majority of logged stock price processes that rejected the hypothesis for an unit root were on the LSE.

ADF test for logged index values and logged stock prices with time trend

If we insert a trend component in the regression model considered above, we end with the following regression model:

T	Without time trend 8	12	With time trend 8	12
FTSE 100	2.83	2.54	1.93	1.74
ASTRAZENECA	0.70	0.72	0.48	0.50
BARCLAYS	7.87	0.48	5.41	0.34
BG GROUP	11.11	2.82	7.59	1.93
BHP BILLITON	12.76	0.96	8.70	0.67
BP	2.12	1.07	1.46	0.74
BRITISH AMERICAN TOBACCO	12.35	1.19	8.43	0.82
BT GROUP	8.49	2.40	5.78	1.64
DIAGEO	8.46	0.92	5.80	0.64
GLAXO SMITH KLINE	7.72	1.45	5.28	1.00
HBOS	5.21	1.38	3.59	0.96
HSBC	6.34	0.85	4.37	0.59
LLOYDS TSB GROUP	8.47	1.79	5.80	1.23
NATIONAL GRID	5.60	1.86	3.85	1.28
RIO TINTO	11.35	1.76	7.77	1.22
ROYAL BANK OF SCOT. GROUP	8.65	1.41	5.93	0.98
ROYAL DUTCH SHELL A	1.72	1.33	1.18	0.91
STANDARD CHARTERED	7.99	1.80	5.50	1.25
TESCO	10.89	1.34	7.44	0.92
UNILEVER	2.35	0.93	1.64	0.65
VODAFONE GROUP	4.23	1.16	2.89	0.79

Table 4.4: KPSS test statistics for logged index values and stock prices in the LSE from 1 Jan 1998 to 31 Dec 2007

$$\Delta \log S_t = c + \mu t + (\alpha - 1) \log S_{t-1} + \sum_{i=1}^{p} \Delta \log S_{t-i} + \epsilon_t \qquad (4.4)$$

For this model, the theoretical critical value is -3.41 for significance level 5%. In our case, the empirical range of the test statistic was $[-0.57, 1.89]$ and $[-0.62, 2.42]$, for cases $p = 0$ and $p = 4$, respectively. It holds that all logged index values and logged stock prices are not trend stationary at significance level of 5% for the case of $p = 0$ and $p = 4$, and thus all processes are governed by a random walk with trend. In other words, all analyzed processes walked in an unpredictable manner.

KPSS tests for logged index values and logged stock prices without trend

For the KPSS test the following regression model is used

$$\Delta \log S_t = c + k \sum_{i=1}^{t} \xi_i + \eta_t. \tag{4.5}$$

The theoretical critical value for this test is 0.46 for significance level 5%. In our case, the empirical range of the test statistic was $[0.70, 12.76]$ and $[0.60, 3.02]$, for cases $p = 0$ and $p = 4$, respectively. One can conclude that all analyzed processes do not show a stationary pattern at significance level of 5% for the cases $T = 8$ and $T = 12$. This means that all processes walk in an unpredictable manner.

KPSS tests for logged index values and logged stock prices with trend

Similarly as before, we shall include a trend component in the regression model in the following way

$$\Delta \log S_t = c + \mu t + k \sum_{i=1}^{t} \xi_i + \eta_t. \tag{4.6}$$

The critical value is now 0.15 for significance level 5%. In the case of German and British stock market the empirical range of the test statistic for returns was $[0.48, 8.70]$ and $[0.42, 2.06]$, for cases $T = 8$ and $T = 12$, respectively. As before, one can conclude that all analyzed processes are not trend stationary at significance level of 5% for the cases $T = 8$ and $T = 12$, and thus are governed by a random walk process.

Since the ADF and KPSS test procedures showed that it is reasonable to accept the unit root hypothesis, one can conclude that the logged index values and the logged stock prices follow a random walk. This means that all index values and stock prices walk in an unpredictable manner in the period under review. Thus, both analyzed stock markets were weak form efficient. Moreover, under the weak form efficiency hypothesis, there does not exist arbitrage opportunities or market imperfections.

Chapter 5

Behaviour of Extreme Returns

For a comprehensive statistical analysis of risk for selected indices and stocks, an important feature of return processes was examined, namely the existence of moments of their distributions.

5.1 Tail Index Regression

To model the thickness of the tails of the return distribution we can assume that the ordered and largest returns follow a Pareto distribution

$$P(r) = P(r_t < r) \sim kr^{-\alpha}. \tag{5.1}$$

When $\alpha > c$ it holds that $E[|r_t|^c] < \infty$, see [1]. This means that the $c-th$ moment of the return distribution is finite. Consequently, we can use the Pareto distribution of the largest returns to check the existence of moments of the return distributions.

