Essays on Financial Markets

Byström, Hans

2000

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Essays on Financial Markets

Hans Byström

Lund Economic Studies number 91

Distributed by Department of Economics
Lund University
P.O. Box 7082
S-220 07 Lund
Sweden
Telephone +46-46-222 00 00
Fax +46-46-222 46 13

ISSN 0460-0029

Printed by KFS i Lund AB
Lund, Sweden, 2000

To my parents

Acknowledgments

This thesis would never have been concluded without the help and support of a number of people, and it is with great pleasure that I wish to thank the following persons.

First, and most importantly, I want to express my gratitude to my supervisor Björn Hansson. From the first day, he believed in me and my background in engineering, and in addition to supporting me on academic questions, he has also helped me through all the practical problems of graduate studies in general and the writing of a thesis in particular. Without his optimism, friendliness, and moral support, this thesis would have been much less of a joy to write.

I am also greatly indebted to my colleague and friend Henrik Amilon, with whom I have cowritten two of the essays in this thesis. Without Henrik's analytical mind and ability to focus on the core of the problem, the quality of these essays would certainly be much lower.

Special thanks go to all participants in the departmental Macro and Finance Seminars. Without their valuable comments, I would not have been as well prepared for the various international conferences where I have presented my work. Among all helpful participants at these conferences, I would like to mention Ray Chou in particular. I also want to express my gratitude to Johan Knif for insightful and detailed comments when acting as discussant on my licentiate dissertation and Johan Lyhagen for useful suggestions when acting as discussant on my final seminar. On mathematical and statistical issues, I have received help from several people, most importantly William A Brock, Zhuanxin Ding, James Dow, David Edgerton, Sören Johansen, Blake LeBaron, and Tobias Rydén.

In addition, I would like to acknowledge all my colleagues in Lund, especially Fredrik CA Andersson, Hossein Asgarian, HP Bermin, Henrik Degrér, Thomas Elger, Andreas Graflund, Therese Hindman, PM Lundgren, Birger Nilsson, Mattias Persson, and Dag Rydorff. I would also like to thank my former colleagues at the Department of Economics at the European University Institute in Florence, Italy.

Other people who deserve to be mentioned are Jörgen Norén and Ove Hagerbo for their help with computer issues, and Mona Hansson, Agneta Kjellgren, Agnetta Kretz, Ulla Olofsson, Jeanie Petersson, Ann-Charlotte Sahlin, and Berit Bäcklin Sjödahl for their help with administrative matters.

A special thanks goes to my family. I want to give my deepest gratitude to my parents by dedicating this book to them. They have always believed in me and whatever I have chosen to do. I also wish to thank my sister Anna for always providing insightful comments on things in general, and my brother Mats for his endless energy and his opinions on every possible issue. With all my heart, I also want to thank Camilla for her never ending support and encouragement and for always being there, as my dearest friend.

The very generous financial support from *Stiftelsen Bankforskningsinstitutet*, *Humanistisk Samhällsvetenskapliga Forskningsrådet*, and *Crafoordska Stiftelsen* is greatfully acknowledged.

Lund, June 2000

Hans Byström

Contents

Chapter 1

Introduction and Summary

1.1 Aim and Scope

This thesis is a collection of five independent, although not unrelated, essays dealing with several issues in empirical financial economics. Through the use of quantitative methods, the aim of the thesis is to study practically relevant problems related to the functioning of financial markets. The issues at hand stretch all the way from options pricing and futures hedging, over stochastic volatility and variance/covariance forecasting models, to issues on chaotic asset markets and market micro structures. The essays have at least three things in common, however; first, they all focus on time series properties of asset prices, second, they all explicitly model the return volatilities of these assets as time varying, and third, they all deal with nonlinear models.

1.2 Background and Contribution of the Thesis

Financial markets are not mere academic inventions or theoretical abstractions. They thrive even in the absence of academic research, and are of real and fundamental importance for the stability and growth of the economy. Throughout the last decades, much of the traditional national bank-based financial system has changed into a globally integrated market-based financial system, and in line with this development, financial markets are not only growing in size but also in complexity. While some of the new financial instruments and new ways of utilizing the markets originally sprung out of the academic world, the reverse has actually been more common; the academic researcher has had to move quickly in order to keep pace with fast moving market participants and their creativity in inventing new financial products. Still, academic research on financial markets is important in order to gain a deeper and more structured general understanding of the markets, and just as academic progress depends on the availability of practical cases to study, the development of the markets themselves does depend on qualified financial models and empirical analysis of market behavior.

Throughout the years, a multitude of different financial markets have been set up, and in this thesis, I focus on one of the oldest, the stock (or equity) market, as well as one of the youngest, the electric power market. When it comes to the stock market, my focus is on individual stocks and stock indices from the Nordic countries. In addition to the stocks and stock indices themselves, I also look at one of the larger derivative markets in the Nordic countries; the Swedish market for stock index options. In the case of electricity, I have chosen to study the first multinational electricity exchange in the world, the Nordic Power Exchange "Nord Pool". Both spot and futures contracts on electricity are traded in this market.

Research on financial markets deals with a number of different issues, from empirical studies on market behavior to the development of general theories on asset pricing, using tools from various academic subjects. In this thesis, however, I deal exclusively with the econometrics of financial markets. More specifically, I look at time series of asset prices and asset price changes (returns); in particular the variability and covariability of such time series. Asset returns are usually supposed to be random variables with associated statistical distributions, and through the use of both discrete time and continuous time models, estimation and forecasting of moments of these asset return distributions is the main thread throughout the thesis. The distinction between unconditional and conditional distributions is very important; unconditional distributions are often assumed to be constant over time while conditional distributions are usually modelled as time dependent. In this thesis, I capture the time variation with so called "ARCH models" and "stochastic volatility models". As an alternative to these stochastic models, deterministic "chaos models" are also applied; the idea is that not too complicated deterministic relationships exist between the returns of different days and that these relationships can be used to explain asset price movements and to improve forecasts of asset returns.

In financial markets, it can often be observed how large returns tend to be followed by large returns, of either sign, and small returns tend to be followed by small returns. To capture this "volatility clustering", Engle (1982) introduced the autoregressive conditional heteroscedastic (ARCH) model. ARCH and its extensions, most notably Bollerslev's (1986) generalized ARCH model (GARCH), have been very successful in modelling time varying variances in individual financial series. In multivariate extensions to the univariate ARCH and GARCH models, however, there are huge problems due to the very large number of parameters that must be estimated. These estimation problems limit the use of traditional multivariate ARCH and GARCH models. In Chapter 2 I therefore study a new multivariate technique, Orthogonal GARCH, of forecasting large covariance matrices in the presence of heteroscedasticity (Ding (1994), and Alexander and Chibumba (1998)). One of the assumptions behind Orthogonal GARCH is supposed to break down under stress scenarios and I therefore demonstrate the forecasting performance of the model on stock indicies from Denmark, Finland, Norway, and Sweden during the Asian Crisis 1997-1998.

Compared to the stock index prices studied in Chapter 2, prices in the electricity market are extremely volatile. In addition, since electricity cannot be stored, it is very difficult to price electricity futures by using arbitrage arguments. These facts, together with the fact that no other study on hedging with electricity futures exists in the literature, have led me to the study of the Nordic power market. In Chapter 3, I show that electricity spot as well as electricity futures returns are several times as volatile as ordinary financial asset returns. I also study whether electricity futures traded on Nord Pool can actually be used to hedge against adverse price changes in the spot electricity market.

Some studies have suggested that deterministic chaos models might be successfully applied to problems in economics and finance. This has lead many authors to claim to have found chaos in different asset markets, and that the movements of the assets are deterministic, not stochastic. In order to investigate whether Swedish stock returns can be described by chaotic models, we apply (in Chapter 4) a statistical test, the BDS test (Brock *et al.* (1996)), that detects deviations from the IID (Identical Independent Distributions) hypothesis. Intradaily, daily, and monthly stock index returns are all used.

If we plot these stock index returns in a scatterplot, with today's return on the x-axis, and yesterday's return on the y-axis, nothing unexpected appears. On the other hand, if we do the same thing with the returns of an individual stock, a very interesting observation can be made. The scatterplot of return pairs is now not only a diffuse cloud of dots, but a clearly defined "compass rose" pattern appears; evenly spaced rays of different thickness radiate from the origin in the directions of an ordinary compass (Crack and Ledoit (1996)). The reason why stocks have a compass rose pattern while stock indices have none is that prices of stocks and stock indices are quoted in different ways. While stock index prices are quoted with many decimals, stocks are usually quoted in integers or fractions of integers. This discreteness creates the observed compass rose. While the compass rose has been studied by other authors, we are the first to study whether the compass rose affects AR and GARCH parameter estimates (Chapter 5). We also discuss forecasts based on the compass rose more rigorously than other authors. Finally, the effect of the compass rose on the BDS test and Savit and Green's (1991) dependability index is studied.

While I have worked with discrete time models in the other chapters, I turn to continuous time models in Chapter 6. These models have turned out to be successful in the context of derivatives and, typically, the popular Black-Scholes framework is used in the pricing of derivatives. The Black-Scholes model to price options is based on a number of assumptions, the most critical one being the assumption of a constant return volatility. Stochastic volatility option pricing models relax this assumption by modelling the volatility as a stochastic process (Hull and White (1987)). In Chapter 6, I have chosen to apply a model based on Fourier-Inversion (Stein and Stein (1991)) to the Swedish stock index call option market. Compared to Stein and Stein (1991), I apply the Fourier-Inversion technique not only to the pricing of options, but also to the estimation of risk-neutral process parameters. To my knowledge, this study is also the first to apply stochastic volatility models to the Swedish options market.

1.3 Summary of the Thesis

In the first essay (Chapter 2), *Orthogonal GARCH and Covariance Matrix Forecasting in a Stress Scenario: The Nordic Stock Markets During the Asian Financial Crisis 1997-1998,* I study a new multivariate technique, Orthogonal GARCH, of forecasting large covariance matrices based on GARCH models. Orthogonal GARCH is built on principal components analysis and makes the creation of covariance matrices of arbitrary size based on any kind of ARCH or GARCH model possible. Using traditional multivariate GARCH models, covariance matrices of more than two to three assets are very difficult to estimate due to the large number of parameters. In Orthogonal GARCH, on the other hand, not only can any number of assets be modelled together in a common framework, but the forecasted covariance matrix also has the important property of always being positive definite. The most important drawback with Orthogonal GARCH is that it builds on assumptions that might sometimes break down; for instance when some of the assets we model behave differently than the other assets, or when the time period considered is very volatile. For that reason, I have chosen to apply the Orthogonal GARCH model to the highly volatile Nordic stock markets during the Asian Crisis 1997-1998, using a number of different forecast evaluation techniques. First, I apply two different statistical evaluation techniques, the traditional RMSE as well as an asymmetrical extension of this loss function. Second, I apply an operational technique suggested by the Basle Committee on Banking Supervision based on the concepts of Value-at-Risk and Risk Management. And third, considering the fact that a better covariance matrix forecast also gives a more accurate price of options on the assets in question, I create a simulated rainbow options market and test which forecasting technique earns most money when used to price and trade options. The results from the different evaluation methods indicate a better performance of the Orthogonal GARCH model compared to traditional unconditional forecasting techniques. The absolute performance of the Orthogonal GARCH model is weaker in the most volatile time periods but its relative performance is still very strong.

In the second essay (Chapter 3), *The Hedging Performance of Electricity Futures on the Nordic Power Exchange Nord Pool,* I study the first multinational power exchange in the world, "Nord Pool". Nord Pool has existed since 1996 and has participants from Norway, Sweden, Finland, Denmark, and England. Both spot and futures are traded on the exchange and in this essay, I investigate whether the futures contracts can be used to hedge short-term positions in the underlying spot market. Due to a number of factors, this question is of particular interest in the electricity market; first, since electricity cannot be stored, it is not obvious that futures prices and spot prices follow each other as well as in other markets where arbitrage is possible, second, the returns in the electricity markets have much higher volatility than in other financial markets, and third, the return distributions are far from being normally distributed. Minimum variance hedges are constructed with hedge ratios estimated in a number of different ways, and standard unconditional hedges like the naive hedge and the OLS hedge are compared to conditional GARCH hedges and moving average hedges in an out-of-sample fashion. The Orthogonal GARCH model described above is used in addition to a constant correlation bivariate GARCH model. The empirical results indicate some gains from hedging with futures, despite the lack of straightforward arbitrage possibilities in the electricity market. I also find the relative performance of the different variance minimizing hedges to depend on whether I evaluate by looking at unconditional or conditional variances. In the former case, the unconditional hedges are the better performers and in the latter case, the conditional hedges are the better performers.

In the third essay (Chapter 4), *The Search for Chaos and Nonlinearities in Swedish Stock Index Returns (*co-authored with Henrik Amilon*),* we search for evidence of chaos and other nonlinearities in Swedish stock return series. Empirical evidence suggests that nonlinear models, including chaotic models, might explain the dynamics of a financial return series. In our essay, we use the BDS test to determine which linear or nonlinear dependences are responsible for the observed rejection of the IID-hypothesis in the Swedish stock market. We look at monthly, daily and 15-minute return series and find clear evidence of nonlinearities in general but no evidence of chaos. Most of the nonlinearities seem to be due to ARCH effects, and a GARCH model with t-distributed errors explains most of the observed nonlinearities.

In the fourth essay (Chapter 5), *The Compass Rose Pattern of the Stock Market: How does it Affect Parameter Estimates, Forecasts, and Statistical Tests? (*co-authored with Henrik Amilon*),* we investigate the discrete nature of stock prices and how the minimum "tick size" on a stock exchange creates a "compass rose" pattern in a scatter plot of stock returns. The effect of the compass rose on GARCH estimates/forecasts as well as on tests for chaos is further studied. For a particular stock to demonstrate a compass rose pattern, its "typical" price changes must be comparable in size to the imposed minimum tick size, and in our Monte Carlo study, we simulate discrete AR-GARCH price series with a data generating process resembling a typical stock on the Stockholm Stock Exchange. We create discrete price series rounded both according to the rules of the Stockholm Stock Exchange and to the nearest integer. Simulations reveal some effects on AR-GARCH estimates as well as forecasts due to rounding, in particular for integer rounding. The same holds for correlation integral based tests for dependences (the BDS test and Savit and Green's dependability index); when price series are heavily rounded (to the nearest integer), we find that large shifts in the null-distributions of the tests render these useless in detecting chaos and other dependences. We also show how non-stationarities and "spurious" dependences in the series are created by the discreteness, and how this gives rise to shifts in the null-distributions of the statistical tests.

In the fifth essay (Chapter 6), *Stochastic Volatility and Pricing Bias in the Swedish OMX-Index Call Option Market,* I study the pricing of European call options when the underlying stock index (OMX-Index) volatility changes randomly over time. This differs from the Black-Scholes approach, where volatility is assumed to be constant over the life-time of the derivative security. Both the stock index and its volatility are modeled as diffusion processes; the stock index as a Geometric Brownian Motion, and the volatility as a mean-reverting square-root process. Stochastic Volatility option prices are calculated with the Fourier-Inversion method and process parameters are backed out from empirical market prices on the Swedish Exchange for Options and

Other Derivative Securities (OM). The stochastic volatility option prices are compared to Black-Scholes prices as well as to market prices, and the pricing bias is examined. Both models overprice options out-of-the-money and underprice options in-the-money. A dynamic hedging strategy reveals some mispricing in this market, and risk-free profit possibilities exist if transaction costs are neglected.

References

Alexander, C.O., and Chibumba, A.M. (1998), "Orthogonal GARCH: An Empirical Validation on Equities, Foreign Exchange and Interest Rates, unpublished manuscript.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, 31, 307-327.

Brock, W., Dechert, W., Scheinkmann, J., and LeBaron, B. (1996), "A Test for Independence Based on the Correlation Dimension", *Econometric Reviews*, 15(3), 197-235.

Crack, T., F., and Ledoit, O. (1996), "Robust Structure Without Predictability: The "Compass Rose" Pattern of the Stock Market", *Journal of Finance*, June, 751-762.

Ding, Z. (1994), *Time Series Analysis of Speculative Returns*, PhD Thesis, UCSD.

Engle, R. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica,* 50, 987-1007.

Hull, J., and White, A. (1987), "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance*, 42, 281-300.

Savit, R., and Green, M. (1991), "Time Series and Dependent Variables", *Physica D,* 50, 95-116.

Stein, E.M., and, Stein J.C. (1991), "Stock Price Distributions with Stochastic Volatility: An Analytical Approach", *Review of Financial Studies*, 4, 727-752.

Chapter 2

Orthogonal GARCH and Covariance Matrix Forecasting in a Stress Scenario: The Nordic Stock Markets During the Asian Financial Crisis 1997-1998

2.1 Introduction

Over the last decades, a great deal of the focus has shifted from the study of means of stock market returns to return volatilities. Correct modelling and forecasting of variances and covariances are important in many practical applications; for instance in option pricing and calculations of Value-at-Risk measures. It is therefore of great interest to compare different forecasting models and to find a model that gives more accurate forecasts. Much research has focused on univariate forecasting problems and the task of finding increasingly accurate volatility forecasts. Less work has been done within multivariate frameworks where covariances as well as variances have to be forecasted.

Traditionally, very simple models are used to describe and predict variances and covariances. One often relies on historical data and extrapolates into the future in a straightforward (unconditional) way. However, the increasingly frequent large shocks in the financial markets in the last decades have spurred the search for, and application of, more elaborate (conditional) models. Some of these models are feasible in univariate settings but difficult to generalize to multivariate specifications. Examples of this are the ARCH-type models developed by Engle (1982), Bollerslev (1986), Nelson (1991) and many others over the last two decades. Conceptually, these models are easily extended to multivariate frameworks, but problems arise when it comes to parameter estimations. The number of parameters explodes when more than two or three assets are modelled together, creating huge estimation problems for a risk manager who wants to estimate large ARCH or GARCH covariance matrices, perhaps for hundreds of assets.

Ding (1994) and Alexander and Chibumba (1998) have shown how this problem can be avoided by using orthogonal factors and principal component analysis in what the first author calls "Principal Component Multivariate ARCH" and the latter two call "Orthogonal GARCH"[1]. These two GARCH models are actually identical and the basic idea behind the model is to "diagonalize" the multivariate problem, so that only univariate GARCH estimations are needed[2]. The most important advantage of the Orthogonal GARCH method is the possibility to create arbitrarily large covariance matrices built on GARCH forecasts. In addition, these matrices have the nice

[1] In his thesis, Ding mentions how the model was suggested to him by Ron Khan at BARRA, Berkeley.

[2] In this paper, I call the model "Orthogonal GARCH" for the simple reason that I prefer that name to the lengthy one suggested by Ding.

feature of always being positive definite; something that is not always the case in, for instance, moving average forecasting models. Compared to classical multivariate GARCH models, Orthogonal GARCH estimations also seem to converge in situations where traditional GARCH models do not. Some financial variables that, traditionally, have been difficult to model might therefore possibly be modelled. The major drawback of the model is expected to reveal itself under extreme market conditions; in particular when unexpected events affect some of the assets in the system more than others. Under such stress scenarios, the orthogonality assumption partly breaks down (Alexander and Chibumba (1998)).

Regardless of how one chooses to model volatilities and correlations, a crucial question in studies on forecasting performance is how to evaluate the forecasts. A statistical approach, where comparisons between model predictions and the realized value are made by, for instance, calculating a measure like the root mean squared error (RMSE), can be adopted. This particular error statistic is symmetric in the loss function and assumes over and underpredictions to be considered as equally bad by a particular investor. In practical applications, an investor can be expected to put different weight on over and underpredictions and such an investor might therefore prefer an asymmetric error statistic. A problem with both the symmetric and the asymmetric statistical approach for volatility evaluation, however, is that realized volatility is not directly observable.

An alternative way to proceed is to look at "operational" procedures. The Bank for International Settlements (Amendment to the Capital Accord to Incorporate Market Risks (1996)) has suggested the use of "backtesting" to evaluate volatility and correlation forecasts in Value-at-Risk models and bank capital requirement calculations. An important difference between this approach and the statistical approach is that while measures like RMSE focus on the entire return distribution, backtesting focuses on the lower tail of the distribution[3]. This makes the backtesting of BIS more suited for Value-at-Risk evaluations.

[3]The realized variance in a measure like RMSE is calculated by using all returns over a certain sample period, positive as well as negative, large as well as small.

Yet another direction to take in order to evaluate forecasts of second moments is to exploit the role of covariances and variances in the pricing and hedging of contingent claims (Engle *et al.* (1993)). In order to price and hedge options whose payoff depends on more than one underlying asset, so called "rainbow options", forecasts of the whole covariance matrix are required. By setting up a hypothetical rainbow derivatives market, where a number of actors with different forecasts trade simulated rainbow options with each other, a preference free approach to the evaluation of covariance matrix forecasts is created (Gibson and Boyer (1998)). In such a setup, forecasting actors producing better forecasts should simply earn higher profits than other actors in the market.

In this essay, I present results from covariance/variance forecasting in a multivariate framework by using Orthogonal GARCH in addition to more traditional methods on four Nordic stock indices between 1996 and 1998. During this period, several countries in East and Southeast Asia were hit by a severe financial crisis that was spread to the rest of the financial world. In the Nordic countries, the effects of the crisis could be observed on plunging stock markets coupled with a rise in stock return volatility, larger than the volatilities during both the 1992 Swedish currency crisis and the Kuwait war. The markets in Finland, Norway, and Sweden were more badly affected than those in Denmark, which led to stock indices from the first three countries being slightly differently affected by the crisis compared to the Danish index. As mentioned, in addition to the high level of volatility, this is expected to weaken the performance of the Orthogonal GARCH method. As mentioned above, a well-known problem in this kind of studies is how to evaluate the different forecasts. Therefore, I include a whole range of different evaluation approaches; each evaluation measure corresponding better to a certain kind of economic agent.

Chapter 2.2 briefly describes the Asian crisis and presents some descriptive statistics on the data. Chapter 2.3 deals with covariance matrix modelling and presents the Orthogonal GARCH model, and chapter 2.4 describes the different evaluation measures. Chapter 2.5 contains the empirical results, and finally chapter 2.6 summarizes and concludes the essay.

2.2 The Nordic Stock Markets during the Asian Crisis

In order to evaluate different forecasting techniques, I use data collected from four Nordic stock exchanges in four different countries. Three out of these four countries (all but Norway) are members of the European Union. The Stockholm Stock Exchange and the Copenhagen Stock Exchange (soon also the Oslo Stock Exchange) use the same trading system and all four exchanges coordinate their distribution of price and turnover statistics.

2.2.1 The Asian Crisis

After 30 years of unprecedented economic performance, a serious crisis hit the East and Southeast Asian countries in mid 1997. The crisis "started" in Thailand during the summer and the crisis was soon spread from Thailand to the "neighboring" countries, Malaysia, South Korea, The Philippines, and Indonesia. In October 1997, the crisis was spread further through the internationally important Hong Kong economy to Europe (and the Nordic countries) as well as to the US. The uncertainties about the Asian economies and the effect of the crisis on the world economy as a whole were well reflected by the gloomy and highly volatile European and US stock markets during the whole second half of 1997. The first half of 1998 was a period of slight recovery in Asia, and the European financial markets temporarily became less nervous. However, what had started as a regional economic and financial crisis in East Asia in the summer of 1997 had slowly developed into a global financial crisis and by the summer of 1998, volatilities were again increasing on a global scale. In the period August to November 1998, the Nordic financial markets were close to being as volatile as during the October crisis 1987.

2.2.2 Data and Preliminary Statistics

The data consists of daily quotes from four Nordic stock indices that serve as underlying assets for options and futures trading in Sweden (OMX), Denmark (KFX), Finland (FOX), and Norway (OBX), respectively[4]. Each index represents a value weighted portfolio of the $25-30$ most traded stocks on the associated stock exchange and can be seen as a proxy for the stock price behavior in the actual country. The four markets have trading hours that almost perfectly overlap and 717 observations of daily returns from February 1, 1996 to December 29, 1998 are used. Descriptive statistics are given in Table 2.1.

All the series show excess kurtosis and some skewness, but the Ljung-Box statistics indicate no significant autocorrelations in the index returns. The squared return series Ljung-Box statistics, however, show large Q-values indicating autocorrelated squared returns, which explains some of the observed kurtosis.

In Figure 2.1, I plot volatilities (standard deviation (%) on an annual basis using a 20-days moving window) for the four Nordic stock index returns for the three-year period 1996-1998. In late 1997 and 1998, the volatilities in all four countries reached levels comparable to the levels during the Kuwait war. The patterns in Sweden, Finland, and Norway are very similar while the volatility in Denmark is lower systematically.

2.3 Covariance Matrix Modelling and Forecasting

For a financial institution, the forecasting of *covariances* between asset returns is important in addition to the forecasting of *variances*. The forecasted covariance matrix, that is forecasts of covariances and variances presented in matrix form, serves many purposes for a practitioner. Financial institutions can create large covariance matrices of all its positions in different assets and use them for a number of its operations; a portfolio manager might want to use them to create optimal portfolios, the option trader can price and hedge different kinds of contingent claims with the help of both variances and covariances of underlying assets, and risk managers easily transform covariance matrices into Value-at-Risk measures.

[4]All data are retrieved from Reuters Inc. A small number of observations were removed from the original dataset due to different dates of holidays in the four countries.

The last point in particular, that is, the use of variance and covariance forecasts in Value-at-Risk models, is of current interest since all financial institutions in the European Economic Community countries have been encouraged to calculate their own Value-at-Risk measures for inspection of their central banks since the end of 1997 ("Amendment to the Capital Accord" (1996a)); these (internal) Value-at-Risk models are supposed to create less stringent capital requirements than standard (external) capital requirement calculations. The importance of linkages between returns as well as volatilities in different markets is also emphasized by the Amendment to the Capital Accord, where natural hedges and correlations between bond markets, stock markets, and foreign exchange markets are suggested to be allowed (Belaisch and Kjeldsen (1997)). According to the rules worked out by The Bank for International Settlements (BIS) and the Basle Committee on Banking Supervision, central banks are supposed to regularly (every 3 months) evaluate their domestic institutions Value-at-Risk models by using operational methods and historical data from at least one year, "Supervisory Framework for the use of Backtesting" (1996b). On the basis of the performance of these Value-at-Risk models, the central bank calculates capital adequacy requirements for the bank in question.

In the light of this regulative aspect on Value-at-Risk models, together with the other uses mentioned above, it is obviously important to be able to forecast large covariance matrices. BIS requires the use of at least one year of historical data in its evaluation methodology but gives no guidelines as to how the forecasts should be calculated. A number of approaches are in use; regression methods, differently weighted moving averages, simulation methods, etc. GARCH models are usually difficult to use, but as suggested in Ding (1994), Alexander and Chibumba (1998), and as shown in this article, Orthogonal GARCH might be a powerful alternative to other simpler forecasting techniques.