Furthermore, the log function $P(r)$ is linear for large and ordered returns, see [1]:

$$\log P(r) \approx \log k - \alpha \log r. \tag{5.2}$$

Under these circumstances, the tail index regression has the following form:

$$\log\frac{i}{n} \approx \log k - \alpha \log r_{(i)}, \tag{5.3}$$

where $r_{(i)}$ represent the ordered returns.

The graphical illustration of the fitted tail index regression line for both market indices, DAX and FTSE 100, is given in Figure 5.1. In the regression model only the largest ten returns per index were used.

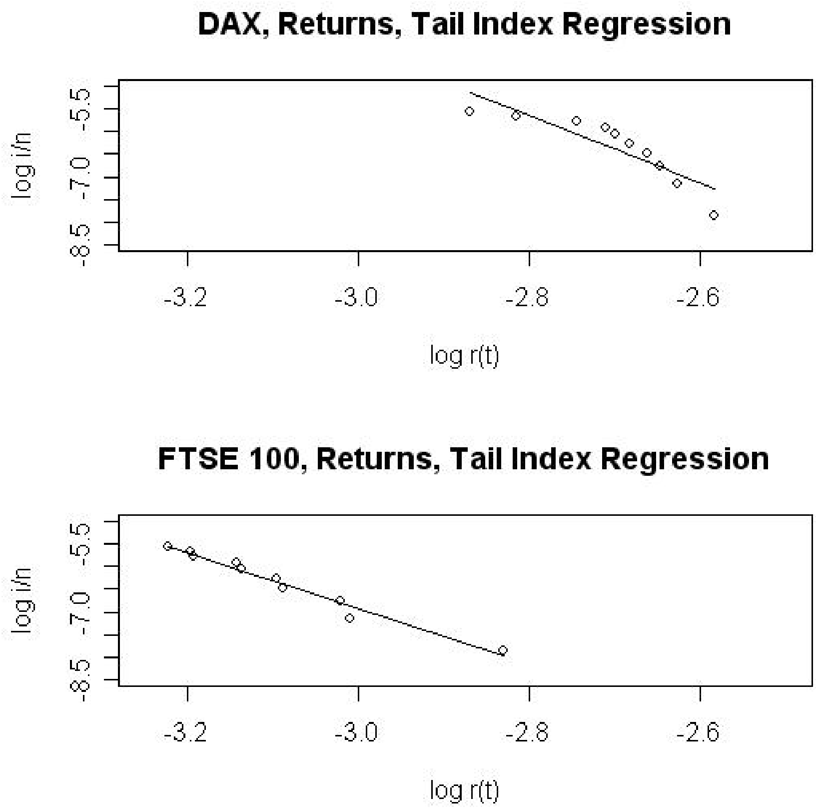

Figure 5.1: The right side of the logged empirical distribution of the DAX and FTSE 100 returns from 1 Jan 1998 to 31 Dec 2007

The slope of the tail index regression for DAX is 7.46, and 6.17 for FTSE 100 index. This indicates that for both index the first four moments of the return distribution are finite.

5.2 Least Square and Hill Estimators

Two estimators of the tail index α of a particular return distribution were analyzed here, namely the Least Square and the Hill estimator. The Tables 5.1 and 5.2 contain estimated values of the tail index α for return distributions with m observations used for the estimation, for FSE and LSE, respectively. According to [1], the ratio m/n should be around 0.5% or 1%. Therefore, two values, 10 and 25, are suitable for the selection of m in our particular case.

	LS		Hill	
m	10	25	10	25
DAX	7.46	4.34	5.44	3.49
ADIDAS	4.22	4.79	5.51	5.24
ALLIANZ	4.09	3.35	3.50	3.29
BASF	3.17	3.79	4.13	5.11
BAYER	1.37	2.08	3.05	3.77
BMW	4.75	4.61	5.12	4.67
COMMERZBANK	2.84	3.76	5.35	4.23
CONTINENTAL	4.66	3.89	6.85	3.71
DEUTSCHE BANK	4.17	4.14	7.30	4.05
DEUTSCHE TELEKOM	4.12	4.88	6.43	5.71
EON	4.98	4.58	5.27	4.23
HENKEL	3.80	4.24	4.13	4.42
LINDE	4.25	4.75	5.79	4.87
LUFTHANSA	2.99	3.71	4.27	3.96
MAN	4.43	5.19	5.57	6.45
MUENCHENER RUECK	3.45	3.19	3.85	2.84
RWE	2.75	3.21	3.79	4.06
SAP	2.54	3.08	3.36	3.85
SIEMENS	3.16	3.82	4.02	4.86
THYSSENKRUPP	4.84	5.75	6.42	7.03
VOLKSWAGEN	7.89	4.81	7.41	4.60

Table 5.1: Least Square (LS) and Hill estimators of the tail index α for returns on DAX index and selected stocks on FSE with m observations used for the estimation

	LS		Hill	
m	10	25	10	25
FTSE 100	6.17	4.74	6.93	5.19
ASTRAZENECA	3.90	3.43	4.12	3.70
BARCLAYS	7.13	6.14	6.13	6.09
BG GROUP	5.44	4.41	4.37	4.90
BHP BILLITON	3.41	3.40	3.08	3.84
BP	3.55	4.39	4.78	5.09
BRITISH AMERICAN TOBACCO	1.75	2.21	3.12	2.44
BT GROUP	6.19	4.62	6.45	3.48
DIAGEO	4.90	3.90	4.96	3.66
GLAXO SMITH KLINE	2.07	2.60	3.04	3.52
HBOS	2.26	2.81	3.70	3.61
HSBC	4.30	4.39	5.86	3.89
LLOYDS TSB GROUP	4.99	4.42	7.35	3.36
NATIONAL GRID	2.73	3.11	3.49	3.79
RIO TINTO	2.46	2.99	5.15	3.44
ROYAL BANK OF SCOT. GROUP	6.41	4.89	7.36	4.45
ROYAL DUTCH SHELL A	4.52	5.40	6.78	5.61
STANDARD CHARTERED	3.88	4.07	5.27	4.61
TESCO	4.38	4.45	5.22	4.30
UNILEVER	6.50	3.86	7.21	3.64
VODAFONE GROUP	4.76	5.00	5.96	5.54

Table 5.2: Least Square (LS) and Hill estimators of the tail index α for returns on FTSE 100 index and selected stocks on FSE with m observations used for the estimation

The Least Square estimator was already presented in the previous subsection. According to [1], Hill (1975) has suggested an estimator using the maximum likelihood method:

$$\hat{\alpha} = \left(\frac{1}{m-1} \sum_{i=1}^{m-1} \log r_{(i)} - \log r_{(m)} \right)^{-1}. \tag{5.4}$$

From the estimated values, one can conclude that all return distributions, except BAYER and BRITISH AMERICAN TOBACCO (case $m = 10$), have finite first and second moments. It is most likely that this exceptions are due to the small number of returns used for estimation.

Additionally, one can empirically find that for nine processes, according to the Least Square estimator, the third moment is not finite. When one considers only the Hill estimator, all processes have defined first three moments.

The knowledge about the existence of moments is important in further modelling, particularly in the modelling of the volatility process. For some models, it is strictly required that higher order moments are finite. Our return processes have finite second moments, so the volatility process can be modelled, among others, with ARCH and GARCH models.

Chapter 6

Conclusion

The analysis of return processes on German and British stock market in the period from 1998 to 2007 show that important statistical concepts play an important role in the empirical analysis of stock markets.

The distribution of returns was symmetric for most analyzed stocks, and leptokurtic for all return processes. The estimated kurtosis was significantly greater than 3 for all return distributions, which partially contributed to the rejection of the normality hypothesis for return distributions.

Majority of the return processes were stationary, but not independent processes, i.e. transformations of the processes show serial dependence. In the analyzed period, their volatility was time dependent. This fact motivates the implementation of ARCH and GARCH models when modelling the volatility process.

The analyzed stock markets prove the Efficiency Market Hypothesis. This investigation was based on the assumption that stock prices reflect all available information on some trading day. Additionally, it is assumed that all investors on a particular stock market have identical information sets, and that there does not exist arbitrage opportunities. Under this setup, one can show that analyzed stock prices and index values walk in an unpredictable manner.

Existence of moments of return distribution plays an important role in the analysis of extreme returns and also in model selection in further analysis. Our empirical analysis of extreme returns shows that the first two moments exist for all return processes. Additionally, when considering only the Hill estimator, one can conclude that the first three moments for all return distributions are finite.

To summarize, the empirical analysis supported the existence of stylized facts on the German and British stock market in the period under review. Both markets are efficient, and all analyzed return processes on these markets are suitable for further modelling.

Bibliography

[1] J. Franke, W. Härdle, and C. Hafner. *Statistics of Financial Markets - An Introduction.* Springer-Verlag, Berlin, 1st edition, 2004.

[2] R.S. Tsay. *Analysis of Financial Time Series.* John Wiley & Sons, Inc., New York, 1st edition, 2002.

Declaration of Authorship

I hereby certify that the thesis I am submitting is entirely my own original work except where otherwise indicated. I am aware of the University's regulations concerning plagiarism, including those regulations concerning disciplinary actions that may result from plagiarism. Any use of the works of any other author, in any form, is properly acknowledged at their point of use.

Student's signature: ____________________

Name (in capitals): Andrija Mihoci

Date of submission: 5 August 2008