2.3.1 Orthogonal GARCH

A well known problem with GARCH models is the difficulty in generalizing the model to a multivariate framework. All theoretical results are easily modified to many dimensions but problems arise when the parameters in the model equations must be estimated. Since the number of parameters explodes with increasing dimensions, model specifications restricting the parameterspace dimension are usually chosen. These specifications are all based on simplifying assumptions, and problems like non-positive definite covariance matrices may arise. To circumvent these problems, Ding (1994), and Alexander and Chibumba (1998) suggest a factor GARCH model using orthogonal factors. The main idea is to use principal components analysis to generate a number of orthogonal factors that can each be treated in a univariate GARCH framework.

I assume there are k return series with t observations (at time t) represented as a $t \times k$ matrix, $\mathbf{Y}_t$ [5]. The $t \times k$ matrix, $\mathbf{P}_t$, of principal components is defined as

$$\mathbf{P}_t = \mathbf{Y}_t \mathbf{W}_t \tag{1}$$

where $\mathbf{W}_t$ is the orthogonal $k \times k$ matrix of eigenvectors of $\mathbf{Y}_t^T \mathbf{Y}_t$ ordered according to the size of the corresponding eigenvalue. Notice that $\mathbf{P}_t$, like $\mathbf{W}_t$, is now an orthogonal matrix. By inverting (1), one gets the principal components representation of the system

$$\mathbf{Y}_t = \mathbf{P}_t \mathbf{W}_t^T.$$

Ω_t, the variance of $\mathbf{Y}_t$ at time t, can now be calculated as

$$\Omega_t = var(\mathbf{Y}_t) = var(\mathbf{P}_t \mathbf{W}_t^T) = \mathbf{W}_t \mathbf{D}_t \mathbf{W}_t^T \tag{2}$$

where $\mathbf{D}_t$ is a diagonal matrix of principal component variances at t and $\mathbf{W}_t$ is assumed to be known at time t[6]. This also enables us to calculate the forecasted covari-

[5] If the return series represent different kinds of assets with different behavior, it is sometimes necessary to first normalize the matrix by subtracting each column by its sample mean and divide it by its sample standard deviation.

[6] $\mathbf{W}_t$ does not change much from day to day and $\mathbf{W}_t$ can be approximated with $\mathbf{W}_{t-1}$ without introducing large errors in the calculation of the covariance matrix. This is particularly the case when calculating the *forecasted* covariance matrix. In this case, I forecast at time $t-1$, only using information up to $t-1$. (including $\mathbf{W}_{t-1}$)

ance matrix $\Omega_{t+1} \mid \Psi_t$, where Ψ_t is the information set at t, by univariate methods; for each principal component, the conditional variance of the principal component i, $var_{t+1}(P_i \mid \Psi_t)$, can easily be forecasted by, for instance, any univariate GARCH model. This gives us the Orthogonal GARCH specification. This particular covariance matrix also has the advantage of always being positive definite, since $\mathbf{D}_t$ is diagonal with positive elements along its diagonal.

For each principal component with the conditional variance modelled as GARCH(1,1) we have:

$$p_t = \alpha_0 + \varepsilon_t$$

$$\sigma_t^2 = \phi_0 + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2 \tag{3}$$

where σ_t^2 is the conditional variance of ε_t, $\varepsilon_t = \sigma_t u_t$, and $u_t \sim N(0,1)$. Most empirical studies suggest that an order of the GARCH model larger than $(1,1)$ is rarely needed. In GARCH modelling, the parameters are restricted; ϕ_0 must be larger than zero, ϕ_1 and ϕ_2 must be zero or larger, and the sum of ϕ_1 and ϕ_2 must be less than one in order to have a finite unconditional variance. The maximum likelihood estimates (using the BHHH algorithm implemented in *GAUSS*) of the GARCH model and some statistics on the standardized residuals are presented in Table 2.2. For all four principal components, we get significant parameter estimates, and $\phi-$parameters fulfilling the restrictions above. The residuals are well behaved and show no signs of large deviations from normality.

It is important to remember that just like most other multivariate GARCH specifications, Orthogonal GARCH is based on some assumptions. When assuming the conditional covariance matrix of the principal components to be diagonal, we foresee that only the *unconditional* covariance matrix is diagonal by construction. The *conditional* covariances need not be perfectly zero. Alexander and Chibumba (1998) recognize this and show how the errors induced by this assumption are quite small, at least for highly correlated return systems in tranquil periods.

2.3.2 Covariance Matrix Forecasts

The forecasts at time t of the volatilities from the univariate GARCH specification of the principal component ii, at different times $t+k$ in the future, are expressed as

$$\sigma^2_{ii,t+k} = \phi_0 \frac{1-(\phi_1+\phi_2)^{k-1}}{1-(\phi_1+\phi_2)} + (\phi_1+\phi_2)^{k-1}\sigma^2_{ii,t+1}, \qquad k \geq 2$$

where

$$\sigma^2_{ii,t+1} = \phi_0 + \phi_1\varepsilon^2_t + \phi_2\sigma^2_{ii,t}$$

This formula comes from iteration of the variance equation (3), and from this expression, I calculate variance forecasts for the principal components over different horizons, h, between one and fifteen days by simply summing up $\sigma^2_{ii,t+k}$ for all $k = 1$ *to* h[7]. Using equation (2) in chapter 2.3.1, I then transform the results for the principal components into the original index return series. In this way, I get explicit values of both variances (volatilities) and covariances (correlations) that can be compared to comparative models forecasts as well as realized values.

The following two variance expressions are used as comparative models[8]: at time t, I forecast $\sigma^2_{ii,t+k}$ (which is independent of k) by using the historical variance (HI)

$$\sigma^2_{ii,t+k} = \frac{1}{t-1}\sum_{s=1}^{t}\left(r_{ii,s} - \frac{1}{t}\sum_{j=1}^{t} r_{ii,j}\right)^2, \qquad k \geq 1$$

and a 20-day equally weighted moving average model (MA), where r_t is the daily return.

$$\sigma^2_{ii,t+k} = \frac{1}{19}\sum_{s=t-19}^{t}\left(r_{ii,s} - \frac{1}{20}\sum_{j=t-19}^{t} r_{ii,j}\right)^2. \qquad k \geq 1$$

[7]Since no autocorrelation in the return series is evident, it entails no correction.

[8]These expressions are very similar in spirit. Both are simple expressions for the unconditional variance (and covariance); only the number of used historical observations differs. In a stress scenario, they act fairly differently, however, and therefore, I emphasize their differences rather than their simularities by giving them different names.

The expressions for the covariances $\sigma_{ij,t+k}$ are much the same.

Variances and covariances over the different forecasting periods are then calculated by simply multiplying $\sigma^2_{ii,t+k}$ and $\sigma_{ij,t+k}$ by the number of days, h, over the forecasting horizon.

2.4 Out-of-Sample Evaluation of Covariance Matrix Forecasts

A difficult question to answer is how to best evaluate the various forecasts out-of-sample. An investor that trades options might care about the influence the volatility forecasts have on the whole return distribution while Value-at-Risk managers might care more about how common extreme returns are. Other investors might want to punish over-predictions more than under-predictions.

2.4.1 Symmetric and Asymmetric Statistical Evaluation

As a symmetric statistical evaluation method, I have chosen the traditional root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\sigma^2_{t,h,sample} - \sigma^2_{t,h,forecast}\right)^2},$$

where N is the number of days in the test sample, and h is, once more, the length of the forecast horizon. $\sigma^2_{t,h,sample}$ is simply the h-day sample variance (or covariance) of the realized returns over the h-day forecasting period commencing at t.

As an asymmetric extension, I follow Brailsford and Faff (1996) in constructing an error statistic that penalizes underpredictions more heavily than overpredictions:

$$ASY(U) = \frac{1}{N}\left[\sum_{t=1}^{O}\left|\sigma^2_{t,h,sample} - \sigma^2_{t,h,forecast}\right| + \sqrt{\sum_{t=1}^{U}\left|\sigma^2_{t,h,sample} - \sigma^2_{t,h,forecast}\right|}\right]$$

where O is the number of overpredictions and U is the number of underpredictions[9]. For instance, this kind of loss function might be appropriate for a seller of a call option since an underprediction of the stock price volatility will lead to a downward biased estimate of the call option price. By switching places for the O and the U in the formula above, we get the corresponding asymmetric error statistic that penalizes overpredictions harder than underpredictions. A forecast model giving "unbiased" estimates might, in some general sense, be considered better; unbiased meaning that the model overpredicts (underpredicts) 50 percent of the time.

2.4.2 Backtesting

Following the Bank for International Settlements, I also choose to look at their proposed operational method, that is backtesting. In this essay, backtesting is applied only to evaluate the diagonal elements of the forecasted covariance matrix, i.e. the variances. Backtesting can be done in different ways and this essay follows Alexander and Leigh (1997) who implement backtesting in two ways; "backward looking" where the current one-day 97.5% as well as 99% Value-at-Risk measures are compared to the last 250 returns for each day in the test period, and "forward looking", where each day in the test period the current one-day Value-at-Risk measure, is compared to the realized return of that day. Since we only study very short time horizon (one-day) Value-at-Risk measures, it is probably best to assume a zero mean return and simply calculate the Value-at-Risk measure as $VaR_{1-\alpha} = Z_\alpha \sigma$, where Z_α is the critical value for the standard normal distribution and σ is the forecasted (one-day) standard deviation.

Assuming moderate excess kurtosis, the average number of exceptions, returns that are more extreme than the Value-at-Risk measures, should be around 2.5 (0.01*250) for 99% VaR and 6.25 for 97.5% VaR in the backward looking version.[10] Due to sta-

---

[9] When using this asymmetric evaluation method, errors of equal size but with different signs do not cancel out as when the simple Mean Error statistic is used.

[10] The excess kurtosis as well as the skewness in the Scandinavian return series could, naturally, affect the critical VaR values measured in standard deviations (1.96 standard deviations from the

tistical problems and the approximative nature of the backtesting method, BIS has defined three zones for the evaluation of Value-at-Risk measures. The use of these zones is then believed to help bank supervisors (central banks) in their calculations of bank capital requirements. If the average number of exceptions in the backward looking version (for the last 250 days and the 99 % measure) is less than five then the model falls into the green zone; between five and ten exceptions puts the model in the yellow zone; and finally more than ten exceptions in the red zone. When calculating capital requirements, green models are OK, and yellow models should lead to capital requirement increases imposed on the bank in question to create incentives to improve forecasting and Value-at-Risk measures. Red models are assumed to seriously underestimate Value-at-Risk measures and should not be used at all. In addition to substantially increased capital requirements, bank supervisors should also require the bank to immediately begin working on improving its model.

Since it is not my intention to impose capital requirements on banks but to evaluate the accuracy of different forecasting models, I emphasize both the zones and the actual number of exceptions.

2.4.3 Pricing of Simulated Options

Finally, I evaluate covariance as well as variance forecasts by their ability to correctly price a certain type of rainbow option, that is, the option to exchange one asset for another (even called Outperformance Option). Among all options depending on more than one asset, the option to exchange one asset for another is one of few options with an analytical solution (Margrabe (1978))[11]. This option has a payoff function equal

mean for the $\lambda_{0.025}$ critical value and the normal distribution). The $\lambda_{0.025}$ critical values (in number of standard deviations) for the four return series for our three-year period are 2.08 for OMX, 2.32 for KFX, 2.15 for FOX, and 2.39 for OBX. In other words, large negative returns are more common than for the normal distribution (fattailedness).

[11]The Margrabe model is an extension of the Black-Scholes model and it assumes non-stochastic variances and covariance. This stands in contrast to what is observed in the real world as well as to my methodology of modelling second moments as time varying. However, the short time to maturity (one day) of the simulated options in my experiment minimizes the difference in option price between

to $\max[0, x_1 - x_2]$, where x_1 and x_2 are the prices of the two underlying assets, and the option price depends on both assets' variances as well as their covariance. In other words, the whole covariance matrix is needed in order to accurately price such an option. All options are at-the-money and their time to maturity is always one day.

In the setup described below, I follow Gibson and Boyer (1998) but instead of trading a single rainbow option, each actor now trades six *different* rainbow options each day[12]. Each day t in the test period, a certain sequence is followed:

1. **Forecasts.** Each actor forecasts, at $t-1$, the covariance matrix using the observations up to that date. In order to get a deeper market, I add a fourth actor to the other three forecasting actors (GARCH, MA, and HI). This fourth actor uses a strategy halfway between MA and HI; he simply uses a MA model but with a 350-day window instead of a 20-day one. This actor is called MA-long.

2. **Option Pricing.** Each actor uses his covariance matrix forecast to price a one-day at-the-money option at time $t-1$ to exchange one asset for a certain amount of an other asset at time t[13].

3. **Option Trading.** Actors trade options among themselves at $t-1$. Each actor trades six options with each of the other three actors. An actor who finds another actor's option to be underpriced buys the option and vice versa. In this

constant volatility and varying volatility models (Engle *et al.* (1993)). As mentioned by Gibson and Boyer (1998), the use of the same constant volatility option pricing model for all different forecasts should bias my results *against* finding any differences among forecasting agents.

[12]While Gibson and Boyer apply their evaluation technique to a two-dimensional problem, my setup has four different assets, which gives me six different options with different pairs of underlying assets. According to Gibson and Boyer (1998), up to 1000 observations are needed to get statistically significant results. Since my test period only has 345 observations, this might be a problem. However, by trading six times as many different options as Gibson and Boyer, I bypass this problem (Byström (1999)).

[13]Index levels in the four countries are all calculated in the same currency, the Swedish krona (SEK). The exchange rates are assumed to be constant at the rates prevailing on February 1, 1996.

way, each actor trades 18 options each day, his "bank account" being credited (if a seller) or debited (if a buyer) with the mean of the two traders' options prices.

4. **Hedging.** For each of the six options, each trader hedges his exposure to the two underlying assets by going short or long in these assets at time $t-1$; an amount equal to his options position's sensitivity to changes in each underlying asset (delta hedging). His bank account is again credited or debited with the amount needed for the trades. It is important to remember that the quality of the different forecasts affects the hedging performance as well as the pricing performance.

5. **Payoff.** Money in the bank account earns one day's interest at the riskfree rate[14]. Using actual returns between time $t-1$ and time t, the payoffs (possibly zero) of the options are calculated. The bank accounts are again credited or debited with these payoffs. The hedge positions of the underlying assets are sold at t and the bank account is once more debited or credited.

6. **Accounting.** Going from $t-1$ to t, the balances of the bank account of each actor's have changed with an amount equal to that day's profit (positive or negative).

   The last day in the test period, the actor with the "best" forecasts should have made the highest accumulated profit.

2.5 Empirical Results

Evaluating forecasts of second moments is very hard and it is not my intention to present a single forecasting model that gives superior forecasts in every situation, for every financial asset, every forecasting horizon, and every evaluation method. Instead, my aim is to throw some light on the performance and the stability of the Orthogonal

[14]The interest rate used is the Swedish 30-day treasury bill rate (statsskuldsväxel).

GARCH approach to forecast large covariance matrices. In addition, it should be remembered that there exist some reasons for the Orthogonal GARCH methodology to perform relatively weakly for my particular choice of time period; the time period is highly volatile (stress scenario), and the assets in question react to shocks in a slightly asymmetric way.

The unconditional *correlation* matrix (over the whole sample) is shown below. The largest correlation is that between the Swedish OMX index and the Finnish FOX index, and the smallest the one between the Swedish OMX index and the Danish KFX index.

	OMX	KFX	FOX	OBX
OMX	**1.00**	**0.58**	**0.74**	**0.66**
KFX	0.58	**1.00**	**0.67**	**0.63**
FOX	0.74	0.67	**1.00**	**0.71**
OBX	0.66	0.63	0.71	**1.00**

When Orthogonal GARCH is put to use it is marginally more time consuming than more traditional regression or moving average methods. As mentioned before, it also has some model inherent advantages compared to simpler models, and the behavior of Orthogonal GARCH forecasts might be of interest, even if no systematic forecasting superiority is found.

To evaluate the performance of the different forecasting models out-of-sample, I divide my set of time series into an (expanding) *estimation period* and a *test period*. The test period contains the last 345 days of the sample and is divided into three equally long subperiods of 115 days each. These subperiods are chosen to coincide with the rise and fall of volatility and to have equal length; they are all fairly volatile in an historical perspective but two of these, August 7, 1997 to January 23, 1998 and July 15, 1998 to December 29, 1998, are clearly more volatile than the third, January 24, 1998 to July 14, 1998. From here on, I call the less volatile period 1998 ”tranquil” and the volatile periods ”volatile 97” and ”volatile 98”, respectively. For each period, I apply the statistical methods, the operational evaluation method, as well as the option methodology described above.

2.5.1 Entire Period

The RMSE results for the entire test period August 7, 1997 to December 29, 1998 are presented in Table 2.3. Before turning to the individual forecasting models and their relative performance, a quick look at the numbers in the table tells us how higher stock index variances as well as longer forecast horizons typically give larger errors[15]. It can also be seen that covariance forecasts are at least as good as variance forecasts.

Focusing on the relative performance of the different models, it is immediately observed how the ranking of the models remains more or less constant over different horizons. It is equally obvious how GARCH overall dominates both the 20-day moving average model (MA) and the historical variance model (HI); GARCH is best in 23 out of 30 cases and in the other 7 cases, it is second best after MA. The weak performance of Orthogonal GARCH in the particular case of the KFX-index is probably due to the slightly asymmetric behavior of the Danish KFX-index compared to the other indices. Reducing the number of principal components would probably improve GARCH forecasts in this case[16].

The results from the asymmetrical statistical evaluation methods are shown in Table 2.4. In the upper part, overpredictions are punished harder than underpredictions, $ASY(O)$, and in the lower part, underpredictions are punished harder than overpredictions, $ASY(U)$. For the $ASY(O)$ measure it is obvious how for all variances, covariances, and horizons, HI systematically dominates GARCH, which, in turn, systematically dominates MA. Not surprisingly, the historical model underpredicts considerably during this volatile period, and the relative ranking of the models would be expected to be the opposite, when underpredictions are punished harder than overpredictions. However, even if the historical model is clearly the worst performer

[15]The larger is the volatility of the asset, the larger the error (FOX>OMX>OBX>KFX).

[16]It must be remembered that each principal componenet is built up from all asset series. KFX forecasts would probably be improved if the principal component associated with the common movements of the other three indicies could be identified and removed.

according to $ASY(U)$, MA does not systematically dominate GARCH this time. In other words, for the two asymmetrical error statistics taken together, GARCH dominates MA. In addition, GARCH (as well as MA) has the attractive feature of acting fairly unbiased (neither heavy over nor heavy underpredictions) compared to the historical method. As for the symmetrical evaluation method, Orthogonal GARCH works less well for the KFX index than for the other indices.

While both variance and covariance forecasts are evaluated with the statistical approaches above, backtesting is performed on variances only. The results from the "backward looking" version are presented in Table 2.5 for the 97.5% VaR measure as well as for the 99% VaR measure. The theoretically predicted average number of exceptions is 6.25 and 2.5, respectively, and the best (smallest value) model in each category is typed in bold. Table 2.5 contains evidence of GARCH dominating MA as well as HI. Orthogonal GARCH clearly dominates the other two models in one-day forecasting of simple variances by giving a low number of exceptions in accordance with the guidelines of BIS. Not in a single situation is GARCH beaten by any of the other two models. Overall, the historical model is worse than the other two models and ends up in the yellow region for all four indices (99% VaR).

In Table 2.6, I present the results from the "forward looking" interpretation of the BIS backtesting technique. Once more, GARCH dominates the other two models, but this time, MA performs almost as well as GARCH. The historical model is, once more, the weakest performer and this time, it ends up in the red zone for all indices. In Figure 2.2, I have plotted 97.5% VaR estimates for the three different models as well as the returns for the OMX index over the entire test period. The two conditional VaR measures, GARCH and MA, increase (decrease) when the size of the returns increases (decreases) and both models catch much of the movement in return volatility. The underestimation of volatility of the historical VaR measure is also evident as is the slow adjustment of the model to changes in volatility. Overall, the models perform quite badly when evaluated with the forward-looking approach and in many cases they end up in the yellow or the red zone. This is not very surprising, considering the very volatile market (in the test period) and the associated difficulties in creating accurate forecasts and VaR measures.

Results from the option evaluation methodology is presented in Table 2.7 and in Figure 2.3. Table 2.7 contains the mean daily profit and its standard deviation. More or less the same information is presented in Figure 2.3, where the accumulated profits from the different forecasting techniques over the test period are shown. It is quite clear how Orthogonal GARCH gives higher profit than the other forecasters, thereby indicating better forecasts of the whole covariance matrix. Traders using Orthogonal GARCH or MA earn positive profits from trading rainbow options, while both HI and MA-long traders lose money. HI is the worst performer and it is clear how incorporating longer series of historical observations in variance and covariance forecasts decreases the profitability of option trading based on these forecasts.

To summarize, Orthogonal GARCH dominates models based on simple historical variances and covariances, and in addition to always giving positive definite covariance matrices, it gives accurate forecasts in a highly volatile market. According to the statistical evaluation, our straight-forward GARCH method dominates both the moving average and the historical method. Looking at the performance of the different models in an operational Value-at-Risk context, where the focus is on the lower tail of the return distributions rather than the whole distribution, gives the same results. Furthermore, GARCH also dominates the other models in the rainbow option trading experiment.

2.5.2 Subperiods

By dividing the entire test period into subperiods, we get an idea about the stability of the results in the former chapter as well as some indications of whether the level of volatility affects the Orthogonal GARCH methodology. In Table 2.8, I present both the RMSE and the two asymmetric error statistics for each of the three subperiods. In Chapter 2.5.1, it was shown how the results for the covariances were very similar to the results for the variances, and for clarity, I only present forecast errors for variances in this chapter. The same holds for the different horizons, and only 5-days forecast errors are presented in this chapter[17]. When comparing the relative performance of the different models, it immediately becomes clear how GARCH dominates both the 20-day moving average model and the historical model for all subperiods, when we evaluate using RMSE. The pattern from the entire period also repeats itself for the asymmetrical evaluation methods; GARCH clearly dominates MA for two of the periods and for the third period, GARCH always performs better than at least one of the other two models. The underprediction by HI for the entire period repeats itself for each of the subperiods. The only period that stands out slightly is the "tranquil" period. In this period, GARCH works even better than in the other periods while the historical method shows less of its heavy underpricing. Overall, for the statistical evaluation measures, most of the results from the entire period remain valid for each of the individual subperiods.

For the operational method, we present the results for both the backward-looking and the forward-looking version of back testing in Table 2.9[18]. Even here, the results are fairly stable over the different subperiods, and at least for the backward looking approach does GARCH dominate the other models in each period. For the forward looking approach, GARCH dominates (or performs equally well as) the other models in two of the three periods while in the volatile period 1997, MA slightly outperforms GARCH as well as HI. As for the statistical evaluation method, GARCH and, in particular, HI works relatively better in the tranquil period.

Turning to Figure 2.3 and the option trading experiment it can be seen how the performance of the different forecasting algorithms changes slightly from period to period. For the two volatile periods the results from the entire period remain; GARCH dominates MA (for the volatile period of 1997, MA does almost as well as GARCH) that, in turn, dominates MA-long and HI. For the relatively tranquil period, the performance of MA deteriorates while HI performs better than in the volatile periods. Overall, the results are more or less in accordance with the results

[17]The results for other horizons as well as for covariances do not differ from those presented to any extent.

[18]97.5% and 99% VaR show very similar results and the 99% measure is therefore left out.

from both the statistical evaluation and the backtesting.

In this chapter, I find that the results from the previous chapter remain more or less valid for each of the three subperiods. Orthogonal GARCH outperforms both of the two simpler forecasting techniques (as well as MA-long for the options approach) irrespective of which evaluation measure and time period is chosen. The moving average works better than the historical method in the more volatile periods, while HI dominates MA in the tranquil period (a period that is in no way tranquil seen in a historical perspective).

2.6 Conclusions

This essay has studied forecasts of variances and covariances from the Orthogonal GARCH model as well as from simpler models in a multivariate framework of Nordic stock index returns. Orthogonal GARCH produces positive definite covariance matrices without the typical estimation problems associated with multivariate GARCH models. A look at the Nordic stock markets during the Asian crisis 1997 and 1998 shows how Orthogonal GARCH performs relatively well in forecasting covariance matrices under volatile market conditions.

Different users of forecasts define "good" and "bad" forecasts in different ways. Therefore, I assess the predictive performance of Orthogonal GARCH, using a number of different evaluation methods. Both symmetrical and asymmetrical error statistics show (although no statistical inference is made) how Orthogonal GARCH produces better forecasts than the use of historical variances and covariances as well as forecasts from an equally weighted moving average model. In addition to statistical evaluation methods, I also apply operational methods. The "backtesting" approach suggested by BIS is particularly well suited to Value-at-Risk contexts and even here, Orthogonal GARCH dominates the historical variance as well as the moving average model in forecasting the whole covariance matrix.

Using simulated options to evaluate forecasts of the second moments also demonstrates the dominance of Orthogonal GARCH over the simpler models. A market actor using Orthogonal GARCH to predict the whole covariance matrix earns a significant profit from trading over/under priced rainbow options with actors using historical variances and covariances as well as moving average models. From this, I draw the conclusion that Orthogonal GARCH predictions are the most accurate.

Overall, Orthogonal GARCH seems to perform well as a covariance matrix forecaster in highly volatile periods. It is difficult to statistically assess the significance of the results, but the results in this essay suggests Orthogonal GARCH to be a good forecaster compared to different standard covariance matrix forecasters. The use of historical averages using data extending far into the past generally gives the worst forecasts in this very volatile period of the Nordic stock markets. Only in a few cases does Orthogonal GARCH perform worse than the other predictors in the study. By using longer estimation periods and more elaborate ARCH-type models, this orthogonalization approach should be able to produce even better forecasts. The essay demonstrates the importance of including GARCH-effects in volatility and covariance forecasts and how the orthogonalizing technique makes multivariate forecasting from GARCH models feasible.

References

Alexander, C.O., and Leigh, C.T. (1997), "On the Covariance Matrices Used in Value at Risk Models", *Journal of Derivatives* 4 no.3, 50-62.

Alexander, C.O., and Chibumba, A.M (1998), "Orthogonal GARCH: An Empirical Validation on Equities, Foreign Exchange and Interest Rates, unpublished manuscript.

"Amendment to the Capital Accord to Incorporate Market Risks." Bank for International Settlements, 1996a.

Belaisch, A, and Kjeldsen, K (1997), "Value-At-Risk Models for Determining Capital Adequacy: Does Allowance for Correlations Within Broad Risk Categories Matter?", Chiefs Economist's Department European Investment Bank, Economic and Financial Report 97/02.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.

Brailsford J. T., and Faff, R.W. (1996), "An Evaluation of Volatility Forecasting Techniques", *Journal of Banking and Finance* 20, 419-438.

Byström, H. (1999), "Using Simulated Currency Rainbow Options to Evaluate Covariance Matrix Forecasts", unpublished manuscript, Lund University.

Ding, Z., (1994), *Time Series Analysis of Speculative Returns*, PhD Thesis, UCSD.

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, 987-1007.

Engle, R.F., Hong, C.S., Kane, A., and Noh, J. (1993), "Arbitrage Valuation of Variance Forecasts with Simulated Options." *Advances in Futures and Options Research,* 6, 393-415.

Gibson, M.S., and Boyer, B.H. (1998), "Evaluating Forecasts of Correlation Using Option Pricing", *Journal of Derivatives*, Winter 1998, 18-38.

Margrabe, W. (1978), "The Value of an Option to Exchange One Asset for Another", *Journal of Finance*, 33, 1, 177-186.

Nelson, D. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347-370.

"Supervisory Framework for the Use of 'Backtesting' in Conjunction with the Internal Models Approach to Market Risk Capital Requirements." Bank for International Settlements, 1996b.

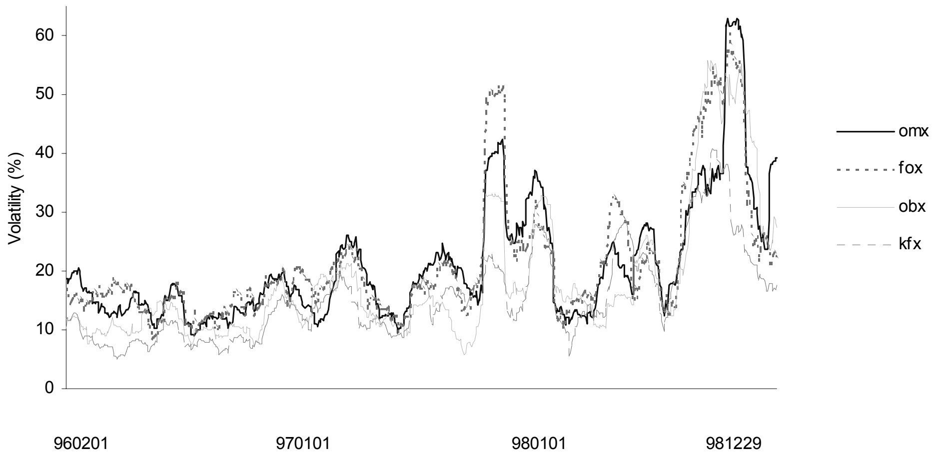

Figure 2.1: Volatilities for the Nordic stock indicies 1996-1998 (20-day window).

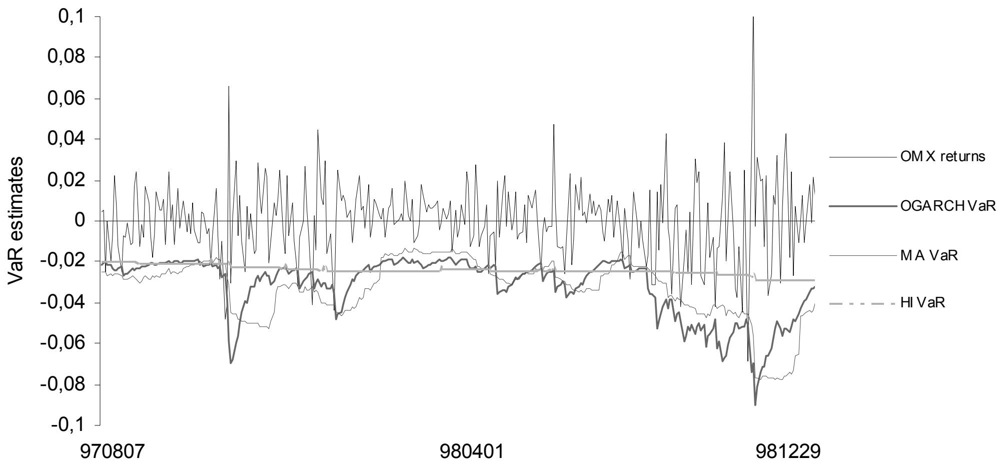

Figure 2.2: OMX VaR estimates (97.5%) and OMX returns over the test period.

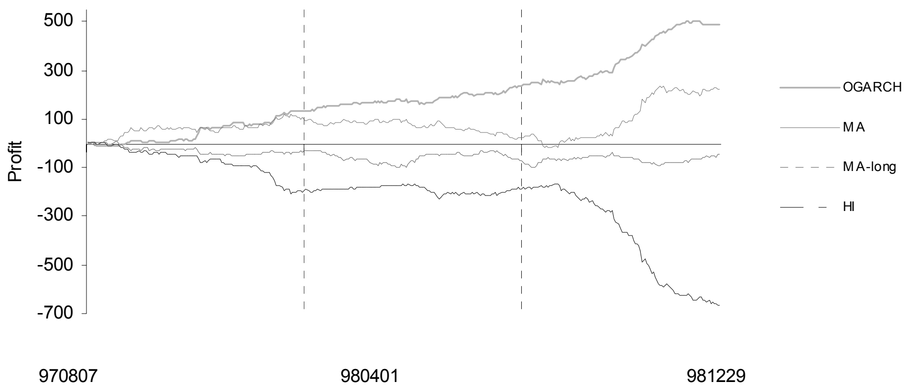

Figure 2.3: Accumulated profits from trading rainbow options.

Table 2.1: Descriptive statistics on daily return data. February 1, 1996 to December 29, 1998.

Statistics	OMX	KFX	FOX	OBX
No. of observations	717	717	717	717
Mean (%)	0.098	0.093	0.114	0.026
Std. Dev. (% on a yearly basis)	23.79	17.30	24.35	21.91
Skewness	0.25	-0.58	-0.69	-0.47
Excess Kurtosis	5.33	2.89	4.39	4.56
Ljung-Box:				
Q(6)	5.25	14.04	8.31	12.31
Q(12)	22.13	20.74	14.79	22.88
Q(18)	52.10	31.60	32.26	36.49
Ljung-Box (Squared Returns):				
Q(6)	140.84	200.10	247.37	362.43
Q(12)	191.82	313.40	347.48	635.29
Q(18)	237.56	417.22	441.30	917.86

99 percent critical values for Ljung-Box are: 16.8, 26.2, and 34.8.

Table 2.2: Typical GARCH(1,1) volatility parameter estimates for the four Principal Components, as well as statistics on the standardized residuals.

	First PC	Second PC	Third PC	Fourth PC
$\phi_0 \cdot 10^6$	21.91 (9.31)	0.10 (0.043)	1.30 (0.95)	0.10 (0.082)
ϕ_1	0.17 (0.038)	0.027 (0.0095)	0.038 (0.016)	0.026 (0.0093)
ϕ_2	0.76 (0.056)	0.97 (0.0093)	0.93 (0.045)	0.97 (0.0090)
Residuals:				
Mean	-0.041	0.012	-0.0042	0.0089
Standard deviation	0.999	1.005	1.005	0.993
Skewness	-0.390	0.189	0.0782	0.0591
Excess Kurtosis	0.279	0.250	0.697	1.178

The first 530 observations are used for these particular estimates. Small figures are standard errors.

Table 2.3: RMSE for variance and covariance forecasts over different horizons. Values are for the entire test period 970807 - 981229.

	5-days			10-days			15-days		
Variances	GARCH	MA	HI	GARCH	MA	HI	GARCH	MA	HI
OMX	**10.46**	*15.75*	18.90	**21.94**	*34.44*	37.20	**34.96**	*54.10*	54.72
KFX	*6.42*	**5.72**	8.94	*13.42*	**12.74**	17.75	*21.13*	**20.71**	26.34
FOX	**13.01**	*15.28*	20.91	**28.47**	*33.75*	41.34	**46.17**	*54.16*	61.09
OBX	**9.70**	*10.89*	19.51	**21.10**	*24.58*	38.82	**35.50**	*40.15*	57.82
Covariances									
OMX/KFX	**5.38**	*5.54*	8.71	**11.38**	*12.31*	17.28	**18.31**	*19.77*	25.67
OMX/FOX	**10.06**	*13.33*	16.27	**21.68**	*29.13*	32.02	**35.11**	*45.96*	47.05
OMX/OBX	**7.27**	*9.54*	13.97	**15.48**	*21.07*	27.60	**25.51**	*33.75*	40.83
KFX/FOX	*7.57*	**7.07**	11.57	*16.46*	**15.92**	22.98	*26.83*	**26.16**	34.15
KFX/OBX	*6.42*	**6.32**	10.89	**13.69**	*14.37*	21.64	**22.19**	*23.82*	32.19
FOX/OBX	**9.59**	11.55	17.39	**21.28**	*25.80*	34.49	**35.07**	*41.93*	51.13

GARCH = Orthogonal GARCH, MA = 20-days moving average model, and HI = historical variance. The smallest value is typed in bold letters and the second smallest value is typed in italics.

Table 2.4: Asymmetric error statistic for variance and covariance forecasts over different horizons. Values are for the entire test period 970807 - 981229.

	Overpred. Punished Harder								
	5-days			10-days			15-days		
Variances	GARCH	MA	HI	GARCH	MA	HI	GARCH	MA	HI
OMX	*97.16*	129.31	**47.52**	*134.87*	199.95	**70.06**	*159.75*	253.36	**86.94**
KFX	*71.02*	96.35	**21.95**	*104.14*	145.54	**32.00**	*130.76*	187.20	**38.63**
FOX	*100.85*	129.71	**54.79**	*146.86*	206.18	**82.13**	*191.50*	271.17	**102.83**
OBX	*85.91*	111.36	**40.41**	*129.66*	183.41	**61.53**	*166.90*	241.73	**78.10**
Covariances									
OMX/KFX	*51.99*	83.71	**25.74**	*77.22*	129.65	**38.55**	*96.70*	162.45	**47.49**
OMX/FOX	*91.37*	120.74	**51.41**	*127.14*	186.34	**76.45**	*159.14*	241.91	**93.82**
OMX/OBX	*60.85*	110.07	**44.49**	*93.10*	171.56	**66.33**	*121.08*	218.67	**80.63**
KFX/FOX	*64.64*	103.70	**38.69**	*96.89*	159.64	**56.03**	*127.57*	203.64	**69.43**
KFX/OBX	*76.48*	96.40	**38.62**	*112.60*	147.61	**56.24**	*140.50*	191.61	**69.33**
FOX/OBX	*88.77*	117.95	**47.63**	*132.89*	183.77	**70.74**	*169.29*	241.14	**87.67**

	Underpred. Punished Harder								
	5-days			10-days			15-days		
Variances	GARCH	MA	HI	GARCH	MA	HI	GARCH	MA	HI
OMX	**139.44**	*139.73*	248.88	**215.78**	*219.61*	357.87	**291.25**	*293.69*	447.94
KFX	*128.08*	**100.10**	200.33	*193.38*	**155.56**	284.28	*247.07*	**202.31**	351.44
FOX	**152.07**	*154.26*	255.87	*241.61*	**238.86**	361.37	*326.27*	**308.67**	443.22
OBX	*128.23*	**121.65**	237.40	*195.24*	**190.52**	341.43	*257.45*	**252.95**	426.75
Covariances									
OMX/KFX	*115.44*	**98.98**	197.63	*177.77*	**155.10**	281.17	*238.87*	**202.37**	347.71
OMX/FOX	**127.68**	*134.90*	225.82	**206.16**	*210.82*	320.48	*285.20*	**277.39**	394.76
OMX/OBX	**115.71**	*116.69*	204.85	**176.94**	*185.48*	291.61	**243.55**	*250.18*	360.20
KFX/FOX	*133.98*	**107.08**	202.55	*206.24*	**166.10**	287.49	*271.03*	**218.22**	354.25
KFX/OBX	**96.75**	*99.21*	173.63	**152.22**	*155.43*	247.81	**199.84**	*201.88*	306.71
FOX/OBX	**113.64**	*118.04*	215.97	*185.68*	**184.58**	308.01	*258.21*	**243.89**	380.54

GARCH = Orthogonal GARCH, MA = 20-days moving average model, and HI = historical variance. The smallest value is typed in bold letters and the second smallest value is typed in italics.

Table 2.5: Operational evaluation. Backtesting "backward looking" according to BIS with 97.5 percent VaR as well as 99 percent VaR. Values are for the entire test period 970807 - 981229.

	Var-97.5 %			VaR-99 %		
	GARCH	MA	HI	GARCH	MA	HI
OMX	**6.88**	7.26	11.44	**4.37**	4.70	7.19
KFX	**6.48**	6.60	11.96	**3.89**	4.04	8.56
FOX	**8.47**	8.97	13.68	**5.76**	5.89	8.16
OBX	**5.91**	7.66	12.63	**3.75**	5.11	8.59

GARCH = Orthogonal GARCH, MA = 20-days moving average model, and HI = historical variance. The smallest value is typed in bold. (Theoretical values are 6.25 and 2.5, resp.)

Table 2.6: Operational evaluation. Backtesting "forward looking" according to BIS with 97.5 percent VaR as well as 99 percent VaR. Values are for the entire test period 970807 - 981229.

	Var-97.5%			VaR-99 %		
	GARCH	MA	HI	GARCH	MA	HI
OMX	**15**	19	31	9	**8**	22
KFX	**15**	**15**	30	11	**10**	22
FOX	**17**	18	31	**12**	**12**	26
OBX	**12**	18	37	**7**	12	30

GARCH = Orthogonal GARCH, MA = 20-days moving average model, and HI = historical variance. The smallest value is typed in bold. (Theoretical values are 8.625 and 3.45, resp.)

Table 2.7: Option trading simulation. Mean daily profit and its standard deviation. Values are for the entire test period 970807 - 981229.

	GARCH	MA	MA-long	HI
Mean Daily Profit	1.626	0.808	-0.085	-2.352
Standard deviation	0.286	0.382	0.240	0.440

GARCH = Orthogonal GARCH, MA = 20-days moving average model, MA-long = 350-days moving average model, and HI = historical variance.

Table 2.8: RMSE as well as asymmetric error statistics for 5-days variance forecasts. Values are for the three equally long subperiods.

	RMSE			ASYM(O)			ASYM(U)		
	GARCH	MA	HI	GARCH	MA	HI	GARCH	MA	HI
Volatile 97									
OMX	**9.62**	*11.56*	12.91	*39.69*	116.71	**13.88**	*172.67*	**132.53**	265.08
KFX	3.51	**2.74**	*3.50*	*49.19*	62.69	**5.68**	*108.05*	**77.75**	157.54
FOX	**15.36**	*20.15*	20.16	*59.26*	145.53	**26.62**	*172.22*	**157.01**	269.97
OBX	**6.34**	8.76	*8.04*	*82.83*	108.57	**46.74**	**81.49**	*87.97*	144.77
Tranquil									
OMX	**3.86**	5.34	*3.93*	108.84	**75.87**	*100.43*	**71.04**	94.10	*78.00*
KFX	**4.14**	*5.21*	5.42	*56.26*	88.45	**48.08**	*104.70*	**88.07**	124.64
FOX	**4.82**	6.63	*5.53*	**83.88**	*97.46*	97.88	**86.72**	114.67	*101.80*
OBX	**2.56**	5.35	*2.74*	*68.99*	83.74	**45.26**	**52.93**	76.07	*64.11*
Volatile 98									
OMX	**15.80**	*26.06*	32.22	*156.27*	170.57	**25.58**	**155.56**	*225.30*	433.10
KFX	*10.01*	**7.87**	14.80	*108.13*	122.49	**12.31**	*183.85*	**138.66**	325.10
FOX	**17.62**	*19.09*	33.22	149.63	*144.77*	**31.63**	**227.39**	*238.32*	466.92
OBX	**16.02**	*16.73*	35.07	**26.67**	147.92	*32.03*	*291.85*	**199.94**	512.78

GARCH = Orthogonal GARCH, MA = 20-days moving average model, and HI = historical variance. The smallest value is typed in bold letters and the second smallest value is typed in italics.

Table 2.9: Operational evaluation. Backtesting (backward as well as forward looking) according to BIS with 97.5 percent VaR. Values are for the three equally long subperiods.

		Backward Looking			Forward Looking		
		GARCH	MA	HI	GARCH	MA	HI
Volatile 97	OMX	3.78	**2.34**	6.44	**5**	**5**	10
	KFX	6.16	**5.05**	8.60	**4**	**4**	8
	FOX	**4.68**	4.70	7.40	7	**6**	10
	OBX	**4.73**	7.95	7.11	**3**	**3**	8
Tranquil	OMX	11.68	15.80	**11.55**	3	3	**2**
	KFX	**8.27**	9.77	10.29	**4**	**4**	**4**
	FOX	**13.24**	14.47	13.69	2	**1**	**1**
	OBX	**9.80**	11.13	11.79	**2**	3	3
Volatile 98	OMX	**3.55**	5.18	16.05	**3**	9	16
	KFX	**4.89**	5.46	16.85	6	**5**	16
	FOX	**5.99**	7.36	19.58	**6**	7	16
	OBX	5.24	**5.00**	18.91	**6**	9	20

The smallest value is typed in bold letters. Theoretical values are 6.25 for backward looking and 2.875 for forward looking.

Chapter 3

The Hedging Performance of Electricity Futures on the Nordic Power Exchange Nord Pool

3.1 Introduction

One of the latest markets to become standardized and organized on an exchange is the electric power market. Allover the world, there are only a handful of operating electricity exchanges, with the Nordic Power Exchange, "Nord Pool", being the only multinational one. On this exchange, there are actors from Norway, Sweden, Denmark, Finland, and England. The early deregulation of their electricity markets is an important factor behind the development of an exchange for trading in electricity in the Nordic countries; in particular the Norwegian and the Swedish markets have become fairly competitive in the last few years.

With the setup of an organized exchange, the situation for electricity producers and distributors has changed; from a situation where a reliable supply of energy was most important, the focus has partly shifted to obtaining optimal financial performance and efficient risk management. On the Nordic Power Exchange, electricity is traded both on a spot market and a futures market. The main reason for trading in futures is for actors to monitor the volatility of their power portfolios and to minimize the negative effect of adverse fluctuations in electricity prices.

This essay investigates the statistical and distributional properties of spot and futures prices on Nord Pool, as well as the short-term hedging performance of the futures. An evaluation of the hedging performance is principally of interest in the electricity market, due to problems with the storage of electricity when arbitrage arguments are used for the pricing of futures. The high volatility in this market, many times as large as in traditional financial markets, also contributes to make hedging important.

When hedging *price risk*, the optimal proportion of the future contract that should be held to offset the cash position is called the optimal *hedge ratio*. This ratio is traditionally estimated by examining the ratio between the *unconditional* covariance between cash and futures prices and the *unconditional* variance of the price of futures. This method can be criticized on a number of grounds. First, the traditional optimal hedge ratio is only utility maximizing under certain assumptions on the futures returns, otherwise it is merely variance minimizing as shown by Myers (1991). Second, since almost all financial assets and commodities have time varying second moments, the hedge ratio will be time varying and possibly best modelled in a dynamic framework as a function of *conditional* covariances and variances (Baillie and Myers (1991)). These conditional covariances and variances are often modelled with the conditionally heteroscedastic ARCH- and GARCH-type models developed by Engle (1982) and Bollerslev (1986). Baillie and Myers (1991), for instance, employ a bivariate GARCH model to estimate hedge ratios for commodities and find that dynamic GARCH-based hedge ratios out-perform those ratios coming from the traditional unconditional approach.

In this essay, I consider time varying variances and covariances of Nordic electricity price returns over the period January 1996-October 1999 and investigate how the time variation affects the hedging performance out-of-sample on the Nordic Power Exchange. In addition to the traditional unconditional hedges, I apply different conditional hedges. On the one hand, I apply continuously updated 50-day moving averages of the second moments, and on the other, I apply two different GARCH models; first, the constant conditional correlation bivariate GARCH model proposed by Bollerslev (1990), and second, a multivariate GARCH model called Orthogonal GARCH where a "diagonalization" of the bivariate problem simplifies the multidimensional GARCH estimation (Ding (1994), Alexander and Chibumba (1998), Byström (1999)).

I find that short-term hedging of electricity spot prices with electricity futures, using different estimates of the optimal (or actually minimum variance) hedge ratio, systematically reduces the variability of the portfolio returns. My empirical findings also confirm that variances and hedge ratios vary significantly over time, but that the two GARCH models only slightly improve the hedging performance out-of-sample compared to the unconditional hedges. The traditional simpler hedging models perform as well as the more elaborate conditional models if the performance of the hedges is evaluated on the basis of their ability to reduce the unconditional (sample) portfolio variance. If we look at the conditional variance of the hedging portfolios instead, then the GARCH based hedges outperform the other hedges.

The essay is organized as follows: chapter 3.2 describes the data and the statistical particularities of electricity prices as well as the general features of the Nordic electricity market and Nord Pool, chapter 3.3 deals with hedging and hedge ratios, and chapter 3.4 presents the hedging results. Chapter 3.5 concludes the essay.

3.2 Spot and Futures on The Nordic Power Exchange Nord Pool

In January 1996, the Swedish electricity market was deregulated and integrated with the previously deregulated Norwegian electricity market. At the same time, the first multinational electricity exchange, Nord Pool, was created. The exchange has participants from Sweden, Norway, Denmark, Finland, and England (January 1998), but only the Norwegian and the Swedish markets are fully integrated. Of all trade in electric power in these two countries, around 25% (January 1998) is managed by Nord Pool; the main part of the electricity trade is still organized as bilateral contracts between producers and consumers.

In this essay, I will deal with the two major markets on Nord Pool, the spot market and the futures market. The spot market is a market for physical delivery of electricity, while the futures market is organized as a purely financial market without physical delivery. In reality, the spot market is also a short-term (one-day) futures market, however. Each day at noon, spot prices and volumes for each hour the *following* day are determined at an auction; what is called a futures contract in the electricity market is actually a futures contract with a *future* as the underlying asset[1]. Otherwise, the futures contracts are highly standardized and defined in terms of a given number of MWhs of electricity for (hypothetical) delivery during a given future week. The contracts are designed to reduce risk and make it possible to secure electricity prices up to three years in advance. It is possible to go short (and long) in the futures market, but the physical properties of the commodity in question, electricity, makes it difficult (or impossible) to go both short and long in the spot market. This complicates the theoretical treatment of these markets as regards pricing and hedging.

3.2.1 Data

Due to the deregulated electricity market being a fairly new phenomenon, data of any useful size and quality has only recently become available. In this essay, I use daily (trading days) spot and futures prices (Figure 3.1) from Nord Pool for almost four years, January 2, 1996 to October 21, 1999[2]. From these daily prices (quoted in Norwegian kroner (NOK/MWh)) I calculate 946 daily returns (log-differences).

[1]Throughout this paper the "spot" price is treated as if it were a true spot price.

[2]The daily spot prices are so called "dygnspriser". These daily prices are the average prices of the 24 hourly quoted prices each day. The futures contracts are so called "veckokontrakt" for (hypothetical) delivery of a certain average amount of power over one week (vecka).

Each day a number of futures contracts with different maturities can be chosen as a hedge instrument and a decision of which contract to use must be taken. My intention is to study the short term hedging performance (one-week holding period), and since short term future contracts are more liquid as well as more correlated with the underlying spot prices than the longer term contracts, futures with three weeks left to maturity are chosen for the hedging investigations[3]. In order to avoid thin market and expiration effects, however, I roll over to the next three-week maturity contract one week prior to the expiration of the current contract. With this roll-over procedure, a time-series is created for the full time period 1996-1999[4]. A drawback from closing out the futures positions before expiration is that it introduces some basis risk to the hedge, since the future price is not directly tied to the underlying spot price prior to the maturity date[5]. This basis risk is particularly serious in the electricity market, where large temporary deviations between spot and future prices appear, due to non-straightforward arbitrage possibilities.

In Figure 3.2, I have plotted daily spot and futures return volatilities over the second half of the data set; GARCH volatilities (from the constant correlation bivariate GARCH model) as well as sample volatilities (annualized standard deviations). The volatilities are apparently very high compared to ordinary asset and currency markets. It can also be seen how the GARCH model successfully captures the swings in volatilities.

The volatility of both spot as well as futures returns clearly varies over time and the assumption of identically and independently normally distributed returns seems unrealistic. Further evidence of this is given by the investigation of the return

[3]In addition to these futures, I have also hedged spot movements with longer maturity futures (still with one-week holding period). In these cases, the hedging performance deteriorates compared to the shorter maturity futures case.

[4]In this way, the futures in my hedges always have a remaining lifetime of between 5 and 15 days.

[5]I already have basis risk due to the complexity associated with the closing of the futures; the futures are price contracts for an entire future week, not a single day. The basis risk due to this phenomenon should be partly eliminated by the roll over procedure though; no actual expiration is ever allowed.

distributions in Figure 3.3 and Figure 3.4; the return distributions are skewed and fat-tailed compared to the normal distribution. A time varying conditional variance might be one reason for the fat-tailed unconditional distributions.

According to Table 3.1, both the spot and the futures price series (logarithms) in this sample are stationary already on the level. The assumption of stationary electricity prices might seem unrealistic in the long run, but for the sample at hand, it cannot be rejected. Error correction due to cointegration is therefore not expected to improve the behavior of the GARCH models.

Table 3.2 reports some statistics on the return series. Both the futures returns and the spot returns have means not significantly different from zero. The unconditional distributions of spot returns and, in particular, futures returns are non-normal, as evidenced by skewness, high excess kurtosis, and highly significant Bera-Jarque statistics[6]. Tests for autocorrelation, using different Q-statistics, indicate that no autocorrelation is present in the spot market while some correlation exists at long lags in the futures market. Finally, a test for ARCH, using the Q^2-statistic, finds significant ARCH effects at all lags in the spot market and at long lags in the futures market.

3.3 Hedging Strategies

In this chapter, the minimum variance hedge ratio is estimated[7]. This hedge ratio determines how many futures contracts should be bought or sold for each spot contract for an investor to minimize the variance of his portfolio returns. There are many alternative ways of estimating this ratio, and this chapter looks at both unconditional and conditional estimates.

---

[6]The Bera-Jarque (B-J) statistic is χ^2_2-distributed under the null of normality. The statistic is $n \cdot [\frac{skewness^2}{6} + \frac{excess\ kurtosis^2}{24}]$,where n is the sample size.

[7]The minimum variance hedge ratio is also the optimal one (maximizing the mean-variance expected utility) if the futures prices either behave as martingales or the investors have infinite risk aversion. Neither of these conditions are expected to hold in the electricity market, however.

3.3.1 The Minimum Variance Hedge Ratio

To secure positions in a spot market, traders use futures as hedging instruments. For each spot contract, the hedge ratio tells us how many futures contracts should be purchased or sold. Let s_{t+1} and f_{t+1} be the changes in spot and futures prices, respectively, between time t and $t+1$, and let h_t be the hedge ratio at time t. Then,

$$x_{t+1} = s_{t+1} - h_t f_{t+1} \tag{1}$$

is the return to a trader going long in the spot market and going short in the futures market at time t. The variance of this return portfolio is

$$var_t(x_{t+1}) = var_t(s_{t+1}) + h_t^2 \cdot var_t(f_{t+1}) - 2 \cdot h_t \cdot cov_t(s_{t+1}, f_{t+1}), \tag{2}$$

and the minimum variance ratio, $h_{t,\min .var.}$, can then be derived by simply minimizing this variance with respect to h_t. We end up with the following expression for $h_{t,\min .var.}$:

$$h_{t,\min .var.} = \frac{cov_t(s_{t+1}, f_{t+1})}{var_t(f_{t+1})}. \tag{3}$$

In the particular case of electricity I invent a scenario where an electricity producer/distributor knows he is to sell electricity on the spot market on a particular day in the near future and wants to hedge with electricity futures. Since the aim is to study the short term hedging performance, a one-week investment horizon is assumed. For each spot position, the producer goes short in $h_{t,\min .var.}$ contracts in the futures market, where $h_{t,\min .var.}$ comes from (3). This operation is then repeated each day in the (out-of-sample) *test period*. As described below, $h_{t,\min .var.}$ is either modelled as constant throughout the test period or as time varying and updated on a daily basis.

3.3.2 Estimating the Hedge Ratio

The hedge ratio is estimated in five different ways; the naive one-to-one hedge ratio, where each spot contract is offset by exactly one futures contract, the OLS-hedge ratio where a regression of the spot returns on futures returns gives the hedge ratio expressed as in (3), but in an invariant unconditional version[8], a dynamic hedge ratio calculated by using continuously updated moving averages (50 days back in time) of variances and covariances and finally, two different dynamic hedge ratios based on bivariate GARCH modelling of the two return series (estimated by using daily data).

Several time series, including our spot and futures return series, exhibit periods of unusually large volatility followed by periods of tranquility. Under such circumstances, the assumption of a constant variance is obviously not appropriate. In order to capture the varying variance, the conditional variance can be modelled as a function of past errors as well as its own lags. This is done in GARCH models, and the first GARCH model to be estimated is the bivariate constant conditional correlation GARCH model introduced by Bollerslev (1990). The second choice is a multivariate GARCH model called Orthogonal GARCH, where the use of principal components analysis "diagonalizes" and simplifies the problem (Ding (1994), Alexander and Chibumba (1998) and Byström (1999)).

In the first GARCH model, spot returns and futures returns (daily returns) are modelled within the bivariate constant conditional correlation framework of Bollerslev (1990). The mean equations are specified as AR(2) processes and the conditional variance equations as GARCH(1,1):

$$
\begin{aligned}
y_{s,t} &= \alpha_{s,1} + \alpha_{s,2} y_{s,t-1} + \alpha_{s,3} y_{s,t-2} + \varepsilon_{s,t} \\
y_{f,t} &= \alpha_{f,1} + \alpha_{f,2} y_{f,t-1} + \alpha_{f,3} y_{f,t-2} + \varepsilon_{f,t}
\end{aligned}
$$

$$
\begin{aligned}
\sigma^2_{s,t} &= \phi_{s,1} + \phi_{s,2}\varepsilon^2_{s,t-1} + \phi_{s,3}\sigma^2_{s,t-1} \\
\sigma^2_{f,t} &= \phi_{f,1} + \phi_{f,2}\varepsilon^2_{f,t-1} + \phi_{f,3}\sigma^2_{f,t-1} \qquad (4) \\
\sigma_{sf,t} &= \rho\sigma_{s,t}\sigma_{f,t},
\end{aligned}
$$

where $\sigma^2_{s,t}$ and $\sigma^2_{f,t}$ are the conditional variances of $\varepsilon_{s,t}$ and $\varepsilon_{f,t}$, $\sigma_{sf,t}$ is the conditional

[8] Regression of futures returns on spot returns over the estimation period gives an OLS estimate of the optimal hedge ratio $h_{opt} = 0.408$, with a $standarderror = 0.049$. Daily returns, instead of weekly, are used since this is expected to improve the estimate (Duffie (1989)).

covariance between $\varepsilon_{s,t}$ and $\varepsilon_{f,t}$, and $\varepsilon_t = \sigma_t u_t$, where $u_t \sim N(0,1)$[9]. The time-varying conditional covariance between spot and futures returns are parametrized to be proportional to the product of the corresponding conditional standard deviations. This assumption greatly simplifies the computational burden in the estimation compared to more elaborate multivariate models. In the upper part of Table 3.3, estimation results from the Maximum Likelihood estimation (using the BHHH algorithm implemented in *GAUSS*) of the bivariate GARCH model are presented. It is shown how all GARCH parameters as well as the correlation coefficient are significantly different from zero. According to the statistics in Table 3.3, both the spot and futures residuals distribution approximate a standard normal distributions fairly well, even though some autocorrelation remains for the futures residuals. Some skewness and kurtosis remain as well.

The second model is the Ding (1994) and Alexander and Chibumba (1998) Orthogonal GARCH model. The main idea is to use principal components analysis to generate a number of orthogonal factors which can each be treated in a simple univariate GARCH framework.

There are *two* daily return series, spot and futures, with t observations represented as a $t \times 2$ matrix, $\mathbf{Y}_t$. The $t \times 2$ matrix, $\mathbf{P}_t$, of principal components is then defined as

$$\mathbf{P}_t = \mathbf{Y}_t \mathbf{W}_t \tag{5}$$

where $\mathbf{W}_t$ is the orthogonal 2×2 matrix of eigenvectors of $\mathbf{Y}_t^T \mathbf{Y}_t$ ordered according to the size of the corresponding eigenvalue. Notice that $\mathbf{P}_t$, just as $\mathbf{W}_t$, is now an orthogonal matrix. By inverting (5), one gets the principal components representation

[9] In this essay we limit ourselves to the standard GARCH model without asymmetrical extensions or dummies, and it is usually sufficient to limit the order of the GARCH model to (1,1). The same model has also been estimated with the conditional distribution for the error term modelled as the Student's t-distribution. No significant change in either the time-varying hedge ratio or the hedgeportfolio variance has been observed.

of the system:

$$\mathbf{Y}_t = \mathbf{P}_t \mathbf{W}_t^T.$$

Ω_t, the variance of $\mathbf{Y}_t$ at time t, can now be calculated as

$$\Omega_t = var(\mathbf{Y}_t) = var(\mathbf{P}_t \mathbf{W}_t^T) = \mathbf{W}_t \mathbf{D}_t \mathbf{W}_t^T, \tag{6}$$

where $\mathbf{D}_t$ is a diagonal matrix of principal component variances at t and where $\mathbf{W}_t$ is assumed to be known at time t[10]. This also enables us to calculate the forecasted covariance matrix $\Omega_{t+1} \mid \Psi_t$ by univariate methods, where Ψ_t is the information set at t; for each principal component, the conditional variance of the principal component i, $var_{t+1}(P_i \mid \Psi_t)$, can easily be forecasted by, for instance, any univariate GARCH model. This gives us the Orthogonal GARCH specification. This particular conditional covariance matrix also has the advantage of always being positive definite, since $\mathbf{D}_t$ is diagonal with positive elements along its diagonal.

In the lower part of Table 3.3, estimation results from the univariate GARCH estimates of the principal components are presented by using the previous AR(2)-GARCH(1,1) model. As for the constant correlation bivariate model, all GARCH parameters are significant, and the residuals are skewed and peaked, indicating (slightly) less than a perfect fit.

It is important to remember that Orthogonal GARCH, just like most other multivariate GARCH specifications, is based on certain assumptions. When assuming the conditional covariance matrix of the principal components to be diagonal, we foresee the fact that only the *unconditional* covariance matrix is diagonal. The *conditional* covariances need not be zero (Alexander and Chibumba (1998)), which could create some problems for the orthogonal GARCH methodology.

---

[10] $\mathbf{W}_t$ does not change very much from day to day and $\mathbf{W}_t$ can be approximated with $\mathbf{W}_{t-1}$, without introducing large errors in the calculation of the covariance matrix. This is particularly the case when calculating the *forecasted* covariance matrix. In this case, we forecast at time $t-1$, only using information up to $t-1$ (including $\mathbf{W}_{t-1}$).

3.4 Hedging Performance

The data set is divided into an *estimation period* of 473 days, January 2, 1996 to November 30, 1997, and an equally long 473 day *test period,* December 1, 1997 to October 21, 1999. Calculating weekly returns from daily data over the test period, test series with 469 *overlapping* weekly spot- and futures returns are constructed. From each of these series, I then construct five different test series with 93 *non-overlapping* weekly spot- and futures returns[11]. To further assess the stability of the results the test period is divided into three equal long subperiods; December 1, 1997 to July 23, 1998, July 24, 1998 to March 8, 1999, and March 9, 1999 to October 21, 1999. For the dynamic models, the (one-week) hedge ratio is updated each day in the *test period;* in the two GARCH cases, GARCH *forecasts* of weekly (5-days) covariances and variances by iteration of the variance equation are used (using an expanding sample when estimating the GARCH parameters). In this way, the hedge ratio variability is captured and we get the three series of varying hedge ratios pictured in Figure 3.5. These ratios are all stationary and the means of the different hedge ratios all close to the constant OLS hedge ratio[12].

3.4.1 Different Evaluation Approaches

The purpose of the entire hedging exercise has been to minimize the variance (uncertainty) of the hedge portfolio (which, in the case of an infinitely risk avert trader or Martingale futures, is equal to utility maximizing). To evaluate the hedging performance out-of-sample, I look at the *test period* and the test series defined above, covering approximately the months December 1997 to October 1999. Throughout the

[11]The series with 469 overlapping observations contain Monday to Monday weekly returns, Tuesday to Tuesday weekly returns, etc. Each of the five series with 93 non-overlapping returns contains returns from only one of the five working days, for instance Monday to Monday weekly returns. In this way, almost all autocorrelation found in the overlapping weekly return series is removed.

[12]The Phillip Perron values with and without trend are -71.7 and -71.9 for the BIGARCH hedge ratio, -110.4 and -111.5 for the OGARCH hedge ratio, and -16.13 and -16.09 for the 50-days MA hedge ratio. The mean hedge ratios are 0.43, 0.47, and 0.64, respectively.

essay, I evaluate out-of-sample and all hedge ratios are *predicted* hedge ratios using *predicted* variances and covariances.

I have chosen to evaluate to what extent the hedges minimize both the unconditional variance and the conditional variances over the test period. The unconditional variance is simply the sample variance of the (one-week) spot returns and hedge portfolio returns over the test period. In calculating the conditional variance I assume either the bivariate GARCH model or the Orthogonal GARCH model to be the underlying "data generating process", and each day I calculate (one-week) conditional spot and hedge portfolio variances that can be compared. The average conditional variances over the test period are then calculated and compared. The close fit of both the GARCH models to data, as demonstrated in Table 3.3 and in Figure 3.2, makes the assumption of data generated by either of these GARCH models plausible. The interesting information from this conditional evaluation is to what additional extent GARCH hedges reduce the conditional portfolio variance, in the presence of heteroscedasticity, compared to hedges that do not incorporate GARCH effects.

Both the unconditional variance approach, Kroner and Sultan (1993), and Park and Switzer (1995), as well as the conditional variance approach, Baillie and Myers (1991), Sephton (1993), and Bera, Garcia and Roh (1997), can be found in the literature.

3.4.2 The Unconditional Variance

This subsection studies the unconditional variance of the spot and portfolio returns by simply calculating the sample variance of the different return series over the test period as in Kroner and Sultan (1993) and Park and Switzer (1995). In Table 3.4, the variances and variance reductions (out-of-sample) of the different hedges compared to the open spot position are presented for the test series with overlapping returns. All hedges reduce the variance. Somewhat surprisingly, the best hedges are the unconditional naive hedge and the OLS hedge. All the dynamic hedge ratios perform worse than the static ones and there do not seem to be any major gains from modelling spot and futures returns on Nord Pool with time-varying volatilities[13]. The finding that the naive hedge performs equally well as (or even slightly better than) the OLS hedge was also found in the US stock index market by Park and Switzer (1995), and is an example of how simpler models sometimes work better than more elaborate ones.

In order to assess the significance of the results in a statistical sense, I turn to "bootstraping" techniques to get the distributions of the portfolio variance estimates. Bootstrapping return series with overlapping returns is possible by systematically picking (with replacement) groups of five (the dependences reach over five days) returns from the series until a new bootstrapped series of equal length as the original series is constructed (Shau and Tu (1995)). This procedure is repeated 1000 times and for each bootstrapped series, a new estimate of the unconditional return variance is found. In Table 3.5, means and standard deviations from these 1000 variance estimates are presented. It can be noticed how, as expected, the mean variances correspond closely to the actual sample variances in Table 3.4. From the size of the standard deviations, it can be concluded that even if the hedge portfolios have a smaller variance for our sample than the spot position, no hedge differs in a statistically significant way (at traditional significance levels) from the unhedged spot position. At the same time, no hedge significantly differs from any other hedge in its hedging performance.

[13]In the case of orthogonal GARCH, the weak performance could at least partly be due to a break down in the orthogonalization methodology.

The overlapping nature of the returns (in the test period) obviously induces dependences in the return series above. To study the effect of this on the relative hedge performance, Table 3.6 shows the same entries as Table 3.4, but now calculated for non-overlapping return series where almost all autocorrelation in the returns is removed. From the five different test series it can be seen how the conclusions drawn from Table 3.4 remain valid. All the different hedges systematically reduce the portfolio variance compared to the spot variance even though the reduction is much smaller for some series (for instance series 4) than for others[14]. Once more, even for the non-overlapping return series, the naive hedge and the OLS hedge perform better than the conditional hedges.

As a final variation of this theme, the 469 observation long test period is split into the three subperiods of equal length. As can be seen in Table 3.7, the hedging performance differs somewhat between the three subperiods but for all hedges and time periods, the hedged positions vary less than the unhedged spot position. For the special case of the Orthogonal GARCH hedge, it can clearly be seen how the performance deteriorates in period two; this is probably due to the high return variance in this period. The high volatility spills over to both hedge ratios and portfolio returns and, as mentioned above, the performance of the Orthogonal GARCH technique is particularly sensitive to highly volatile time periods.

It has been shown how the non-hedged spot position is, overall, more volatile than the hedge portfolios, indicating how hedging spot positions with futures contracts on the Nordic Power Exchange can be profitable for a variance-minimizing trader. To further evaluate the performance of the different hedges, I have chosen to look at another performance measure; how often the weekly portfolio return is actually smaller (in an absolute sense) than the weekly spot return. In Table 3.8, I count the number of times (out of 469) the absolute weekly return of the hedged positions is smaller than the absolute weekly spot return[15]. While the naive hedge now performs

[14]The Orthogonal GARCH hedge actually increases the variance for the fourth non-overlapping series.

[15]This is an alternative measure of the variability of the different portfolios. The large variance of

relatively worse, the two GARCH hedges perform equally well as the OLS hedge. Even though the moving average hedge is just slightly better than the naive hedge, it is clear how the conditional hedges tend to reduce the "variance" *as often* as the unconditional hedges, even if they, on average, do not remove *as much* of the variability as the unconditional hedges. It should also be noticed that the portfolio returns are actually larger than the spot returns in about 35 to 40% of the weeks.

So far, we can conclude that it has been possible to use futures hedges on Nord Pool to reduce the variance in the Nordic electricity market in the last two years, while the performance of the different hedge portfolios does not generally differ significantly from each other. The fairly weak performance of the more elaborate GARCH hedges compared to the simple OLS hedge might be explained by estimation problems or the fact that GARCH models do not always perform as well out-of-sample as in-sample. There are no indications of severe estimation problems, however, and the problems associated with forecasts from GARCH models estimated on daily data should be quite limited at such short horizons as one week. When it comes to Orthogonal GARCH, which has primarily been developed for large highly correlated multidimensional problems, the weak performance could partly be explained by a breakdown of the orthogonalization technique in periods of high volatility and asymmetrically behaving weakly correlated assets.

3.4.3 The Conditional Variance

Studying the reduction of the conditional instead of the unconditional variance gives somewhat different results. Assuming that the true return processes, and the conditional covariance matrix, are generated by either of the two GARCH models gives us the possibility to compare the relative performance of the data-generating GARCH model and the other models in minimizing the conditional variance (Baillie and Myers (1991), Sephton (1993) and Bera, Garcia and Roh (1997)). Hedging the spot position

---

a particular portfolio might be due to a small number of very large returns, even though most of the returns are smaller than the other portfolios. Such a portfolio should perform relatively better with this alternative variability measure.

each day in the test period and taking the average value of the conditional variance of the spot and hedge portfolio returns, provides a measure for each portfolio that should be as small as possible.

In Table 3.9, the (average) conditional variances and variance reductions (out-of-sample) of the different hedges compared to the open spot position are presented for two different choices of underlying covariance matrix[16]. The constant correlation GARCH model is (not unexpectedly) the best model when it is itself assumed to have generated the data. The OLS hedge is the second best, dominating the other GARCH model, Orthogonal GARCH. On the other hand, when the Orthogonal GARCH model is assumed to have generated the data it is also producing the best hedge. This time the constant correlation GARCH hedge performs better than the OLS hedge. The ranking of the very best hedges not surprisingly changes when changing the assumption of the covariance matrix, but those hedges that perform badly in one case, are also shown to perform badly in the other case. The best hedge in the unconditional variance evaluation, the naive hedge, is now instead the worst performer, and while the moving average model earlier performed as well as the GARCH models, it now barely reduces the variance at all. In Table 3.9, I assume, a priori, that return series follow GARCH models and the relative ranking of the different models is therefore of no major interest; the two GARCH models obviously dominate the other models. Instead, the focus should be on the absolute reduction in variance compared to the spot position, as well as the absolute performance of the GARCH models compared to the simpler models. From this, it should appear whether modelling the hedge ratio in this market with GARCH models is worth while. The answer to this question finally depends on how costly a dynamic GARCH hedge is, in terms of transaction costs etc., compared to the simpler static hedges.

As in the unconditional evaluation in the last subsection, we continue by looking at non-overlapping return series as well as subsamples of the whole test period. In

[16] In Tables 9, 10, and 11 the percentage reduction of the average (over the test periods) conditional variances is presented. Instead, looking at the average of the percentage reductions over the test periods gives exactly the same relative performance and relative ranking of the different hedge portfolios.

Tables 3.10 and 3.11, we see how the results from Table 3.9 remain stable both for the non-overlapping series and the different subsamples[17]; the performance of the different hedges changes very little from series to series in Table 3.10 and even if there is some change in performance from subsample to subsample in Table 3.11, this does not affect the relative ranking of the different hedges.

In a similar way as done in Table 3.8 for the unconditional evaluation, I now look at the number of times the conditional portfolio variance is smaller than the non-hedged spot position conditional variance over the test period. In Table 3.12, it is observed how the ranking from the earlier Tables 3.9, 3.10, and 3.11 remains unchanged, and that the GARCH model that is assumed to have generated the conditional covariance matrix *always (100%)* reduces the variance compared to the spot position.

When comparing the results from the conditional evaluation with the results from the unconditional evaluation, there are both similarities and differences. In both cases, and throughout the essay, it is quite obvious how hedging in this market can reduce the variance. However, while the naive model performed well in the unconditional evaluation context it is, by far, the worst performer in the conditional context. In some cases, it even increases the variance. The opposite holds for the two GARCH models; while not improving on the simple OLS hedge in the unconditional evaluation, the inclusion of heteroscedasticity and volatility clustering in calculating the hedge ratio clearly contributes towards an optimal hedge when looking at the conditional variance. The disappointing results from the last subsection are therefore reversed and the dynamic modelling of the hedge ratio seems to improve the hedging performance; when spot and futures returns can be modelled as GARCH processes, which is shown to be a plausible assumption in this market, then the hedge ratio should be modified and continuously updated according to these GARCH processes. It is hard to tell from the results in this chapter however, which GARCH model actually captures most of the variability in the market, since the relative ranking of the two GARCH models strongly depends upon the choice of the true covariance matrix.

---

[17]In order to save space, only the variance reductions, not the actual portfolio variances, are presented in Table 3.10.

3.5 Conclusions

In this essay, I have looked at a possible scenario of an actor on the Nordic Power Exchange that hedges a position (at a one-week horizon) in the spot market with futures contracts. The traditional unconditional methods of calculating the minimum variance hedge ratio are extended to time-varying moving average and GARCH hedge ratios. A constant correlation bivariate GARCH framework is compared to the new multivariate Orthogonal GARCH approach. Out-of-sample evidence presented in this essay indicates how both the traditional unconditional naive hedge, the unconditional OLS hedge, and the dynamic conditional GARCH hedges reduce the variance of the hedge portfolio compared to the spot position. The relative performance of the different hedges depends on the evaluation measure, however. The OLS hedge and the two GARCH hedges reduce both the unconditional and the conditional variance, while the naive hedge successfully reduces only the unconditional variance. Among the dynamical hedges, the moving average model is dominated by the GARCH models. Overall, there seem to be some gains from including heteroscedasticity and time-varying variances in the calculation of hedge ratios, but the constant OLS hedge ratio is nearly as successful in reducing the portfolio variance.

References

Alexander, C.O., and Chibumba, A.M (1998), "Orthogonal GARCH: An Empirical Validation on Equities, Foreign Exchange and Interest Rates, unpublished manuscript.

Baillie, R. T., and Myers, R. J. (1991), "Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge", *Journal of Applied Econometrics*, 6, 109-124.

Bera, A. K., Garcia, P., and Roh, J-S (1997), "Estimation of Time-Varying Hedge Ratios for Corn and Soybeans: BGARCH and Random Coefficient Approaches", OFOR Paper Number 97-06 December 1997.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.

Bollerslev, T. (1990), "Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model", *The Review of Economics and Statistics*, 31, 498-505.

Brailsford J. T., and Faff, R.W. (1996), "An Evaluation of Volatility Forecasting Techniques", *Journal of Banking and Finance* 20, 419-438.

Byström, H., (1999), "Orthogonal GARCH and Covariance Matrix Forecasting in a Stress Scenario: The Nordic Stock Markets During the Asian Financial Crisis 1997-1998", unpublished manuscript, Lund University.

Ding, Z., (1994), *Time Series Analysis of Speculative Returns*, PhD Thesis, UCSD.

Duffie, D. (1989), *Futures Markets*, Prentice-Hall.

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, 987-1007.

Figlewski, S. (1984), "Hedging Performance and Basis Risk in Stock Index Futures", *Journal of Finance*, 39, 657-669.

Kroner, K. F., and Sultan, J. (1993), "Time-varying Distributions and Dynamic Hedging with Foreign Currency Futures", *Journal of Financial and Quantitative Analysis*, 28, 535-551.

Markowitz, H. M., (1952), "Portfolio Selection", *Journal of Finance*, 7, 77-91.

Myers, R. J. (1991), "Estimating Time-varying Optimal Hedge Ratios on Future Markets", *Journal of Futures Markets*, 11, 39-53.

Park, T. H., and Switzer, L. N. (1995), "Time-varying Distributions and the Optimal Hedge Ratios for Stock Index Futures", *Applied Financial Economics*, 5, 131-137.

Pilipovic, D. (1998), *Energy Risk*, McGraw-Hill.

Sephton, P. S. (1993), "Optimal Hedge Ratios at the Winnipeg Commodity Exchange", *Canadian Journal of Economics*, 26, No. 1, 175-193.

Shau, J., and Tu, D. (1995), *The Jackknife and Bootstrap*, Springer.

Storset, S., (1995), "Risikostyring i Kraftmarkedet", *Derivatet*, No. 3.

Figure 3.1: Electricity Prices at the Nordic Power Exchange Nord Pool.

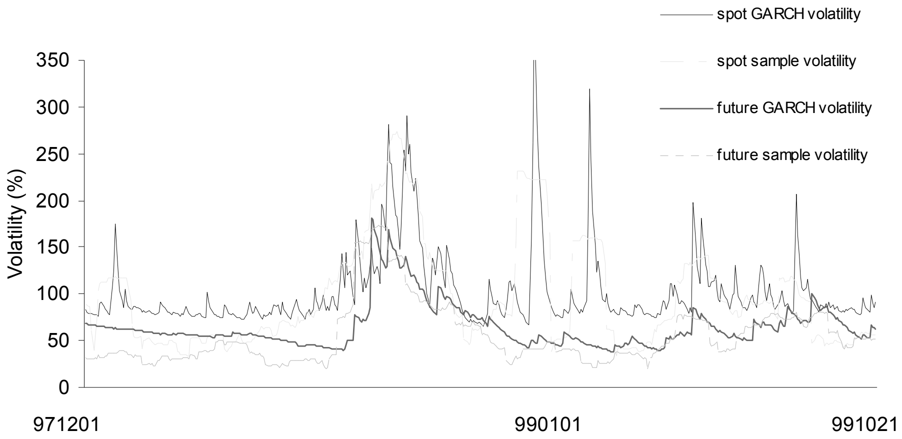

Figure 3.2: Spot Return Volatility and Futures Return Volatility. The GARCH volatility comes from the constant correlation bivariate GARCH model. The sample volatility is annualized in percent (20-day window).

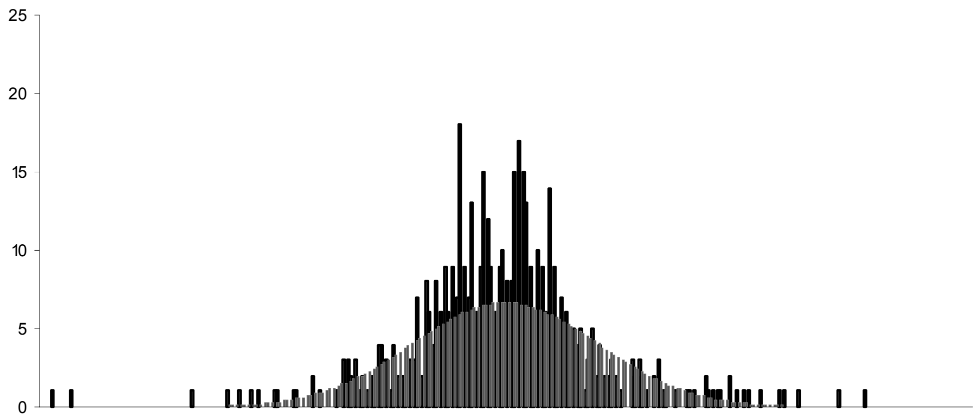

Figure 3.3: Spot Return Distribution vs. Normal Distribution.

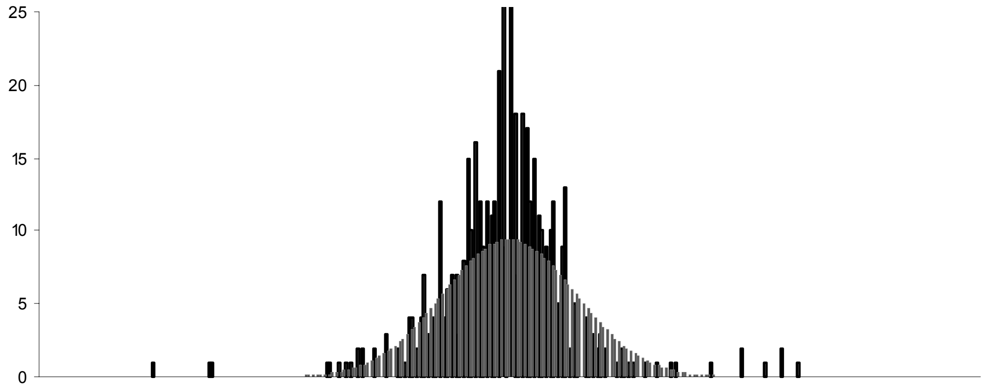

Figure 3.4: Future Return Distribution vs. Normal Distribution.

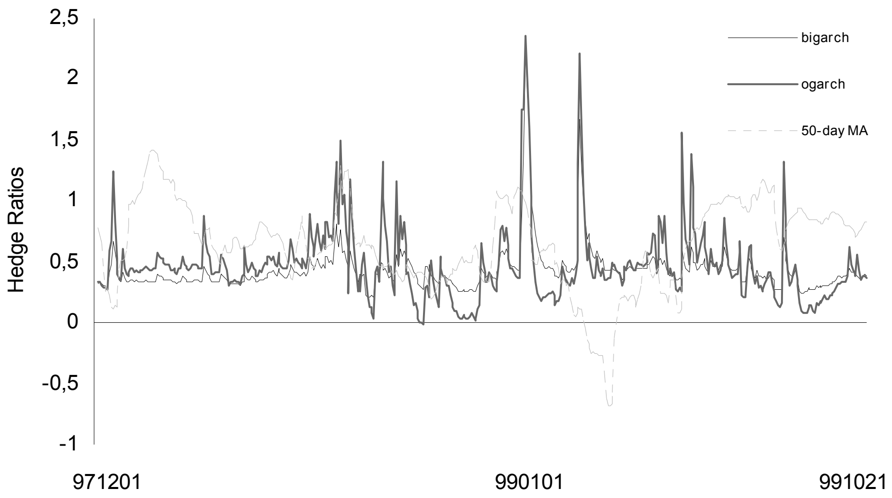

Figure 3.5: Conditional hedge ratios over the test period.

Table 3.1: Stationarity of logprices over the sample period January 2, 1996 to October 21, 1999.

	Phillips-Perron (no trend)	Phillips-Perron (with trend)
Spot prices	-38.97	-84.16
Future prices	-24.71	-42.12

The 99 percent critical values for the Phillips-Perron test with and without trend are -3.96 and -3.43.

Table 3.2: Return statistics January 2, 1996 to October 21, 1999.

	mean	variance	kurtosis	skewness	B-J	Q(6)	Q(18)	$Q^2(6)$	$Q^2(18)$
Spot	-0.000273 0.066	0.00429	6.698	-0.055	1768.83	6.92	20.56	146.16	175.20
Future	-0.000210 0.042	0.00175	9.211	1.004	3503.14	4.57	52.72	8.74	175.40

Small figures denote standard errors. Kurtosis is the excess kurtosis. B-J is the Bera-Jarque test for non-normality and Q(.) is the Ljung-Box test for autocorrelation. 99 percent critical value for Bera-Jarque is 9.21 and 99 percent critical values for Ljung-Box are 16.8, and 34.8.

Table 3.3: Typical GARCH parameter estimates. Maximum Likelihood estimates as well as standardized residual statistics.

	Bivariate GARCH	
	Spot	*Futures*
α_1	0.000464	-0.000128
	0.00185	0.00129
α_2	0.0750	-0.0661
	0.0392	0.0470
α_3	-0.0121	-0.0420
	0.0421	0.0471
ϕ_1	0.000299	0.0000570
	0.000173	0.0000341
ϕ_2	0.199	0.0758
	0.0305	0.0110
ϕ_3	0.736	0.898
	0.0284	0.0146
ρ	0.290	
	0.0270	
Mean	-0.0189	-0.0179
Standard Deviation	1.003	1.003
Skewness	-0.024	0.326
Excess Kurtosis	3.005	6.822
$Q(6)$	4.49	12.73

	Orthogonal GARCH	
	First Principal Component	*Second Principal Component*
α_1	-0.000254	0.000588
	0.00181	0.00127
α_2	0.104	-0.0846
	0.0411	0.0456
α_3	-0.00146	-0.0802
	0.0440	0.0461
ϕ_1	0.000442	0.0000461
	0.000269	0.0000278
ϕ_2	0.229	0.0637
	0.0388	0.0105
ϕ_3	0.681	0.910
	0.0418	0.0158
Mean	0.0159	-0.0020
Standard Deviation	1.001	0.999
Skewness	-0.127	0.050
Excess Kurtosis	2.982	7.361
$Q(6)$	8.05	18.98

The first 722 daily observations are used. Small figures are standard errors and Q(6) denotes the Ljung-Box test for the first six lags.

Table 3.4: Unconditional portfolio return variance and percentage variance reduction out-of-sample compared to the non-hedged spot position. Overlapping weekly portfolio returns.

	Spot	Naive	OLS	BIGARCH	OGARCH	50-days
Variance	0.0227	0.0193	0.0195	0.0201	0.0208	0.0200
% Variance Reduction	–	17.79	16.44	12.87	8.97	13.46

Table 3.5: Means and standard deviations of the bootstrapped (1000x) unconditional variance distributions. Overlapping weekly portfolio returns.

	Spot	Naive	OLS	BIGARCH	OGARCH	50-days
Mean Variance	0.0230	0.0193	0.0195	0.0201	0.0209	0.0203
Standard Deviation	0.0033	0.0035	0.0033	0.0034	0.0039	0.0034

Table 3.6: Unconditional portfolio return variance and percentage variance reduction out-of-sample compared to the non-hedged spot position. Non-overlapping weekly portfolio returns

		Spot	Naive	OLS	BIGARCH	OGARCH	50-days
	series 1	0.0169	0.0145	0.0140	0.0144	0.0149	0.0147
	series 2	0.0271	0.0191	0.0213	0.0218	0.0221	0.0210
Variance	series 3	0.0253	0.0227	0.0226	0.0229	0.0241	0.0229
	series 4	0.0270	0.0263	0.0252	0.0263	0.0274	0.0267
	series 5	0.0191	0.0154	0.0159	0.0163	0.0170	0.0163
	series 1	–	14.19	17.03	14.50	11.61	12.65
	series 2	–	29.27	21.15	19.40	18.49	22.50
% Var. Red.	series 3	–	10.30	10.63	9.58	4.82	9.41
	series 4	–	2.74	6.68	2.47	-1.64	0.95
	series 5	–	19.50	16.62	14.59	11.18	14.60

Series 1 represents Monday to Monday returns, series 2 represents Tuesday to Tuesday returns, etc. In this way, I work with non-overlapping returns with almost all autocorrelation removed.

Table 3.7: Unconditional portfolio return variance and percentage variance reduction out-of-sample compared to the non-hedged spot position. The three subperiods.

		Spot	Naive	OLS	BIGARCH	OGARCH	50-days
	subperiod 1	0.0118	0.0101	0.0102	0.0100	0.0102	0.0114
Variance	subperiod 2	0.0376	0.0372	0.0354	0.0370	0.0388	0.0374
	subperiod 3	0.0176	0.0106	0.0123	0.0125	0.0130	0.0113
	subperiod 1	–	17.25	15.58	18.03	15.98	3.89
% Var. Red.	subperiod 2	–	0.99	6.04	1.52	-3.09	0.43
	subperiod 3	–	66.91	43.08	40.60	36.01	55.73

The three subperiods correspond to the three time periods of equal length 1997-12-01 to 1998-07-23, 1998-07-24 to 1999-03-08, and 1999-03-09 to 1999-10-21.

Table 3.8: The number of times (out of the 469 weekly returns) that the hedge portfolios absolute weekly return is smaller (in an absolute sense) than the absolute weekly spot return.

	Naive	OLS	BIGARCH	OGARCH	50-days
number of times	280	306	306	305	281
% of the full sample	59.70	65.25	65.25	65.03	59.91

Table 3.9: Portfolio return conditional variance (average value over the test period) and percentage reduction of this average conditional variance out-of-sample compared to the non-hedged spot position. Overlapping conditional weekly portfolio variances.

	Spot	Naive	OLS	BIGARCH	OGARCH	50-days
Cond. Var., BIGARCH Cov. Matrix	0.0250	0.0268	0.0235	0.0232	0.0237	0.0250
Cond. Var., OGARCH Cov. Matrix	0.0247	0.0243	0.0223	0.0213	0.0208	0.0239
% Red., BIGARCH Cov. Matrix	–	-7.13	6.16	7.11	5.11	-0.07
% Red., OGARCH Cov. Matrix	–	1.11	9.65	13.72	15.82	3.06

Table 3.10: Percentage reduction of the conditional portfolio variance (average value over the test period) out-of-sample compared to the conditional spot variance. Non-overlapping conditional weekly portfolio variances.

		Naive	OLS	BIGARCH	OGARCH	50-days
	series 1	-7.16	6.17	7.11	5.53	0.17
	series 2	-7.32	6.32	7.10	5.34	-0.36
% Red., BIGARCH Cov. Matrix	series 3	-8.42	6.15	7.11	4.27	-0.54
	series 4	-6.93	6.23	7.09	5.22	-0.34
	series 5	-5.75	5.93	7.16	5.13	0.34
	series 1	-0.04	9.27	13.06	14.71	3.27
	series 2	-1.05	9.09	11.70	13.53	2.50
% Red., OGARCH Cov. Matrix	series 3	2.11	10.02	14.14	17.11	1.98
	series 4	1.03	9.74	12.86	14.84	2.72
	series 5	3.78	10.24	16.83	18.96	4.60

Table 3.11: Portfolio return conditional variance and percentage reduction of the conditional variance out-of-sample compared to the non-hedged spot position. The three subperiods.

		Spot	Naive	OLS	BIGARCH	OGARCH	50-days
	subperiod 1	0.0161	0.0175	0.0152	0.0152	0.0154	0.0170
Cond. Variance,	subperiod 2	0.0390	0.0412	0.0366	0.0360	0.0370	0.0386
BIGARCH Cov. Matrix	subperiod 3	0.0195	0.0212	0.0182	0.0181	0.0184	0.0191
	subperiod 1	0.0165	0.0156	0.0147	0.0145	0.0142	0.0160
Cond. Variance,	subperiod 2	0.0367	0.0356	0.0328	0.0303	0.0294	0.0355
OGARCH Cov. Matrix	subperiod 3	0.0204	0.0217	0.0190	0.0187	0.0184	0.0198
	subperiod 1	–	-8.38	5.54	5.82	4.19	-5.66
% Var. Reduction,	subperiod 2	–	-5.70	6.16	7.64	5.30	1.17
BIGARCH Cov. Matrix	subperiod 3	–	-9.07	6.70	7.09	5.50	2.02
	subperiod 1	–	5.74	11.02	11.95	13.73	3.01
% Var. Reduction,	subperiod 2	–	3.06	10.54	17.36	19.95	3.19
OGARCH Cov. Matrix	subperiod 3	–	-6.33	6.86	8.37	9.80	2.87

Table 3.12: The number of times that the conditional hedge portfolio variance is smaller than the conditional spot variance (out of the 469 weekly variances).

	Naive	OLS	BIGARCH	OGARCH	50-days
no. of times, BIGARCH Covariance Matrix	81	468	469	464	347
no. of times, OGARCH Covariance Matrix	177	416	428	469	322
% out of 469, BIGARCH Covariance Matrix	17.27	99.79	100.00	98.93	73.99
% out of 469, OGARCH Covariance Matrix	37.74	88.70	91.26	100.00	68.66

Chapter 4

The Search for Chaos and Other Nonlinearities in Swedish Stock Index Returns

4.1 Introduction

Common assumptions in financial economics are that financial variables like stock returns and exchange rates can be described by stochastic processes, and that economic systems are linear. From a theoretical point of view, nothing indicates that this must be the case, and numerous empirical studies have shown that different nonlinear dependences might exist in financial time series. The high frequency of crashes and booms that can be observed in stock markets is an example of behavior non-consistent with a linear model[1] with normally distributed disturbances.

Nonlinear models explaining this type of behavior can be both stochastic and deterministic. An example of a stochastic nonlinear model is Engle's (1982) autore-

[1]By a linear model, we mean a model with a linear mean and additive IID error terms.

gressive conditionally heteroscedastic (ARCH) model with extensions. These models have been successful in explaining volatility clustering in stock returns and in modelling interest rates. A chaotic process, on the other hand, is perfectly deterministic but its behavior is indistinguishable from pure randomness. Chaotic systems are always nonlinear and have been important in explaining seemingly random behavior in the sciences. A very important fact regarding chaotic models, which is different from the usual situation in economics, is that all movements are generated within the model. No external shocks need to be introduced.

The lack of explanative power in linear models has led numerous researchers to the study of chaos in financial economics, (Scheinkman and LeBaron (1989), Peters (1991), Larrain (1991)). Hsieh (1989, 1991) instead focuses on nonlinear stochastic models, in particular GARCH models. Varson and Doran (1995) try to distinguish chaos from random nonlinearities with the Grassberger-Procaccia correlation dimension. The disadvantages of this method are that very long data series are needed, and that it does not constitute a statistic test. Therefore, Brock, Dechert, and Scheinkman (1987) proposed a related statistic test, the BDS test, based on the correlation integral. This is a measure of spatial correlation, and it can be used to detect deviations from the IID-hypotheses. Since the BDS test does not distinguish between different causes for rejections of IID, Hsieh (1991) first filters data with different linear and nonlinear filters in order to detect what kind of model can explain the nonlinearities.

In this paper, we look at Swedish stock index returns, and try to detect the presence of nonlinearities in this market. Can nonlinear effects, for instance, help explaining the large jumps in stock prices that occur with fairly high frequency? Our data covers monthly returns from 1919 to 1996, daily data from 1977 to 1996, and intradaily data (15 minutes) from January 1992 to August 1993.

Since the movements cannot be successfully described by linear models, we try to determine how to model the nonlinearities that must be introduced. Typically, nonlinearities can enter in the process governing returns (nonlinearities in the conditional mean) or in the process of the time-varying conditional variance (nonlinearities in the conditional variance). In order to have a chaotic returns process, there must exist nonlinearities in the conditional mean. To test this hypothesis, we investigate the predictability of linear versus nonlinear models fitted to data by using neural networks. The weak evidence of nonlinearities in the conditional mean makes us continue our search for causes of the non-IID stock returns. We turn to the BDS test, and try to separate the different causes of a rejection of the IID-hypothesis in our market; chaos, heteroscedastic conditional variance, nonstationarity, or simple linear autocorrelation. We reject the hypothesis of conditional mean changes or nonstationarities causing the BDS test to reject the IID hypothesis. Instead, we find strong evidence of conditional variance dependences in the stock index returns. We do not only find significant GARCH effects but also the stronger result that GARCH effects alone contribute to more or less all the non-IID behavior in our data.

In chapter 4.2 we give an introduction to chaos, nonlinearities, and the BDS test, in chapter 4.3 our data is described, in chapter 4.4 we present our results, and chapter 4.5 concludes our paper.

4.2 Chaos and How to Detect It

What is chaos? A clear cut definition has not yet been generally accepted, but three fundamental properties must be included (see Strogatz (1996)):

- Chaotic motion must have an aperiodic long-term behavior. This means that there are trajectories[2] which do not settle down to fixed points, periodic orbits, or quasiperiodic orbits as $t \to \infty$.
- Chaotic motion exhibits sensitive dependence on initial conditions, which means that nearby trajectories diverge exponentially fast.
- Chaotic motion is purely deterministic. The irregular behavior stems from the nonlinearity of the system and not from random effects.

A dynamic system evolving in a chaotic way is also said to have a *strange attractor*. It is possible to quantify the sensitivity on the initial conditions above by defining the

[2]The phase space is the space spanned by the state variables. A trajectory is the path of evolution of a point in the phase space.

Liapunov exponents for chaotic systems in continuous time (flows), and discrete time (maps). For a flow in n dimensions, an infinitesimal sphere will be distorted into an infinitesimal ellipsoid during its evolution in phase space. If $\epsilon_k(t)$, $k = 1, ..., n$ denotes the length of the kth principal axes of the ellipsoid, then $\epsilon_k(t) \propto \epsilon_k(0)\, e^{\lambda_k t}$, where the λ_k are the Liapunov exponents. The n different Liapunov exponents therefore describe the deformation of the system. For a strange attractor, at least one of the λ_k must be positive because of the separation of neighboring trajectories. The sum of the λ_k describes the contraction of volume, which must be negative since, by definition, an attractor (strange or otherwise) should attract all trajectories starting in a sufficiently small open set containing the attractor. Chaotic behavior demands a positive Liapunov exponent, which implies the existence of a time horizon beyond which prediction breaks down, since any discrepancy in the estimate of an initial state will grow exponentially.

From where does chaos come? In the "definition" of chaos, two properties seem hard to combine. How can trajectories on a strange attractor diverge endlessly and yet remain bounded? The answer is that strange attractors result from a stretching and folding process. To be more concrete, consider again the sphere in the phase space. A strange attractor is generated when the system contracts the sphere in some directions and stretches it in others. To remain bounded, the distorted sphere must be folded back onto itself. After a large number of iterations, the sphere is spread throughout a bounded region in the phase plane. The mechanisms involved reflect the volume contraction, the sensitivity on initial conditions and the boundedness of the attractor.

4.2.1 The Correlation Integral

How do we test for the presence of chaos? One way would be to calculate the largest Liapunov exponent and see if it is positive, but this is not easily done in real world situations. Moreover, the other Liapunov exponents are even harder to estimate so this procedure cannot distinguish between different types of strange attractors. Grassberger and Procaccia (1983) therefore proposed a different procedure based on spatial correlation.

Consider a set $\{\mathbf{x}_t\}_1^T$ of points on an attractor. Define the *correlation integral* as $C(l) = \lim_{T\to\infty} C^T(l)$,

$$C^T(l) = \frac{1}{T^2}\sum_{s,t}\theta(l - |\mathbf{x}_t - \mathbf{x}_s|),$$

where $\theta(a) = 1$ if $a > 0$, and 0 otherwise. The quantity $C(l)$ measures the fraction of the total number of pairs $(\mathbf{x}_t, \mathbf{x}_s)$ whose distance $|\mathbf{x}_t - \mathbf{x}_s|$ is less than l. Grassberger and Procaccia found that $C(l)$, and its sample estimate $C^T(l)$, are proportional to l^d where d is called the *correlation dimension.* In practice, d is estimated as the slope in a plot of $\log C^T(l)$ versus $\log(l)$. When dealing with finite data sets, the power law, $C^T(l) \propto l^d$, only holds over an intermediate range of l, since at large l, all points will be within a distance of l, while on the other hand, at extremely small l, no pair of points will be within l. Typically, strange attractors have a fractal structure and therefore, a non-integer correlation dimension.

4.2.2 Attractor Reconstruction

A time series, $\{\mathbf{x}_t\}$, of all state variables $\mathbf{x} \in R^n$ is often not accessible. In many cases, it is not even possible to specify either the relevant components of $\mathbf{x}$ or its dimension, which might be high. Nevertheless, such systems might have low-dimensional attractors. Fortunately, Packard *et al.* (1980) and Takens (1981) show that the dynamics in the full phase space can be reconstructed from measurements of a single-variable time series $\{\{x_t\}; x_t \in R, x \subset \mathbf{x}\}$. The main idea is to construct m-dimensional vectors

$$\boldsymbol{\xi}_t = (x_{t,} x_{t+\tau}, ..., x_{t+(m-1)\tau})$$

for some delay $\tau > 0$. If $\{\mathbf{x}_t\}$ possesses an attractor that can be embedded in an n-dimensional $\mathbf{x}$-space, then the topological structure of the attractor remains unchanged when embedded in $\boldsymbol{\xi}$-space, provided that the *embedding dimension* m is large enough. A necessary condition is $m \geq n$ and a sufficient condition is $m \geq 2n+1$. In other words, it is possible to make some kind of "variable substitution" between the unobservable variables in $\mathbf{x}$ and the lagged observables.

Since $\{\mathbf{x}_t\}$ and $\{\boldsymbol{\xi}_t\}$ are observationally equivalent, they also have the same correlation dimension, provided that m is large enough. Thus, $C_m^T(l)$ can be calculated, now with a subscript m, using $\boldsymbol{\xi}_t$ instead of $\mathbf{x}_t$. If $\{\boldsymbol{\xi}_t\}$ is obtained from a chaotic time series, then the computed correlation dimension will level off at its true value when the embedding dimension is large enough, so that there is enough room for the attractor to unfold. The point to be made is that if $\{\boldsymbol{\xi}_t\}$ is obtained from a purely random (IID) sequence, the correlation dimension keeps increasing with m, since noise always fills up the space in which it is embedded. In principle, estimates of d can thus be used to distinguish chaos from noise.

The delay τ should not be chosen too small since then $x_t \approx x_{t+\tau} \approx ...$ and the attractor might have problems in disentangling. If, on the other hand, τ is chosen too large, problems arise because the spatial correlation is low for distant values in the time series.

Unfortunately, the Grassberger-Procaccia method has some drawbacks. First, it must be emphasized that the method breaks down when the embedding dimension is too large due to the sparsity of data, which causes statistical sampling problems.

Second, Grassberger and Procaccia based their results on time series of 10000-30000 points, although they argued that only a few thousand points would be necessary to obtain reasonable estimates. Ramsey and Yuan (1989) have shown that there is a tendency to underestimate the slope in data sets with as many as 2000 points, thus indicating chaos when none is present.

Third, the graphical procedure is not put on firm statistical ground and it might be difficult to interpret when noise is added to the system.

4.2.3 Influence of Noise

In real life, systems can seldom be isolated from the unpredictable environment. The observed dynamics will then consist of two parts; the intrinsic deterministic dynamics of the system and the influence of random noise. One common way of modelling the randomness is as *dynamical noise*:

$$x_t = f(x_{t-1}, ...) + \varepsilon_t, \tag{1}$$

where ε_t is IID, x_t is the observable, and f is a deterministic, possibly nonlinear, function. If f is chaotic, much of what has been said above about chaos is still valid when noise is added to the system, *provided the noise level is not too high.* Otherwise, the trajectories could jump out of the basin of attraction, or jump between different parts of the attractor in such a way that it becomes impossible to separate the behavior from true noise; the attractor has lost its structure. An acceptable noise level is hard to determine since different chaotic systems exhibit very different sensitivities to noise, but generally, only small amounts of noise can be introduced.

The Grassberger-Procaccia method must be modified when dealing with noise (see Ben-Mizrachi *et al.* (1984)). For length scales, l, below those where the random component blurs the structure, the slope of $\log C_m^T(l)$ versus $\log(l)$ is proportional to the embedding dimension, while for length scales above, the slope is equal to the correlation dimension of the deterministic system ($C_m^T(l) \propto l^d$).

Equation (1) can also be used to describe different stochastic processes. If, for instance, f is a nonlinear (or linear) map with a stable attracting fix point, then the output corresponds to passing random noise through a nonlinear (linear) recursive filter, thereby obtaining stationary aperiodic behavior. In principle, the Grassberger-Procaccia method could be used to identify the properties of f by calculating the slope at different l and m, and comparing it to the slope of an IID process of the same length and with the same moments, i.e. the time series *randomly permuted* or *scrambled.*

Figure 4.1 shows $\log(C(l))$ versus $\log(l)$ for the Hénon (1976) map

$$x_{k+1} = 1 + y_k - ax_k^2, \tag{2}$$

$$y_{k+1} = bx_k, \tag{3}$$

for $m = 2, 3$, and 4 with uniformly distributed noise dynamically added. In this case the attractor can be completely reconstructed when $m = 3$.[3] From Figure 4.1, it is clear that the slopes are different for length scales less than the magnitude of the

[3]The slope for $m = 2$ is only slightly less than the slopes in higher embedding dimensions, thereby suggesting that the attractor has a non-trivial structure already when $m = 2$.

imposed noise. Since the noise is uniformly distributed in $(-0.03, 0.03)$, the slopes for $l < 0.06$ ($\log(l) < -2.8$) are equal to the corresponding embedding dimensions. As expected, the correlation dimensions of the scrambled time series show no saturation as m grows. This graphical procedure should, in principle, also be valid for residuals from ARCH-type models[4], where $\varepsilon_t = g(x_{t-1}, ...)u_t, u_t$ IID, but it would be hard to tell if any deviation from the scrambled counterpart is caused by the conditional mean process f or the conditional variance process g or both (see Scheinkman and LeBaron (1989)).

Brock, Dechert, and Scheinkman (1987) modified the Grassberger-Procaccia method in order to circumvent some of its drawbacks by developing the BDS test. A revised version of this method is found in Brock *et al.* (1996).

4.2.4 The BDS Test

The test is based on the observation that for an IID sample

$$C_m^T(l) = [C_1^T(l)]^m.$$

The identity should be understood in a statistical sense. Brock, Dechert, and Scheinkman (BDS) derived a normalization factor[5] $V_m^T(l)$ in order to make a correct statistical quantification of the departure from IID. More specifically, they showed that the BDS statistic

$$W_m^T(l) = \sqrt{T}\frac{\left[C_m^T(l) - [C_1^T(l)]^m\right]}{V_m^T(l)}$$

converges in the distribution to $N(0,1)$ as $T \to \infty$, for $l > 0$ and $m > 1$, under the null hypothesis of IID.

[4] By ARCH-type models, we mean models that exhibit autoregressive conditional heteroscedasticity.

[5] The normalization factor is a complicated function of correlation integrals in different dimensions. It is not very illuminating and therefore not presented in this paper, but it can be found in Brock *et al.* (1996).

The BDS test has proved to be quite powerful in finding departures from IID in a number of Monte Carlo simulations; see, for example, Brock, Hsieh, and LeBaron (1991) and Hsieh (1991). It has also been verified that the asymptotic distribution well approximates the finite sample distribution for sample sizes above 1000 observations. As for the Grassberger-Procaccia method, the results depend on the magnitude of the additive noise component. It is therefore recommended (see Brock, Hsieh, and LeBaron (1991)) to apply the test for $l = 0.5 - 1.5$ times the sample standard deviation of the time series.

The BDS test suffers from the fact that a rejection of the null does not provide any hints about the cause of the rejection. It has been shown in Brock and Potter (1993) and de Lima (1996), however, that the null distribution of the test is not affected by applying the BDS test to the estimated residuals from a general class of parametric models with additive IID errors:

$$x_t = f(x_{t-1}, ...; \boldsymbol{\alpha}) + u_t,$$

where $\boldsymbol{\alpha}$ is a parameter vector and u_t is IID, *provided* that a $\sqrt{T}$-consistency estimation of the parameters is possible. The last requirement means that $\sqrt{T}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \rightarrow N(\mathbf{0}, \Sigma)$ for some covariance matrix Σ, and is fulfilled in maximum likelihood estimations. Moreover, de Lima (1996) shows that the nuisance parameter free properties of the BDS test remain valid for residuals from some multiplicative models

$$x_t = f(x_{t-1}, ...; \boldsymbol{\alpha}) + g(x_{t-1}, ...; \boldsymbol{\beta}) u_t, \tag{4}$$

where u_t is IID, which covers many, if not all, of the known ARCH-type models, *if* the test is applied to the transformed residuals $\hat{v}_t = \ln(\hat{u}_t^2)$ where $\hat{u}_t$ are the estimated residuals in (4). The key idea is that the transformation gives rise to a model with additive IID errors:

$$\tilde{x}_t = \ln\left((x_t - f(x_{t-1}, ...; \boldsymbol{\alpha}))^2\right) = \ln\left(g^2(x_{t-1}, ...; \boldsymbol{\beta})\right) + \ln(u_t^2) = \tilde{f}(x_{t-1}, ...; \boldsymbol{\beta}) + v_t,$$

where v_t is IID because u_t is IID. It must be kept in mind that the asymptotic properties of $\hat{v}_t$ and v_t are equal only if $\boldsymbol{\alpha}$ is known, i.e. when a GARCH model is fitted to the data without the conditional mean process. However, the bias introduced is expected to be small when $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are jointly estimated "...because of the lack of predictability of the first moment compared to the second central moment..." (see Brock and Potter (1993)), as is often the case in financial time series[6].

Furthermore, the BDS test does not require any moments for the time series studied to exist, contrary to other nonlinearity tests, but when the test is applied to estimated residuals $\sqrt{T}$-consistency of the parameters sometimes demand the existence of higher moments in the disturbance process, typically a finite variance. The assumption of the existence of higher order moments might otherwise be severe in the nonlinearity testing of financial time series, whose distributions are often heavily tailed. The disadvantage is that the BDS test may require longer samples to be as effective as some other tests.

Altogether, the BDS test may not only be useful in detecting chaos, but also in diagnostic testing of estimated model residuals.

4.3 Data

The data we use is log-returns from two different Swedish stock indices, the OMX-Index and the Affärsvärldens Generalindex. The former was created as an underlying security for trading in standardized stock index options and forward contracts, and consists of a value-weighted combination of the 30 most traded stocks at the Stockholm Stock Exchange. The latter is the most used stock index in Sweden and it is a value-weighted index of the majority of all stocks quoted on the Stockholm Stock Exchange.

The time interval between observations is sometimes of importance in financial time series. Therefore, we look at series with different sampling frequency; monthly, daily, and 15 minutes.[7] In this way, we can detect chaotic and nonlinear behavior on

[6]With the possibility of bootstrapping small sample distributions, the asymptotic properties of the test are nowadays of less importance.

[7]The 15-minute time span between data points should be large enough that possible micromarket structure dependences do not shine through. Such dependences can, for instance, arise from the sequential execution of orders in the traders' limit order book, or by thin trading in some of the

different time scales. In addition, the problem of too short or too long delay times τ in the reconstruction of the chaotic attractor mentioned in chapter 4.2.2 can be mitigated.

The monthly data (Affärsvärldens Generalindex) is extended over the period 1919-1996, while the daily data (OMX-Index) is for the period 1984-1996. The intradaily series (OMX-Index), with data collected every 15 minutes, covers the period January 1992-August 1993 and is divided into three series with approximately equal length to catch any instability in the estimated parameters.[8] At the same time, the length of the series becomes comparable to the length of the daily series. The first observation every day has been removed due to the auction-like start up procedure at the exchange. All data is without dividends, but we have also looked at monthly (1919-1996) and daily (1977-1991) returns with dividends included (Affärsvärldens Generalindex).[9] The results for these series are not reported but are similar to those without dividends and any differences will be commented upon. Table 4.1 displays some sample statistics of the return data. A high degree of kurtosis and skewness, in particular for the 15 minute data is observed.

4.4 Empirical Results

In the following chapter, we apply the BDS test to the raw data to test if the different time series are IID. If the null hypothesis of IID is rejected (and it is), we try to discover the cause of the rejection by applying the BDS test to the estimated residuals from different models. More precisely, we examine whether the rejection origins from nonlinear (possibly chaotic) or linear dependences in the mean processes, nonstationarities, or if the rejection can be explained by nonlinear stochastic models exhibiting

underlying stocks.

[8]The second subperiod covers the volatile months when Sweden abandoned the fixed exchange rate regime.

[9]For monthly returns, the series with dividends has 936 observations and for the daily returns, 3489 observations.

conditional heteroscedasticity.

4.4.1 Test of the Raw Data

Table 4.2 shows the BDS statistics when the embedding dimension is between 2 and 10, and l is chosen as 0.5, 1 and 1.5 times the sample standard deviation of the different time series. As can be observed, the test strongly rejects the hypothesis of IID stock returns at any conventional significance level and for all m and l, even though little can be said about the cause of the rejection. If the time series came from a noisy chaotic system with a higher dimensional attractor ($m \leq 10$), then the test statistic would typically be small for low embedding dimensions and large in dimensions with enough "room" for the attractor. Obviously, this is not the case since the rejections are significant already at $m = 2$.

Considering the results, we proceed to study how dependences in the conditional mean might cause the rejection of IID.

4.4.2 Test of Dependences in the Conditional Mean

Suppose stock returns are generated as:

$$x_t = f(x_{t-1}, ...) + \varepsilon_t \tag{5}$$

where ε_t is IID and f is the conditional mean process. The expression in (5) can be used to describe noisy chaotic models as well as different stochastic processes. Now, we will investigate whether f is nonlinear (and possibly chaotic), and whether any dependences in the conditional mean can explain the rejection of IID.

If f is nonlinear, it can be modelled by nonparametric regressions, e.g. artificial neural networks. This powerful technique is exhaustively described elsewhere (see Hertz *et al.* (1991) and Kuan and White (1994)), and thus, we will here merely say that a multi layer perceptron (MLP), with one hidden layer of sigmoid transfer functions, can consistently model a time series generated by an arbitrary function, provided that the function is bounded and uniformly continuous. On the other hand, if the MLP only contains linear transfer functions, it can only pick up linear dependences, i.e. the nonparametric modelling reduces to a linear regression. We therefore proceed in the following way:

We split each time series into three parts of unequal size. The first 50% are used for the regression, or training, while the consecutive 25% are used as validation sets. The last 25% ($t = 1, ..., T_{test}$) are used as test sets for out-of-sample evaluation. For each returns series, we train 10 MLPs with 3, 5 and 7 hidden neurons, each with sigmoid (than) transfer functions, in a single hidden layer and a linear output neuron. As inputs we use $x_{t-1}, ..., x_{t-9}$ corresponding to an embedding dimension $m = 10$. For each time series, we also train linear perceptrons to capture the linear dependences. In all cases, the sum squared errors of the validation sets are minimized, and for each time series, we pick the network with the best validation performance. The different MLPs are then evaluated using the out-of-sample prediction error normalized with the out-of-sample variance of x_t:

$$E_p = \frac{\frac{1}{T_{test}} \sum_{t=1}^{T_{test}} \left(x_t - \hat{f}\left(x_{t-1}, ..., x_{t-9}\right) \right)^2}{\sigma_{test}^2}.$$

If f is nonlinear, E_p should be smaller for nonlinear than for linear networks . Furthermore, if f is chaotic, then the performance of nonlinear networks should improve as the forecast horizon shortens, that is, when we go from a sample interval of one month down to 15 minutes, and it is a consequence of the positive Liapunov exponent(s) in a chaotic system.

The results are reported in Table 4.3. Three things can be noted. First, the out-of-sample prediction performance is very weak, with E_p slightly below one (and occasionally above). Second, there is no support in the data for nonlinear regression being superior to linear regression. Sometimes, it is superior but then the improvement is negligible. Third, the performance of the nonlinear networks does not improve as the sample interval is reduced.

We also analyze the dependence of the output on the input variables by inspecting derivatives after completed regression. This is done by calculating

$$S_k(t) = \frac{\partial \hat{f}(x_{t-1}, ..., x_{t-9})}{\partial x_{t-k}}, \qquad k = 1, ..., 9$$

for $t = 1, ..., T_{test}$ and then computing

$$S_k = \frac{1}{T_{test}} \sum_t \mid S_k(t) \mid . \tag{6}$$

The measure in (6) shows which variables are the most important. For a linear network, this is just equal to (the absolute values of) the coefficients from a linear regression.

The results are found in Table 4.4. We see that both types of networks identify approximately the same variables as being important, often the first. The magnitudes of S_k differ somewhat between linear and nonlinear networks, but the point is that there seem to be no substantial dependences in the data not captured by the linear networks.

Although this is not a statistical test of nonlinearities in the conditional mean process, we draw the conclusion that there are none. Either the dependences in f are truly linear, or they are so blurred with noise that the conditional mean can best be approximated by a linear process. Since chaos can only arise from a nonlinear mapping, this suggests that there is no chaotic behavior in the mean process.

Can the linear dependences in f cause the BDS test to reject IID? We try to answer this question by applying the BDS test to the residuals from the linear networks above. As appears from Table 4.5, the test strongly rejects the residuals as being IID at any reasonable significance level. The test statistics of the linearly filtered data do not differ a great deal from those of the raw data, indicating that the rejection is not only due to linear dependences[10].

In the following, we try to sort out what kind of nonlinearities the BDS test in fact detects in our data.

[10]We also applied the BDS test to the residuals from the nonlinear networks. As expected, the test statistics did not differ a great deal from those in Table 4.5, but to our knowledge, it is not clear whether the nuisance free properties of the test is valid for residuals from nonparametric regressions.

4.4.3 Test of Nonstationarities

Financial time series are not necessarily stationary. There might be a number of structural changes creating nonstationary series: policy changes, changes in financial structures as well as important technological innovations. Therefore, it is possible that a rejection of IID by the BDS test is caused by nonstationarities and a closer study of whether this is the case is motivated. The type of nonstationarities we discuss in this section arise from switching between different linear models; that is, we allow the parameters of a linear model to change from regime to regime.

To investigate the influence of nonstationarity, we can either apply the BDS test to the residuals from a linear switching model, or we can try to judge it more qualitatively (following Hsieh (1991)). Since we have time series with different extensions in time, we do the latter. Our longest return series is 77 years and our shortest is about 7 months. In between, we have a time series of about 13 years. When we proceed to increasingly shorter time intervals, we expect the effect of structural changes to disappear (we simply assume that less and less changes occur) and consequently, any rejection of the IID hypothesis due to nonstationarities alone should disappear.

There is no support for this idea. The linearly filtered statistics in Table 4.5 are as high for the 15 minute data as for the monthly and daily data. As long as the linear regime shifts occur with a frequency low enough to leave our 15 minute data unaffected, they cannot be the only cause of the BDS test rejecting the IID hypothesis.

4.4.4 Test of Dependences in the Conditional Variance

As mentioned, there are many ways of generating non-IID data; we have treated linear and nonlinear autoregression, chaos, and nonstationarity. Now, we turn to models exhibiting nonlinearities in the conditional variance. Examples of models with this behavior are autoregressive conditionally heteroscedastic models like ARCH, GARCH, and their extensions.

Many time series exhibit periods of unusually large volatility followed by periods of tranquility. Under such circumstances, the assumption of a constant variance is obviously inappropriate. In order to detect the varying variance, the conditional variance can be modeled as a function of past errors. This is the ARCH model. Further extensions of the model, allowing the conditional variance to also be a function of its own lags, gives us Bollerslev's (1986) generalized ARCH model, GARCH(p, q), where p and q are the number of lagged conditional variance and error components, respectively. In the case of an AR(2) process with the conditional variance modeled as GARCH(1,1) we have:

$$
\begin{aligned}
x_t &= \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \varepsilon_t \\
\sigma_t^2 &= \phi_0 + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2,
\end{aligned}
$$

where σ_t^2 is the conditional variance of ε_t, $\varepsilon_t = \sigma_t u_t$, and $u_t \sim N(0,1)$. Most empirical studies suggest that p and q larger than one are rarely needed. One problem with GARCH modelling is that the parameters are restricted; ϕ_0 must be larger than zero, ϕ_1 and ϕ_2 must be zero or larger, and the sum of ϕ_1 and ϕ_2 must be less than one in order to have a finite unconditional variance. The sum $\phi_1 + \phi_2$ measures the persistence of conditional variance to shocks, which approaches infinity as the sum approaches one from below. The case where the sum of ϕ_1 and ϕ_2 equals one is referred to as Integrated in GARCH or IGARCH.

We proceed one more step, this time using an asymmetric specification for the conditional variance process, by extending the standard ARCH-model to Nelson's (1991) EGARCH-model:

$$
\ln \sigma_t^2 = \phi_0 + \phi_1 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \phi_2 \left(\left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - E\left[\left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| \right] \right) + \theta \ln(\sigma_{t-1}^2)
$$

where as above σ_t^2 is the conditional variance of ε_t and $\varepsilon_t \mid \varepsilon_{t-1} \sim N(0, \sigma_t)$. Unlike simple ARCH and GARCH, EGARCH can capture the asymmetric response of the variance to the direction of ε_t, that is, a higher variance when ε_t is negative, and a lower variance when ε_t is positive. In this model, $\phi_1 < 0$ represents the asymmetric effect while $\phi_2 > 0$ produces the ARCH effect. The θ-parameter determines the persistence in variance. Unlike the GARCH model, there are no non-negativity restrictions on the parameters in the EGARCH model.

A crucial assumption in the models above is the normality assumption on the standardized error term u_t; the conditional heteroscedasticity of ε_t alone is expected to explain the observed kurtosis in the return distribution. Since empirical evidence strongly rejects the idea that financial returns are normally distributed, we compare estimates with normally distributed errors with t-distributed ones. In this way, a larger part of the excess kurtosis in the stock returns might be captured.[11]

Baillie and DeGennaro (1990) as well as Poon and Taylor (1992) clearly demonstrate that the t-distribution gives a better fit to the error term of financial return series. For the t-distribution one additional parameter must be estimated, the degree of freedom, υ. The t-distribution is symmetric, but has fatter tails and a higher kurtosis than the normal distribution. When υ goes to infinity, the t-distribution approaches a normal distribution.

GARCH Estimation

Below, we estimate the parameters in different GARCH models. EGARCH models have also been estimated but the results are not presented since these are very similar to those of the GARCH models. For each time series, a number of different GARCH and EGARCH models have been estimated, and the models with the best fit have been chosen, on basis of parsimony, likelihood value, and behavior of the standardized residuals.

For the t-distribution, one problem that has earlier been found for the 15-minute data is the low degrees of freedom (below 4), which means that fourth-order moments do not exist (Hansson and Hördahl (1994)). One possible explanation for these results is that the density for the 15 minute returns has very fat tails, driving the degrees of freedom of the t-distribution to low levels. In fact, these problems are solved by adding a daily overnight dummy to the 15-minute series in order to account for the increased variance in overnight returns.[12] Using a process like the following for the

[11]For the EGARCH model, one must be aware of the fact that the conditional variance expression changes with the distribution; $E\left[\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right|\right]$ depends on the chosen distribution. For the normal distribution $E\left[\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right|\right]=\sqrt{\frac{2}{\pi}}$, and for the t-distribution, our calculations give $E\left[\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right|\right]=\frac{\Gamma\left(\frac{\upsilon-1}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)}\frac{\sqrt{\upsilon}}{\sqrt{\pi}}$. For large values of υ (over 30), this expression is very close to $\sqrt{\frac{2}{\pi}}$.

[12]It would also be straightforward to include a weekend dummy but its impact should be smaller

conditional variance in the GARCH model,

$$\sigma_t^2 = \phi_D D + \phi_0 + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2$$

where D is the dummy variable with parameter ϕ_D, we capture a considerable part of the kurtosis and get well-behaved residuals.[13]

The maximum likelihood estimates (the BHHH algorithm) of GARCH models with normally distributed, as well as t-distributed, errors are presented in Tables 4.6 and 4.7. For all our models and return series, we get significant parameter estimates, except for α_0 for the 15-minute returns. For the monthly data, we cannot reject the null of an infinite unconditional variance (we cannot reject the null of an IGARCH in these GARCH models) even though intuition suggests a mean-reverting variance. This conflicting evidence might be reconciled by allowing for fractional orders of integration (Bollerslev and Mikkelsen (1996)). Including AR(2) terms in the conditional mean gives better fit to data and less (linearly) correlated standardized residuals for some data series. In general, the coefficients of the AR(2) terms are very small, however. The estimates of the degrees of freedom parameter v for the t-distribution are clearly finite, indicating non-normally distributed errors. The degrees of freedom parameters differ between time series but are similar for the GARCH and EGARCH models, respectively. Overall, the parameter estimates are fairly non-sensitive to the specification of the error distribution.

For our daily OMX-Index data, Hansson and Hördahl (1997) have fit EGARCH models where the errors are described by the normal distribution as well as the generalized error distribution (GED). A comparison of our estimates with the results from Hansson and Hördahl (1997) shows almost identical estimates for ϕ_1, ϕ_2 and θ when the errors are assumed to be normal. The same holds for our parameters in the t-distributed model, when compared to the estimated GED models by Hansson and

---

than the overnight dummy; the number of nontraded days each weekend is much smaller than the number of nontraded 15-minute periods overnight.

[13]In the EGARCH-case, a dummy is introduced in a similar manner; $\ln \sigma_t^2 = \ln(1 + \phi_D D) + \phi_0 + \ldots$

Hördahl.[14]

In practice, all models retain some skewness in the residuals, a finding common in empirical studies of conditionally heteroscedastic models. A comparison of the values of kurtosis between the models shows that, as expected, the t-distributed model has residuals with higher kurtosis than the normally distributed one. While the normally distributed standardized residuals show excess kurtosis for all data series, the t-distributed standardized residuals show a lower than theoretically predicted kurtosis of about the same size for all but the daily series.

BDS Statistics on Transformed Standardized Residuals

With a comparison of BDS statistics for the linearly filtered financial data and for the GARCH residuals, the extent to which the rejection of IID residuals are due to GARCH effects is captured. As discussed in chapter 4.2.4, we circumvent the problem in GARCH filtering related to the issue of nuisance parameters by looking at the transformed standardized residuals. In Table 4.8 and Table 4.9, we present the BDS statistics for the transformed standardized residuals from the fitted GARCH models in Table 4.6 and Table 4.7, respectively.

The most obvious result is the substantial decrease in rejecting frequency after GARCH filtering[15]. We cannot reject the null hypothesis for monthly and daily data with normal errors. Some test statistics are outside the often used 95% confidence intervals, but that is not unexpected with as many as 27 test statistics for each return series. We still reject IID returns for the 15-minute data in period 2 and the overall period. For the third period, one statistic is very large. We should reject IID for this period, but the asymptotic properties of the test might not be applicable for m as high as 10, with these sample sizes. A proper bootstrapping may clear these doubts,

---

[14]It is important to note that, as Nelson (1991) shows, EGARCH models with a t-distributed u_t give rise to strictly stationary although not covariance stationary σ_t^2 and ε_t; i.e. σ_t^2 and ε_t have no finite unconditional moments.

[15]The BDS test is asymptotically distributed as $N(0,1)$ and considering our long data series, we assume the same holds in finite samples. Any rejection mentioned is at the 5% significance level. For a two-sided test at this level, the critical values are ± 1.96.

however.

For the t-distributed residuals, we have no rejection for monthly and daily data, and a slightly better fit of 15-minute returns, but we still have problems with period 2. It seems that the data series contain linear or nonlinear dependences that we have not succeeded in catching. Treating overnight effects with only one dummy might be too simple, at least for this period.

There might be another explanation, however. Table 4.8 and Table 4.9 clearly suggest that there are GARCH effects in all our data, although the specification for period 2 is not correct. As mentioned above, this period includes a highly volatile part in the Swedish financial history when the fixed exchange rate regime was abandoned. In chapter 4.4.3, we concluded that *linear* regime shifts could not explain the rejection of IID in our returns, but we cannot exclude *nonlinear* regime shifts in the volatility. Such nonlinearities may be captured by the Markov switching model, SWARCH, where a Markov-chain governs the transition probabilities between the volatility states (Hamilton and Susmel (1994)). In the SWARCH model, the conditional variance is described by an ARCH model which is the same in the different regimes, but the scale of the conditional variance differs across regimes. Moreover, the regimes are solely identified from the data in the estimation procedure and involves no subjective classifications other than the number of volatility states.

It might be argued that if we have regime shifts in our 15-minute data, then the residuals from ordinary GARCH models for the monthly and daily data covering that period should not be IID. However, nonstationarities detected in one time scale are not necessarily detectable in another. For example, a shift to a higher volatility regime for some months would probably not be detectable in the analysis of monthly returns, but would most likely be detected when analyzing 15-minute returns. Some indications of nonstationarites are indeed observed in our parameters; in Tables 4.6 and 4.7 the variation in parameter estimates for the 15 minute return sample and its subsamples can be noted.

4.5 Conclusions

We have searched for nonlinearities and chaos in Swedish stock index returns. We find strong evidence of nonlinear behavior, but no evidence of low-dimensional chaos.

We use the BDS test to detect deviations from IID errors in time series. This test detects both deterministic and stochastic dependences, and lacks some of the disadvantages of the more commonly used Grassberger-Procaccia method.

Applying this test to Swedish stock index returns of different sampling frequencies (monthly, daily, 15 minutes) strongly rejects the IID hypothesis. Since many possible dependences in the data, linear as well as nonlinear, may be the cause of this rejection, we apply the test to residuals from different models.

Linear dependences do not explain much of the deviation from IID. The same applies to nonstationarity; IID is not more strongly rejected for low frequency data. In order to detect chaos in stock indices, we compare the predictability of nonlinear and linear neural networks. No significant improvement is found, giving low importance to nonlinearities in mean or chaotic dependences in the time series. However, this does not exclude the possibility of chaos in individual security returns, since a combination (like in a stock index) of chaotic time series might very well lose its chaotic structure, as examined by Atchison and White (1996).

Instead, heteroscedastic conditional variance models are found to explain much of the rejection of IID. The BDS statistics decrease substantially when GARCH or EGARCH models with normal or t-distributed errors are fitted to the time series. In particular, the t-distributed GARCH model explains the rejection of IID for monthly and daily data, while no model fully explains the nonlinearities found for the 15-minute data, but there is little doubt that the answer to a better fit lies in the specification of the conditional variance. If the rejections are caused by (not too many) regime shifts in the volatility process, these might be successfully captured by SWARCH models.

We find our results quite robust and nonsensitive to the chosen type of GARCH model, and the inclusion of dividends or not. Our paper increases the strength of the already strong support for GARCH-type models of financial returns. For future research, we suggest searching for nonlinearities in other financial variables like implicit option volatilities, interest rates, and currencies. Analyzing nonlinear dependences both in time series and panel data might be of interest. The availability of transaction data also gives the possibility to study nonlinear effects in the micromarket structure.

References

Atchison, M., and White, M. (1996), "Disappearing Evidence of Chaos in Security Returns: A Simulation", *Quaterly Journal of Business and Economics*, 35(2), 21-37.

Baillie, R.T., and DeGennaro, R.P. (1990), "Stock Returns and Volatility", *Journal of Financial and Quantitative Analysis*, 25, 203-214.

Ben-Mizrachi, A., Grassberger, P., and Proccacia, I. (1984), "Characterization of Experimental (noisy) Strange Attractors", *Physical Review A* 29 (February), 975.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, 31, 307-327.

Bollerslev, T., and Mikkelsen, H. O. (1996), "Modelling and Pricing Long Memory in Stock Market Volatility", *Journal of Econometrics* 73, 151-184.

Brock, W., Dechert, W., and Scheinkmann, J. (1987), "A Test for Independence based on the Correlation Dimension", University of Wisconsin at Madison, University of Houston, and University of Chicago.

Brock, W., Dechert, W., Scheinkmann, J., and LeBaron, B. (1996), "A Test for Independence Based on the Correlation Dimension", *Econometric Reviews*, 15(3), 197-235.

Brock, W., Hsieh, D., and LeBaron, B. (1991), *Nonlinear Dynamics, Chaos and Instability: Theory and Evidence*, M.I.T. Press, Cambridge.

Brock, W., and Potter, S. (1993), "Nonlinear Time Series and Macroeconometrics", in *Handbook of Statistics Volume 11: Econometrics*, ed. by Maddala, G., Rao, C., and Vinod, H., 195-229.

de Lima, P.F. (1996), "Nuisance Parameter Free Properties of Correlation Integral Based Statistics", *Econometric Reviews*, 15(3), 237-259.

Engle, R. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica* 50, 987-1007.

Grassberger, P., and Procaccia, I. (1983), "Measuring the Strangeness of Strange Attractors", *Physica D* 9, 189.

Hamilton, J., and Susmel, R. (1994), "Autoregressive Conditional Heteroscedasticity and Changes in Regime", *Journal of Econometrics* 64, 307-333.

Hansson, B., and Hördahl, P. (1994), "Stock Returns and Conditional Volatility: An Empirical Investigation of the Swedish Stock Market", Working Paper, Department of Economics, University of Lund.

Hansson, B., and Hördahl, P. (1997), "Forecasting Variance Using Stochastic Volatility and GARCH" in Hördahl (1997).

Hénon, M. (1976), "A Two-dimensional Mapping of a Strange Attractor", *Commun. Math. Phys.* 50, 69.

Hertz, J., Krogh, A., and Palmer, R. (1991), *Introduction to the Theory of Neural Computation,* Addison-Wesley, Redwood City, Ca.

Hsieh, D. (1989), "Testing for Nonlinear Dependence in Daily Foreign Exchange Rates", *Journal of Business*, vol. 62, no. 3, 339-368.

Hsieh, D. (1991), "Chaos and Nonlinear Dynamics: Application to Financial Markets", *The Journal of Finance*, vol. XLVI, no.5, 1839-1877.

Hördahl, P. (1997), "Financial Volatility and Time-Varying Risk Premia", Ph.D.-Thesis, Department of Economics, University of Lund.

Kuan, M. and White, H. (1994), "Artficial Neural Networks: An Econometric Perspective", *Econometric Reviews*, 13(1), 1-91.

Larrain, M. (1991), "Testing Chaos and Nonlinearities in T-Bill Rates", *Financial Analysts Journal*, September-October, 51-62.

Nelson, D. (1991), "Conditional Heteroscedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347-370.

Packard, N., Crutchfield, J., Farmer, J., and Shaw, R. (1980), "Geometry from a Time Series", *Phys. Rev. Lett.* 45, 712.

Peters, E. (1991), "A Chaotic Attractor for the S&P 500", *Financial Analysts Journal*, March-April, 55-62.

Poon, S-H, and Taylor, S.J. (1992), "Stock Returns and Volatility: An Empirical Study of the UK Stock Market", *Journal of Banking and Finance*, 16, 37-59.

Ramsey, J., and Yuan, H. (1989), "Bias and Error Bias in Dimension Calculation and their Evaluation in Some Simple Models", *Physical Letters A* 134, 287-297.

Scheinkmann, J.A., and LeBaron, B. (1989), "Nonlinear Dynamics and Stock Returns", *Journal of Business*, vol.62, no.3, 311-337.

Strogatz, S.H. (1996), *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering,* Addison-Wesley, Reading, MA.

Takens, F. (1981), "Detecting Strange Attractors in Turbulence", *Lect. Notes in Math.,* 898, 396.

Varson, P.L., and Doran, P. (1995) "The Search for Evidence of Chaos in FTSE-100 Daily Returns", *European Financial Management*, vol.1, no.2, 201-210.

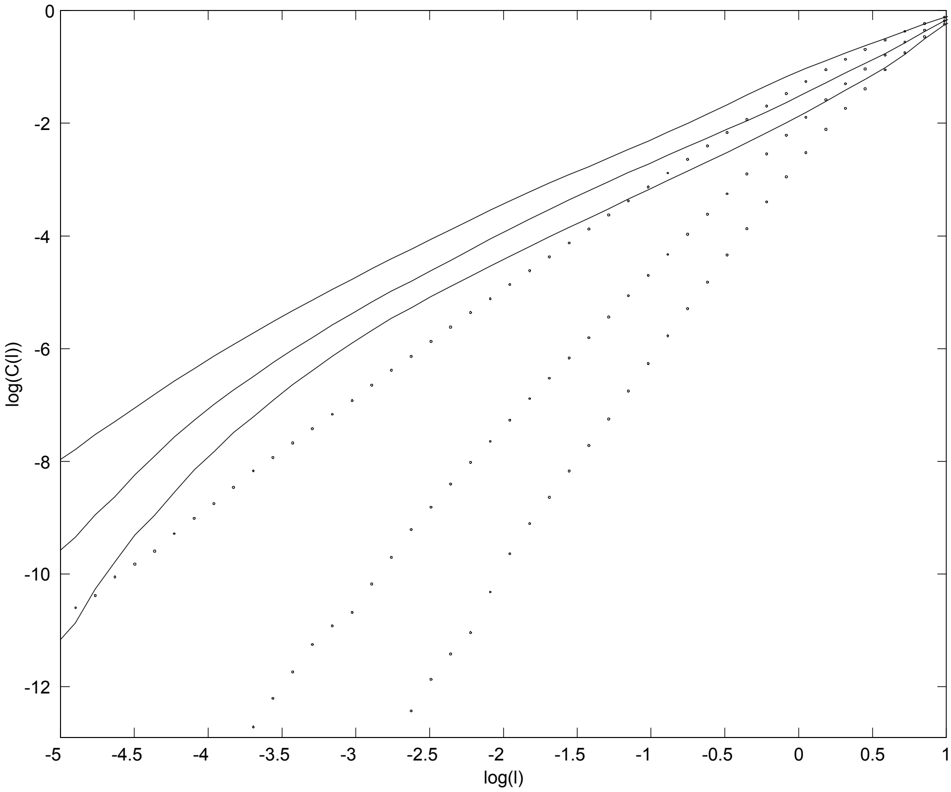

Figure 4.1: Plot of $\log(C(l))$ vs. $\log(l)$ for the Hénon map with $a = 1.4$ and $b = 0.3$ (thick lines) and its scrambled counterpart (dotted), for $m = 2, 3$, and 4. Higher lying lines correspond to lower m. Noise uniformly distributed on $(-0.03, 0.03)$ is dynamically added to the x-component. The changes in slopes for $l < 0.06$ ($\log(l) < -2.8$) are clearly visible. The time series is normalized to unit sample variance, and consists of 2000 observations.

Table 4.1: Sample statistics of the return data.

	monthly	daily	15 min. 1	15 min. 2	15 min. 3	15 min. all
nr. of observations	936	2949	2579	2571	2767	7917
Mean ($\cdot 10^2$)	0.49	0.023	-0.0015	0.0072	0.0088	0.0049
Variance ($\cdot 10^4$)	21.61	0.30	0.03	0.13	0.036	0.06
Skewness	-0.52	-0.18	1.05	4.53	0.27	4.36
Kurtosis	8.04	9.30	32.72	96.43	9.58	130.11

Table 4.2: BDS statistics for the raw data.

		m								
	$\frac{l}{\sigma}$	2	3	4	5	6	7	8	9	10
monthly	0.5	8.33	10.83	13.57	16.61	20.13	26.49	33.56	48.21	67.21
	1	8.43	10.74	13.39	15.72	17.81	20.54	23.64	27.34	32.20
	1.5	7.70	9.83	12.04	13.79	15.20	16.73	18.24	19.76	21.49
daily	0.5	12.36	14.66	17.55	21.05	24.82	30.05	36.82	46.94	59.21
	1	12.53	14.40	16.35	18.53	20.49	22.82	25.29	28.17	31.87
	1.5	13.25	15.12	16.73	18.13	19.21	20.34	21.47	22.56	23.86
15 min. 1	0.5	11.43	13.13	12.82	13.58	13.81	13.81	13.67	13.67	13.00
	1	12.76	13.53	12.88	12.61	12.25	11.73	11.13	10.42	9.51
	1.5	11.80	11.79	10.89	10.19	9.48	8.82	8.21	7.48	6.72
15 min. 2	0.5	16.71	19.11	20.87	22.47	24.03	25.38	26.34	27.23	27.37
	1	14.34	15.67	15.78	15.74	15.32	14.78	14.01	13.25	12.36
	1.5	9.51	10.01	9.67	9.19	8.53	7.85	7.16	6.48	5.81
15 min. 3	0.5	8.89	11.16	11.93	13.39	15.15	16.79	18.13	19.29	20.77
	1	8.36	10.59	11.23	12.40	13.48	14.27	14.79	14.95	15.14
	1.5	7.49	8.98	9.35	10.11	10.72	11.00	11.05	10.85	10.54
15 min. all	0.5	24.98	29.87	32.63	36.31	40.49	44.78	49.12	53.58	58.02
	1	24.09	6.89	27.75	28.68	29.28	29.57	29.65	29.52	29.21
	1.5	20.12	21.57	21.39	21.29	20.92	20.40	19.79	19.13	18.43

l is the length scale, σ is the sample standard deviation, and m is the embedding dimension. The test statistic is asymptotically distributed as $N(0, 1)$.

Table 4.3: Out-of-sample evaluation for nonlinear and linear neural networks.

	E_p			
	MLP3	MLP5	MLP7	MLPL
monthly	1.0512	1.0418	1.1635	1.0357
daily	0.9913	1.0132	1.0046	1.0108
15 min. 1	0.9903	0.9871	1.0127	0.9852
15 min. 2	0.9942	1.0452	1.0015	0.9943
15 min. 3	0.9926	0.9869	1.0029	0.9820
15 min. all	0.9901	0.9925	0.9909	0.9895

MLP3 refers to a MLP with 3 hidden neurons etc. MLPL refers to a linear perceptron.

Table 4.4: Sensitivity dependences in the nonparametric regressions.

		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
monthly	MLP	0.1153	0.1208	0.0360	0.0929	0.0409	0.1091	0.0263	0.0515	0.1228
	MLPL	0.1634	0.1131	0.0123	0.0230	0.0001	0.1681	0.0129	0.1049	0.1217
daily	MLP	0.1217	0.0171	0.0462	0.0249	0.0107	0.0574	0.0475	0.0325	0.0242
	MLPL	0.1843	0.0124	0.0569	0.0106	0.0210	0.0659	0.0898	0.0780	0.0281
15 min. 1	MLP	0.0537	0.0507	0.0511	0.0381	0.0191	0.0046	0.0206	0.0183	0.0105
	MLPL	0.0812	0.0365	0.0816	0.0189	0.0292	0.0069	0.0053	0.0111	0.0157
15 min. 2	MLP	0.0939	0.0499	0.0223	0.0023	0.0665	0.0572	0.0040	0.0147	0.0219
	MLPL	0.0714	0.0486	0.0298	0.0014	0.0105	0.0195	0.0381	0.0251	0.0224
15 min. 3	MLP	0.0341	0.0262	0.0102	0.0055	0.0134	0.0231	0.0068	0.0156	0.0056
	MLPL	0.0585	0.0095	0.0153	0.0270	0.0039	0.0003	0.0109	0.0062	0.0216
15 min. all	MLP	0.1330	0.0385	0.0091	0.0325	0.0070	0.0227	0.0132	0.0062	0.0303
	MLPL	0.1371	0.0393	0.0053	0.0348	0.0054	0.0162	0.0025	0.0181	0.0314

MLP refers to a nonlinear multi layer perceptron. MLPL refers to a linear perceptron.

Table 4.5: BDS statistics for linearly filtered data.

		m								
	$\frac{l}{\sigma}$	2	3	4	5	6	7	8	9	10
monthly	0.5	7.28	10.09	12.59	15.04	17.70	22.11	26.18	31.35	36.66
	1	7.04	9.77	12.55	14.86	16.87	19.46	22.36	25.85	30.39
	1.5	6.50	8.97	11.30	13.19	14.65	16.22	17.73	19.26	20.97
daily	0.5	11.83	14.41	17.04	20.54	24.45	30.48	37.75	47.93	58.89
	1	12.18	14.27	16.16	18.52	20.69	23.26	26.03	29.21	33.28
	1.5	12.94	15.05	16.68	18.25	19.44	20.65	21.87	23.09	24.53
15 min. 1	0.5	10.58	12.21	11.98	12.60	12.90	12.77	12.48	12.10	11.14
	1	11.27	11.99	11.46	11.18	10.86	10.34	9.74	8.99	8.08
	1.5	10.46	10.34	9.64	8.97	8.32	7.69	7.12	6.41	5.69
15 min. 2	0.5	14.34	16.87	18.52	19.83	21.12	22.25	22.78	23.33	23.34
	1	14.27	15.45	15.53	15.51	15.16	14.67	13.95	13.19	12.27
	1.5	9.69	9.99	9.61	9.08	8.42	7.78	7.12	6.44	5.74
15 min. 3	0.5	7.99	10.21	10.80	12.02	13.49	15.04	16.65	17.92	19.78
	1	7.53	9.83	10.52	11.69	12.72	13.44	13.93	14.05	14.18
	1.5	6.67	8.37	8.87	9.76	10.42	10.73	10.77	10.56	10.23
15 min. all	0.5	22.42	27.25	30.09	33.43	37.12	40.88	44.46	48.19	52.13
	1	22.52	25.29	26.25	27.13	27.70	27.95	27.96	27.78	27.42
	1.5	19.56	20.82	20.62	20.49	20.12	19.62	19.01	18.34	17.59

l is the length scale, σ is the sample standard deviation, and m is the embedding dimension. The test statistic is asymptotically distributed as $N(0,1)$.

Table 4.6: GARCH with normally distributed errors.

	$\alpha_0 \cdot 10^5$	α_1	α_2	$\phi_0 \cdot 10^5$	ϕ_1	ϕ_2	$\phi_D \cdot 10^5$
monthly*	0.01	0.16	-0.088	5.87	0.17	0.81	
	0.002	0.035	0.034	1.830	0.024	0.025	
daily	33.20	0.20		0.14	0.11	0.83	
	8.710	0.019		0.022	0.008	0.012	
15 min. period 1	-1.36	0.21		0.07	0.10	0.27	1.89
	2.23	0.021		0.0024	0.013	0.017	0.119
15 min. period 2	4.55	0.26		0.19	0.20	0.17	13.75
	3.640	0.019		0.0068	0.021	0.013	0.452
15 min. period 3	4.43	0.14		0.12	0.15	0.34	1.70
	2.993	0.022		0.0063	0.017	0.025	0.163
15 min. all	1.98	0.21		0.13	0.24	0.17	6.35
	1.650	0.012		0.002	0.013	0.0073	0.099

* indicates that the null hypothesis of an IGARCH cannot be rejected. Small figures indicate standard errors.

Table 4.7: GARCH with t-distributed errors.

	$\alpha_0 \cdot 10^5$	α_1	α_2	$\phi_0 \cdot 10^5$	ϕ_1	ϕ_2	$\phi_D \cdot 10^5$	v
monthly*	0.01	0.16	-0.078	6.60	0.16	0.82		9.13
	0.002	0.035	0.035	2.56	0.031	0.033		2.580
daily	34.00	0.17	-0.037	0.09	0.11	0.86		7.28
	7.490	0.019	0.019	0.020	0.015	0.018		0.630
15 min. period 1	-1.28	0.17		0.06	0.10	0.34	1.54	4.90
	1.910	0.019		0.005	0.024	0.033	0.190	0.455
15 min. period 2	2.24	0.24		0.16	0.21	0.26	7.77	4.66
	3.246	0.019		0.013	0.034	0.027	0.925	0.438
15 min. period 3	5.10	0.11		0.11	0.16	0.35	2.07	4.62
	2.656	0.019		0.011	0.031	0.042	0.158	0.438
15 min. all	1.35	0.17		0.10	0.25	0.31	3.07	4.02
	1.436	0.011		0.006	0.023	0.018	0.235	0.196

* indicates that the null hypothesis of an IGARCH cannot be rejected. Small figures indicate standard errors.

Table 4.8: BDS statistics for transformed GARCH residuals, normal distribution.

		m								
	$\frac{l}{\sigma}$	2	3	4	5	6	7	8	9	10
monthly	0.5	-1.13	-1.10	-0.98	-0.79	-0.43	0.04	0.35	-0.17	-2.24
	1	-0.69	-0.46	-0.28	-0.15	-0.13	-0.21	-0.20	-0.27	-0.13
	1.5	-0.49	-0.22	-0.08	-0.07	-0.06	-0.24	-0.29	-0.46	-0.46
daily	0.5	0.82	1.18	1.37	1.74	2.01	1.84	1.41	0.60	0.13
	1	0.50	0.65	0.80	1.29	1.38	1.29	1.09	0.97	0.92
	1.5	0.67	0.37	0.41	0.87	0.86	0.71	0.57	0.50	0.45
15 min. 1	0.5	0.19	-0.26	-0.48	-0.45	-0.25	-0.11	0.02	1.04	1.38
	1	-0.52	-0.44	-0.29	-0.08	0.21	0.55	0.91	1.28	1.46
	1.5	-0.89	-0.36	0.05	0.37	0.80	1.20	1.59	1.96	2.16
15 min. 2	0.5	1.30	1.47	1.38	1.59	1.69	1.61	0.80	-0.52	-2.69
	1	1.51	1.37	1.33	1.52	1.60	1.80	1.95	2.14	2.00
	1.5	1.83	1.86	1.89	1.97	2.06	2.22	2.31	2.43	2.39
15 min. 3	0.5	-0.44	-0.93	-1.04	-0.68	-0.71	-0.43	-0.87	-1.49	-2.94
	1	-0.27	-0.71	-0.98	-0.72	-0.57	-0.55	-0.59	-0.60	-0.59
	1.5	-0.39	-0.90	-0.97	-0.67	-0.49	-0.55	-0.59	-0.62	-0.61
15 min. all	0.5	0.40	0.03	0.01	0.45	0.70	0.90	0.95	0.83	0.84
	1	0.27	0.06	0.10	0.63	1.01	1.40	1.66	1.91	2.04
	1.5	0.38	0.30	0.56	1.03	1.43	1.80	2.11	2.38	2.51

l is the length scale, σ is the sample standard deviation, and m is the embedding dimension. The test statistic is asymptotically distributed as $N(0, 1)$.

Table 4.9: BDS statistics for transformed GARCH residuals, t-distribution.

		m								
	$\frac{l}{\sigma}$	2	3	4	5	6	7	8	9	10
monthly	0.5	-1.14	-1.27	-1.16	-0.94	-0.68	-0.38	-0.43	-0.81	-1.31
	1	-0.73	-0.49	-0.28	-0.12	-0.11	-0.20	-0.19	-0.25	-0.19
	1.5	-0.49	-0.20	-0.03	-0.04	-0.06	-0.26	-0.32	-0.51	-0.53
daily	0.5	0.72	1.12	1.22	1.47	1.76	1.57	0.93	0.28	0.27
	1	0.55	0.68	0.84	1.21	1.23	1.13	0.97	0.84	0.73
	1.5	0.72	0.55	0.65	1.04	0.94	0.79	0.65	0.60	0.59
15 min. 1	0.5	0.27	0.02	-0.07	-0.00	0.16	0.53	1.04	1.64	2.77
	1	-0.39	-0.38	-0.21	-0.04	0.30	0.54	0.77	0.99	1.13
	1.5	-0.76	-0.28	0.16	0.48	0.99	1.36	1.65	1.89	2.02
15 min. 2	0.5	1.27	1.26	1.18	1.21	1.20	0.72	-0.14	-0.75	-2.45
	1	1.38	1.19	1.27	1.48	1.64	1.86	2.04	2.22	2.18
	1.5	2.00	1.97	2.06	2.16	2.31	2.57	2.72	2.89	2.88
15 min. 3	0.5	-0.58	-0.93	-1.03	-0.82	-0.56	-0.28	-0.62	-0.80	-1.69
	1	-0.34	-0.74	-1.03	-0.83	-0.70	-0.68	-0.74	-0.80	-0.78
	1.5	-0.57	-0.97	-1.07	-0.85	-0.70	-0.75	-0.80	-0.85	-0.83
15 min. all	0.5	0.35	-0.03	-0.15	0.15	0.28	0.51	0.71	1.07	1.47
	1	0.32	0.02	-0.00	0.44	0.88	1.25	1.47	1.67	1.75
	1.5	0.64	0.40	0.55	0.93	1.37	1.72	1.93	2.13	2.19

l is the length scale, σ is the sample standard deviation, and m is the embedding dimension. The test statistic is asymptotically distributed as $N(0,1)$.

Chapter 5

The Compass Rose Pattern of the Stock Market: How Does it Affect Parameter Estimates, Forecasts, and Statistical Tests?

5.1 Introduction

Typically stock prices, as well as numerous other financial time series, move in small-integer multiples of a minimum "tick size". This discrete nature of stock prices restricts stock returns to take on a limited number of values only, a restriction which is one of the necessary conditions for creating the so-called "compass rose" pattern, a geometrical pattern in a scatter plot of returns versus lagged returns. Crack and Ledoit (1996) were the first to discover this pattern in return plots, finding compass rose patterns in all stocks they investigated on NYSE. Crack and Ledoit describe the compass rose pattern as "subjective" and impossible to use for predictive purposes. Contrary to the beliefs of Crack and Ledoit, Chen (1997) demonstrates some evidence of how the compass rose of a stock can be used to improve stock return forecasts. Crack and Ledoit also suggest that the discrete nature of stock prices might affect GARCH estimates as well as statistical tests. So far, no studies of discrete prices and GARCH estimates have been made but Krämer and Runde (1997) show how the compass rose seriously distorts the null distribution of the BDS test (Brock *et al.* (1996)). The BDS test is a widely used statistical portmanteau test detecting deviations from IID in general and the existence of chaos in particular. Based on their own findings, Krämer and Runde suggest that this test should be used with caution whenever discrete data is used. However, we believe their return series to be highly unrealistic and less suitable for drawing conclusions.

There is more to the compass rose than meets the eye and in this paper, we use Monte Carlo simulations to test whether the existence of a compass rose and the associated discreteness affect estimates, forecasts, and correlation integral based statistical tests. Possible effects on parameter estimates are investigated as is the use of the compass rose to enhance forecasts in a GARCH framework (as suggested by Chen) which is tested and further developed in a more theoretical framework. We also test how the BDS test and the associated Savit and Green (1991) dependability index are influenced by the discreteness of stock prices. When researchers look for evidence of chaos in the stock market, they often study stock indexes and not individual stocks. However, as shown by Atchinson and White (1996), an aggregation of chaotic processes may very well be non-chaotic. This turns the focus to the study of individual stocks and therefore, it is important to clarify whether correlation integral based tests can be used when examining stock returns with a discrete nature and compass rose patterns.

Chapter 5.2 describes the compass rose. In chapter 5.3 we investigate how rounding affects AR-GARCH parameter estimates. Chapter 5.4 deals with our compass rose enhanced forecasts of stock returns. Chapter 5.5 shows how correlation based statistical tests and their distributions can be affected by discrete prices and the compass rose. Chapter 5.6 concludes the paper.

5.2 The Compass Rose

A compass rose pattern sometimes appears when returns from a financial time series are plotted against lagged returns. The pattern is characterized by several evenly spaced lines radiating from the origin of the graph; the thickest lines pointing in the major directions of a compass. Crack and Ledoit (1996) were the first to recognize and explain this phenomenon, induced by discreteness in stock price data. They found that the compass rose appears clearly if the stock in question satisfies three conditions:

1. Daily stock price changes are small relative to the price level.
2. Daily stock price changes occur in discrete jumps of a small number of ticks.
3. The stock price varies over a relatively wide range.

The derivation of these three conditions is straightforward and the details can be found in Crack and Ledoit (1996)[1]. The patterns in Figures 5.1a-d show the appearance of the compass rose and when it appears. In Figure 5.1a, we plot the log-returns from a stock, "Atlas Copco A Fria" 1977-1984, listed on the Stockholm Stock Exchange.[2] The compass rose pattern can be seen clearly. In Figures 5.1b-d, we plot simulated returns based on simulated prices from an AR-GARCH model (see chapter 5.3) fitted to the same Atlas Copco stock. In Figure 5.1b, it can be seen how the original simulated returns show no pattern. This is radically changed in Figure 5.1c, where we have rounded the simulated prices to mirror the behavior on

[1]Szpiro (1998) shows how a more rigorous treatment of the compass rose makes the assumption that asset prices change by only small amounts relative to the price level superfluous.

[2]Crack and Ledoit derive their three conditions on basis of the assumption of percentage returns. In this paper, we use the more widely used log-returns and we see how the compass rose is not a consequence of the use of percentage returns.

the Stockholm Stock Exchange[3]. This pattern is very similar to the one in Figure 5.1a. It is important to notice, however, that the discreteness induced by the official tick size alone is not the cause of the striking patterns in Figure 5.1a and Figure 5.1c. In Figure 5.1d, we plot the same rounded returns as in Figure 5.1c but randomly permuted (scrambled). The compass rose pattern disappears, which indicates that something must be added to the discreteness to create the pattern. The remaining cross-shaped pattern is merely a consequence of the large number of zero returns that remain in the scrambled series.

5.3 Parameter Estimation

Crack and Ledoit (1996) hypothesize that Autoregressive Conditional Heteroscedasticity (ARCH) models (Engle (1982)) might be influenced by the compass rose. Chen (1997) uses information contained in the compass rose to improve forecasts from an ARMA-GARCH model. He does not study the effect of the discreteness on his parameter estimates, though, which is exactly what we try to investigate with Monte Carlo simulations in this chapter.

In order to study the effects of the discreteness, return series differing in nothing but their level of discreteness must be found. Since this is not easily achieved with empirical data, we have chosen to simulate series which we then round in order to get discrete time series[4]. The simulated series used in this section all come from the same AR-GARCH model:

$$\begin{aligned} r_t &= \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-2} + \varepsilon_t \qquad (1) \\ \sigma_t^2 &= \phi_0 + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2, \end{aligned}$$

[3]On the Stockholm Stock Exchange, the prices are allowed to take only certain discrete values. The level of discreteness depends on the price level; below 5 SEK the smallest price jump is 0.01 SEK, between 5 SEK and 10 SEK the jump is 0.05 SEK, between 10 and 50 SEK it is 0.1 SEK, between 50 SEK and 500 SEK it is 0.5 SEK, and above 500 SEK it is 1 SEK.

[4]When we use the word discrete, we mean discrete in value space, not in time space. All statistical tools used are based on common discrete-time models.

where r_t is the stock log-return, σ_t^2 is the conditional variance of ε_t, and where the parameter vector $\theta = \{\alpha_0, \alpha_1, \alpha_2, \phi_0, \phi_1, \phi_2\} = \{5.2 \cdot 10^{-4}, 0.12, -0.053, 2.2 \cdot 10^{-5}, 0.11, 0.83\}$ comes from an AR-GARCH estimation of the "Atlas Copco A fria" stock on the Stockholm Stock Exchange over the period 1977-1990[5]. Atlas Copco is one of the largest companies on the exchange and the stock is liquid and representative for the time period. Moreover, it also shows a distinct compass rose pattern. Our choice of simulated series is different from the treatment in Krämer and Runde (1997). Their simulated stock price series come from IID normally distributed returns with means and variances chosen without resemblance to stocks in the real world. Therefore, any relationships between the level of rounding and the potential effects detected are not useful.

Using the model in (1), we simulate 1000 return series, each being 2000 observations long. The return series are exponentiated and 1000 price series are computed, all with the starting value of 50. These "original" price series are not discrete and do not show any compass rose pattern. To get price series with varying degrees of discreteness, we round the original price series; either *Integer Rounding* to the nearest integer or *Stock Exchange Rounding* to mirror the stocks on the Stockholm Stock Exchange, as described in chapter 5.2. After calculating log-returns from these series we have $three \times 1000$ return series originating from non-rounded prices, realistically discretized prices, and heavily rounded prices.

Our purpose is to study how the parameter estimates from the $three \times 1000$ return series change with the *three* different degrees of discreteness. The model we apply to these series is exactly the AR-GARCH model in (1) and therefore, we expect the estimated parameter vectors, $\widehat{\theta}$ for the original series, $\widehat{\theta}_S$ for the stock exchange rounded series, and $\widehat{\theta}_I$ for the integer rounded series, to come rather close to θ. However, it is important to note that the AR-GARCH model we use is misspecified when applied to the rounded series, since ε_t cannot be continuous if r_t is discretely distributed around the conditional mean.

In Table 5.1, we present some statistics such as the mean, the 95% confidence

[5]Totally, 3489 daily observations.

interval, the minimum, and the maximum, of the *three*×1000 AR-GARCH estimates. The AR-GARCH parameters were jointly estimated by Maximum Likelihood methods (the BHHH algorithm)[6]. From Table 5.1, we see that the small sample distributions of some of the parameters in $\widehat{\theta}$, $\hat{\phi}_0$ and $\hat{\phi}_2$, have changed, since the true parameters, ϕ_0 and ϕ_2, are outside the 95% confidence intervals of the means of $\hat{\phi}_0$ and $\hat{\phi}_2$. The distribution of the parameter estimates in $\widehat{\theta}_S$ and $\widehat{\theta}_I$ change with the level of discreteness. The confidence intervals are wider and the means of the estimates differ compared to $\widehat{\theta}$, which is most obvious for $\widehat{\alpha}_1$. It is important to emphasize that a combination of two effects is observed. First, in infinite samples, we know that $\widehat{\theta}$ will converge to θ, but since we have rounded the series, we do not know to what $\widehat{\theta}_S$ and $\widehat{\theta}_I$ will converge asymptotically. Second, as mentioned above, the continuous-state AR-GARCH framework is not the proper one to use in modelling discrete return series. Hence, we cannot separate the effect of the rounding on the asymptotics from the effect of the misspecified AR-GARCH model on modelling performance; we can only investigate the compound effect of discreteness on parameter estimation[7].

To say something about how the distributions of $\widehat{\theta}$, $\widehat{\theta}_S$, and $\widehat{\theta}_I$ differ, we analyze the deviation vectors, $\widetilde{\theta}_S = \widehat{\theta} - \widehat{\theta}_S$ and $\widetilde{\theta}_I = \widehat{\theta} - \widehat{\theta}_I$. Comparing individual estimates from the Stock Exchange rounded and Integer rounded series with the original non-rounded series, we get *two*×1000 deviation vectors. In Table 5.2, we look at the distributions of these deviations. The confidence intervals are much wider for $\widetilde{\theta}_I$ compared to $\widetilde{\theta}_S$. The absolute values of the mean deviations increase with the level of discreteness, and the sign of the mean deviations differ between parameters, but do not change with the level of rounding. According to the 95% confidence intervals of the mean deviations, all means of $\widetilde{\theta}_S$ and $\widetilde{\theta}_I$ are quite small but significantly different from zero.

In Table 5.1, further evidence of the effect of discrete prices on model estimation is given by the log-likelihood value statistics from the Monte Carlo simulations. The log-

[6]Estimating the AR parameters prior to the GARCH parameters, did not considerably affect the estimates and these results are therefore not presented.

[7]An extension of the GARCH framework to handle cases where the return series are discrete is given in Amilon (1999).

likelihood values are lower for discrete time series, in particular for heavily rounded series with a mean log-likelihood value clearly below the original series. Looking at the distribution of the log-likelihood deviations in Table 5.2, it becomes even clearer that the rounded series have significantly lower log-likelihood values. In fact, for the integer rounded series, the likelihood values from the original series are never smaller than the corresponding likelihood values from the rounded series. Since we have rounded the series, we would expect the log-likelihood values to change. In this case, they are almost always smaller, which is a consequence of the rounded series no longer fulfilling the assumptions underlying the model in (1).

5.4 Enhanced Forecasts

Another way of assessing the importance of discrete stock prices is to study whether the compass rose can be used to enhance forecasts. Crack and Ledoit (1996) argue that the compass rose contains no information that can be used for predictive purposes. Chen (1997) on the other hand tries to contradict this empirically by incorporating the compass rose pattern in his ARMA-GARCH return forecasts (not modifying his estimates, though) and in this way improving his forecasts[8]. In summary, Chen argues that there are feasible regions (the rays radiating from the origin) and unfeasible regions (the white spaces between the rays) in the compass rose. Chen's enhanced procedure simply tries to force his ARMA-GARCH forecasts to lie within the feasible regions. When a forecast is being made, Chen goes through the constantly updated historical compass rose pattern created from all historical return pairs (r_t, r_{t+1}) and replaces the *original forecasted return* pair $(r_{today}, r_{forecast})$ by the pair from the feasible region closest to the actual forecasted pair in an Euclidean sense. In our paper, we call such enhanced forecasts *rose-enhanced forecasts*.

In addition to Chen's method, we suggest an alternative enhancement considering the fact that prices move in discrete jumps and can only take on certain values.

[8]Chen's results are based on out-of-sample predictions of six stocks during a certain period of time, by using the mean absolute error and the root mean squared error as performance measures. No statistical significance of his results are presented.

Instead of looking at the historical return pattern, we suggest that more emphasize should be put on the feasibility of the forecasted returns, implied by the feasibility of the forecasted prices. Since only a limited number of discrete prices are feasible forecasts for tomorrow's stock price at a certain point in time, we replace our actual forecasted price by one rounded to the nearest possible price[9]. From this forecasted price, we then calculate an associated forecasted return, called *tick-enhanced forecast*, that can be compared to the original forecast and the rose-enhanced forecast as well as the realized out-of-sample return. This is in contrast to Chen's method, which adjusts the forecasts to belong to the set of realized returns, although these are not attainable at all price levels.

Since we do not incorporate the discreteness in our parameter estimates, it can be shown how the enhancements will affect the performance. Suppose that we have estimated our AR-GARCH parameters from a discretized time series. When forecasting, tomorrow's return, r_t, is assumed to be a continuous stochastic variable symmetrically distributed around the conditional mean return, m_t, which is described by the AR parameters. Since r_t is actually discrete, so is m_t. Suppose further that the possible states at time t are $ix, i = 0, \pm 1, \pm 2, ...$, where, for simplicity, we let $x = 1$ [10]. Let p_j denote the conditional probability that state j occurs. Because of the symmetry of r_t, it follows that if $j < m_t < j + 1/2$, then $p_j > p_{j+1} > p_{j-1} > ... > p_{j+N} > p_{j-N}$, where we limit the possible states to $2N + 1$, where N states are larger than j, and N states are smaller than j. We can let $N \to \infty$, or we can truncate the distribution by letting $p_{j\pm N} = \sum_{l=1}^{\infty} p_{j\pm(N+l)}$. In the same way, if $j - 1/2 < m_t < j$, then $p_j > p_{j-1} > p_{j+1} > ... > p_{j-N} > p_{j+N}$. We disregard the low probability events of $m_t = j$, and $m_t = j \pm 1/2$, although they can be treated in a similar manner. Figure 5.2 illustrates the notation for $j = 0, N = 2$, and $m_t = 0.25$.

[9]The possible prices depend on the level of discreteness; either we allow only integer prices or also the rounded prices mirroring the Stockholm Stock Exchange prices.

[10]Since we are using log-returns calculated from discrete prices, this is only approximately correct.

Now, if our forecast is f, the Mean Squared Error will become:

$$\text{MSE} = \sum_{i=j-N}^{j+N} p_i(f-i)^2 = E[(f-i)^2] = V[f-i] + (E[f-i])^2 = V[i] + (f-E[i])^2,$$

which is minimized for $f = E[i]$ which, in turn, should be close to m_t. Thus, the original forecasted return will always have the lowest MSE, although this forecast is not feasible. Let us consider the Mean Absolute Error on the interval $k \leq f \leq k+1$:

$$\begin{aligned} \text{MAE} &= \sum_{i=j-N}^{j+N} p_i|f-i| = \sum_{j-N}^{k} p_i(f-i) + \sum_{k+1}^{j+N} p_i(i-f) = \\ & f(\sum_{j-N}^{k} p_i - \sum_{k+1}^{j+N} p_i) + \sum_{k+1}^{j+N} p_i i - \sum_{j-N}^{k} p_i i. \end{aligned} \tag{2}$$

If $\sum_{j-N}^{k} p_i < \sum_{k+1}^{j+N} p_i$, the MAE is minimized when f is chosen as large as possible on the interval and consequently, if $\sum_{j-N}^{k} p_i > \sum_{k+1}^{j+N} p_i$, then f should be as small as possible. What happens when we move from one interval to the next? If we let $f_1 = k - \epsilon$ and $f_2 = k + \epsilon$, with ϵ being small, it is easily shown that

$$\text{MAE}_2 - \text{MAE}_1 = 2\epsilon(\sum_{j-N}^{k-1} p_i - \sum_{k+1}^{j+N} p_i). \tag{3}$$

In the case of $p_j > p_{j+1} > p_{j-1} > ... > p_{j+N} > p_{j-N}$, (3) is negative if $k \leq j$, and positive if $k > j$. The MAE is therefore minimized when $k = j$, and according to (2), f should be chosen as small as possible, that is $f = j$. Similarly, when $p_j > p_{j-1} > p_{j+1} > ... > p_{j-N} > p_{j+N}$, (3) is negative for $k < j$, and positive otherwise. The minimizing state is then $k = j - 1$, and (2) is, once more, minimized for $f = j$. In a MAE sense, choosing the most probable feasible outcome is favored compared to the expected forecast. Thus, our enhanced forecast will outperform the original one, using the MAE as a performance measure.

Let us examine our theoretical results empirically in the following way: AR-GARCH return forecasts are nothing but the returns forecasted from the mean process in (1). Using the *two* $\times$ 1000 different parameter sets $\widehat{\theta}_S$ and $\widehat{\theta}_I$, estimated in chapter 5.3, we forecast returns by using the different forecasting methods. As a test sample, we use *two* $\times$ 1000 different return series simulated from (1), each 1000 observations long, and rounded in the two different ways[11]. For each of the *two* $\times$ 1000 parameter sets, Root Mean Squared Errors (RMSE) as well as Mean Absolute Errors (MAE) are calculated over the test sample for the different forecasting methods and different roundings. In our setting, the RMSE and the MAE are:

$$\begin{aligned} \text{RMSE} &= \left[\frac{1}{1000}\sum_{i=1}^{1000}(r_i - \hat{r}_i)^2\right]^{\frac{1}{2}}, \\ \text{MAE} &= \frac{1}{1000}\sum_{i=1}^{1000}|r_i - \hat{r}_i|, \end{aligned}$$

where r_i is the actual return at day i, $\hat{r}_i$ is the forecasted return at day i, and the number of days in the test period is equal to 1000.

In Table 5.3, we present the means and the 95% confidence intervals of the ratios of the enhanced and the original forecast errors. Starting with the MAE statistic, we can see how the mean ratios are generally close to one but that our tick-enhanced method clearly outperforms the original non-enhanced method; the mean of the ratio is significantly smaller than one on the 5% level for both the exchange rounded and the integer rounded series. Even for the RMSE statistics, the mean ratios are close to one but otherwise, the results are somewhat reversed and the original method seems to outperform the tick-enhanced one; for the exchange rounded series, the mean ratio is significantly larger than one on the 5% level, while for the integer rounded series the mean ratio is neither significantly larger nor smaller than one on the 5% level. The results are qualitatively the same for the rose-enhanced forecasts; a mean ratio significantly smaller than one for the MAE, and a mean ratio significantly larger than one for the RMSE. These results confirm our theoretical conclusions about the choice

[11] Each of the *two* $\times$ 1000 parameter sets (estimated by using the observations in the estimation period) is tested on exactly one of the *two* $\times$ 1000 (out-of-sample) test series. The model parameters in each set are kept constant over the entire test period.

of evaluation measure; the MAE measure indicates gains from using our enhanced method while the RMSE seems to favor the original method. It is also worth noticing how the relative performance of the enhanced methods improves when we round more heavily; throughout Table 5.3, we see that the heavier we round, the smaller are the error ratios. Finally, if the tick-enhanced method were to be directly compared with the rose-enhanced method, significant results indicating a dominance of our tick-enhanced method over Chen's rose-enhanced method would most probably be found; in three cases out of four, the mean ratios are smaller for the tick-enhanced method.

The results depend, a priori, on the choice of performance measure, so which should be chosen? If constraining one's forecasts to belong to the set of possible outcomes is believed to be reasonable, then the forecasts should be adjusted to be feasible, as we have done, and an evaluation measure favoring a feasible forecast, such as the MAE, should be chosen. It is obvious that the continuous AR-GARCH framework is not the proper one to apply to discrete return series, but we see no theoretical or empirical justification why Chen's adjustments should result in better forecasts. However, the discreteness of the returns should already be incorporated in the parameter estimates to fully improve forecasts, an issue treated in Amilon (1999).

5.5 Correlation Integral Statistics

In this chapter, we will examine how the distribution of some correlation integral based statistics, defined below, are influenced by the fact that prices only move in discrete ticks. We are primarily concerned with two issues: Do the null distributions of the test statistics change, when applied to return series originating from discrete prices? How are the powers of these test statistics affected when trying to detect explicit time series dependences in such return series?

5.5.1 The BDS Test and Savit and Green's Dependability Index

Grassberger and Procaccia (1983) introduced a quantity called the correlation integral, in order to identify structures and dependences in data series. In the case of time series analysis, the correlation integral is computed by first forming m-histories from the time series considered:

$$\mathbf{x}(t) = (r(t), r(t-1), ..., r(t-m+1)) = (x_1(t), x_2(t), ..., x_m(t)).$$

If the distance between the kth components of two m-histories, $\mathbf{x}(t)$ and $\mathbf{x}(s)$, is defined as

$$l_k(t,s) = |x_k(t) - x_k(s)|, \qquad k = 1, 2, ..., m,$$

then the correlation integral at embedding dimension m and tolerance ϵ can be expressed as

$$C_m(\epsilon) = \frac{1}{N_{pair}} n(l_1 \leq \epsilon, ..., l_m \leq \epsilon), \tag{4}$$

where N_{pair} is the total number of pairs, and $n(l_1 \leq \epsilon, ..., l_m \leq \epsilon)$ is the number of pairs with all components within ϵ apart.

The BDS test is based on the observation that for an IID sample,

$$C_m(\epsilon) = [C_1(\epsilon)]^m.$$

The identity should be understood in a statistical sense. Brock *et al.* (1996) derived a normalization factor, $V_m(\epsilon)$, in order to make a correct statistical quantification of the departure from IID. More specifically, they showed that the BDS statistic

$$W_m(\epsilon) = \frac{C_m(\epsilon) - [C_1(\epsilon)]^m}{V_m(\epsilon)}, \tag{5}$$

converges in distribution to $N(0,1)$, as the time series become infinitely long, for $\epsilon > 0$, and $m > 1$, under the null hypothesis of IID. With the possibility of bootstrapping small sample distributions as in Efron (1979), the asymptotic properties are nowadays of less importance. The BDS test has proved to be quite successful in finding departures from IID in a number of Monte Carlo studies (see, for example, Brock *et al.* (1991)), but the test gives no hints of at what time lags there are dependences in the data, information that is most useful in time series modelling and analysis.

Savit and Green (1991) filled this gap by introducing a dependability index computed from the correlation integrals in different embedding dimensions. As seen from (4), $C_m(\epsilon)$ is nothing but the joint probability of two m-histories being no more than ϵ apart in all their Cartesian components, that is $C_m(\epsilon) = \Pr(l_1 \leq \epsilon, ..., l_m \leq \epsilon)$. This holds under the assumption of time-invariance or stationarity, so that we can compare pairs from different parts of the time series. It is also possible to form the conditional probabilities of two observations being close, given that their m-histories are close:

$$\Pr(l_1 \leq \epsilon | l_2 \leq \epsilon, ..., l_m \leq \epsilon) = \frac{\Pr(l_1 \leq \epsilon, ..., l_m \leq \epsilon)}{\Pr(l_2 \leq \epsilon, ..., l_m \leq \epsilon)} = \frac{C_m(\epsilon)}{C_{m-1}(\epsilon)},$$

since $\Pr(l_2 \leq \epsilon, ..., l_m \leq \epsilon) = \Pr(l_1 \leq \epsilon, ..., l_{m-1} \leq \epsilon)$ by construction. In the same way,

$$\Pr(l_1 \leq \epsilon | l_2 \leq \epsilon, ..., l_{m-1} \leq \epsilon) = \frac{C_{m-1}(\epsilon)}{C_{m-2}(\epsilon)}.$$

If $r(t)$ does not depend on $r(t-m+1)$, then

$$\Pr(l_1 \leq \epsilon | l_2 \leq \epsilon, ..., l_m \leq \epsilon) = \Pr(l_1 \leq \epsilon | l_2 \leq \epsilon, ..., l_{m-1} \leq \epsilon),$$

that is, $C_m(\epsilon)/C_{m-1}(\epsilon) = C_{m-1}(\epsilon)/C_{m-2}(\epsilon)$. It is now possible to define a dependability index

$$\delta_{m-1}(\epsilon) = 1 - \frac{C_{m-1}^2(\epsilon)}{C_m(\epsilon)C_{m-2}(\epsilon)} \tag{6}$$

for $m > 1$ ($C_0 \equiv 1$), which is zero (in a statistical sense), if there is no dependence of $r(t)$ on the lag $t - m + 1$. In contrast to the BDS test, the delta indexes provide information on at what time lags there are dependences causing the rejection of IID. An asymptotic distribution for the δ's under the null hypothesis of IID has not been working out, but can easily be estimated by a bootstrap procedure.

5.5.2 Rounding of IID Series

In order to investigate the ability of the correlation integral based statistics to pick up the micromarket dependences caused by the rounding of prices we do the following Monte Carlo simulation: First, we generate 1000 return series of the length of 2000 from $N(\mu, \sigma)$, with the same unconditional moments as implied by the AR-GARCH coefficients in chapter 5.3, that is $\mu = 5.57 \cdot 10^{-4}$, and $\sigma = 0.019$. After exponentiating to prices, rounding and calculating log-returns as described in chapter 5.3, we have $three \times 1000$ return series originating from non-rounded prices, realistically discretized prices, and heavily rounded prices. An integer rounding of prices around and below 50 must be regarded as unrealistic, at least for modern financial data. Henceforth, we denote these series R, RS, and RI.

In Table 5.4, we report some distributional statistics of the W's and the δ's, such as the mean, the standard deviation, the skewness, the kurtosis, the minimum, the maximum, and the 2.5% and the 97.5% percentiles. The embedding dimensions are $m = 2, 3, 4$, and 5, corresponding to time lags 1 to 4, and the tolerance parameter ϵ equal to the standard deviation of each time series, as is often suggested.[12]

For R, the distribution of the BDS statistics is close to $N(0, 1)$, as expected for sample sizes of this magnitude. The δ's also seem to be normally distributed around zero, indicating no dependences on past lags. In the case of RS, the picture is very much the same. Both the BDS statistics and the deltas are normally distributed, but the distribution appears to be shifted somewhat upward, which is especially notable for W_5, and δ_2. The percentiles and the mean for W_5 are -1.74, 2.35, and 0.16 as compared to -1.87, 2.10, and -0.03 for the R statistics. For realistically discretized data, the rounding mechanism is picked up by the correlation integral statistics, but the introduced shifts of the distributions are not very severe. This is certainly not true when examining RI. Neither the W's nor the δ's are normally distributed, and the means as well as the critical values are heavily shifted upward for all m, although most heavily when $m = 2$.

[12] Throughout this paper, we report our results for $\epsilon = \sigma$ only. We have also examined the cases of $\epsilon = 0.25\sigma, 0.5\sigma, 1.5\sigma$, and 2.0σ, with similar findings.

The δ's are all large, indicating dependences, but our simulated series have no explicit time dependences. How is this possible? The reason is that the rounded series are not time-invariant, which is assumed in the derivation of the correlation integral statistics. At a given price, a stock will only jump a certain amount of ticks, giving rise to a finite number of return realizations. As the stock price evolves, the number of probable realizations will change, increasing if the stock goes up and decreasing otherwise (if the tick size is the same for all prices). In addition, the nonstationarity in returns is linked to the price level, which creates new complicated dependences. Their appearance is best illustrated by an example.

Let $r_1, r_2 \in N(0, \sigma)$ denote the returns at times $t = 1$ and $t = 2$, let $\sigma = 0.019$, and let the starting price be $P_0 = 100$. The prices at $t = 1$ and $t = 2$ are then $P_1 = 100\exp(r_1)$ and $P_2 = 100\exp(r_1 + r_2)$, the rounded prices $\bar{P}_1 = [100\exp(r_1)]$ and $\bar{P}_2 = [100\exp(r_1 + r_2)]$, where $[\cdot]$ denotes integer rounding, and the resulting rounded returns are $\bar{r}_1 = \log\left(\bar{P}_1/P_0\right)$ and $\bar{r}_2 = \log\left(\bar{P}_2/\bar{P}_1\right)$. Suppose we want to calculate the probability of $\bar{r}_1 = 0$, and $\bar{r}_2 = \log(101/100)$. The probability that $\bar{r}_1 = 0$ is equal to:

$$\Pr(\bar{r}_1 = 0) = \Pr\left(\bar{P}_1 = 100\right) = \Pr\left(100 - \frac{1}{2} < P_1 < 100 + \frac{1}{2}\right) =$$
$$= \Pr\left(\log\left(\frac{100 - \frac{1}{2}}{100}\right) < r_1 < \log\left(\frac{100 + \frac{1}{2}}{100}\right)\right).$$

The joint probability that $\bar{r}_1 = 0$ and $\bar{r}_2 = \log(101/100)$ is given by:

$$\Pr\left(\bar{r}_1 = 0, \bar{r}_2 = \log\left(\frac{101}{100}\right)\right) = \Pr\left(\bar{P}_1 = 100, \bar{P}_2 = 101\right) = \tag{7}$$
$$= \Pr\left(\log\left(\frac{100 - \frac{1}{2}}{100}\right) < r_1 < \log\left(\frac{100 + \frac{1}{2}}{100}\right), \log\left(\frac{101 - \frac{1}{2}}{100}\right) < r_1 + r_2 < \log\left(\frac{101 + \frac{1}{2}}{100}\right)\right)$$

However, the probability that $\bar{r}_2 = \log(101/100)$ is given by the following:

$$\Pr\left(\bar{r}_2 = \log\left(\frac{101}{100}\right)\right) = \Pr\left(\frac{\bar{P}_2}{\bar{P}_1} = \frac{101}{100}\right) = \sum_k \Pr\left(\bar{P}_1 = k, \bar{P}_2 = k\frac{101}{100}\right). \tag{8}$$

Since we use integer rounding, the probabilities in the sum are zero unless $k = 100, 200, ..., \infty$. Besides, the probability that $\bar{P}_1 = 200$ or more is almost zero, so (8) reduces to

$$\Pr\left(\bar{r}_2 = \log\left(\frac{101}{100}\right)\right) \simeq \Pr\left(\bar{P}_1 = k, \bar{P}_2 = k\frac{101}{100}\right)|_{k=100} = \Pr\left(\bar{r}_1 = 0, \bar{r}_2 = \log\left(\frac{101}{100}\right)\right)$$

If the rounded returns $\bar{r}_1, \bar{r}_2$ were independent, then

$$\Pr\left(\bar{r}_1 = 0, \bar{r}_2 = \log\left(\frac{101}{100}\right)\right) = \Pr\left(\bar{r}_1 = 0\right) \times \Pr\left(\bar{r}_2 = \log\left(\frac{101}{100}\right)\right),$$

which is obviously not satisfied here. It is these kind of dependences, which could be more or less pronounced in different time series, that are detected by the correlation integral statistics[13]. The rounded returns are no longer IID. In Table 5.5, we show the BDS test statistics, and the autocorrelation in returns and squared returns, for a typical series (series no. 347) in R, RS, and RI. We see that the BDS test strongly rejects integer rounded IID returns, a finding not discovered by just examining the autocorrelations, as is the common approach in econometric analysis.

To summarize, the rounding of prices has two effects. It makes the rounded return series time-variant, and it introduces complicated dependences in the series. The BDS test correctly rejects the rounded returns as IID variables, but the lag-dependences identified by the δ's may be spurious, since the rounded series are nonstationary.

[13]The probability in (7) can be calculated by numerically integrating

$$\frac{1}{\sqrt{2\pi}}\int_{r_1=\log\frac{100-1/2}{100}}^{\log\frac{100+1/2}{100}} \exp\left(-\frac{r_1^2}{2\sigma^2}\right)\left(\Phi\left(\frac{\log\left(\frac{101+\frac{1}{2}}{100}\right)-r_1}{\sigma}\right)-\Phi\left(\frac{\log\left(\frac{101-\frac{1}{2}}{100}\right)-r_1}{\sigma}\right)\right)dr_1.$$

With the numbers chosen here, $\Pr\left(\bar{r}_1 = 0\right) \simeq 0.2076$ and $\Pr\left(\bar{r}_1 = 0, \bar{r}_2 = \log\left(\frac{101}{100}\right)\right) = \Pr\left(\bar{r}_2 = \log\left(\frac{101}{100}\right)\right) \simeq 0.0703$.

5.5.3 Rounding of AR-GARCH Series

We turn to exploring the power of our test statistics when applied to our simulated AR-GARCH return series from chapter 5.3, denoted G, GS, and GI. We have established that the correlation integral statistics are distorted by the nonstationarities and dependences introduced by price rounding. Here, we are interested in the ability of detecting explicit time series dependences. In Table 5.6, we report the minimum, the maximum, and the 95% confidence interval of the test statistics, together with the frequencies of rejecting the null hypothesis of rounded IID observations, at the 5% significance level, when it is false. In the power tests, we are using the simulated critical values from R, RS, and RI in Table 5.4.

Somewhat surprisingly, no sign of upward shifts is distinguishable in the confidence intervals of the BDS statistics when we increase the level of discreteness. The δ's, on the other hand, show more of the previous positive shifts, at least for the integer rounded GI. Because of the upward shifts presented in RS and RI, in combination with the tendency of downward shifts in GS and GI, the power of the test statistics deteriorates, due to the rounding effects. The empirical rejection frequencies of the BDS statistics are 100% for both G and GS. The δ's clearly also identify dependences at all time lags, as should be the case for GARCH series. The ability of detecting lagged dependences weakens as m increases, and also when comparing G and GS. This should not be confused with any superiority of the BDS test over Savit and Green's dependability index. The former is a true portmanteau test, answering the question if a time series is IID or not. The latter, on the other hand, determines whether including additional lags raises the conditional probability of two observations being close, given that their m-histories are close. Deciding whether a time series is IID or not just by examining a certain lag, rather than using information at all lags, is quite a different task. The BDS test and the dependability indexes therefore give different, but complementary, information usable in time series analysis.

When examining GI, the power of the test statistics falls dramatically to around 22% for the W's and to 8% for the δ's. Trying to determine whether an integer rounded time series (at the price levels examined here) has any explicit time series dependences based on the null distributions of the test statistics would most likely lead to the wrong conclusion. It is somewhat surprising that the power of W_2 and δ_1 are so different, since the numerator of W_2 in (5) is equal to $C_2\delta_1$. Obviously, this is related to the influence of the denominator of the BDS test statistic[14].

Let us examine the standardized residuals from the maximum likelihood estimations of the AR-GARCH series, denoted S, SS, and SI. As shown in Brock *et al.* (1996), the asymptotic properties of the BDS test are the same, whether the test is applied to IID series or residuals from linear (and some nonlinear) stochastic models. This nuisance parameter free property of the BDS test is not valid for GARCH residuals, which is clearly visible in Table 5.7. The test statistics are still normally distributed, but with smaller variances, resulting in narrower confidence intervals (see Hsieh (1989)).

The distributional properties of S and SS are again very similar, confirming that realistic rounding of prices does not change the null distribution of the correlation integral based statistics to a very large extent. Once again, these similarities are lost when examining SI. We have upward shifts in the distributions, giving rise to totally different confidence intervals for both the W's and the δ's. Furthermore, there is no longer only a scaling factor of the standard deviation separating the distributions of RI and SI, as is the case between the distributions of R and S. If one wishes to investigate the standardized residuals from a heavily rounded return series, there seem to be no shortcuts. A proper Monte Carlo simulation, similar to ours, must be performed in order to extract the critical values to compare with the test statistics of the residuals of one's GARCH model. Even so, the power of such a test would most probably be quite low.

[14]When investigating the power of $C_2\delta_1$, compared to the integer rounded null distribution of the same quantity, the power raises to 12.1%, but is still far from 22.8%.

5.6 Conclusions

Throughout this paper, we try to shed some light on the consequences of the trade induced compass rose pattern in stock returns. We examine how AR-GARCH parameter estimates change with the level of rounding in stock prices, and find that the distribution of the estimates differ, in particular for the AR(1) parameter. The differences arise from two effects. First, when rounding the series, we may change the dynamics of the processes, and second, the rounded series no longer fullfil the assumptions underlying the continuous-state AR-GARCH model. Obviously, the higher the level of discreteness, the more pronounced will either of the two effects be.

Further, we show theoretically that incorporating discretization in return forecasts (not in estimation), as outlined in chapter 5.4, improves the performance in an MAE sense, while the opposite holds using the RMSE as an evaluation measure. Simulations reveal that the out-of-sample performance is better, significantly at the 5% level, when using the MAE measure, and we argue for the use of this measure if wishing to favor feasible forecasts.

We also investigate how the distributions of some correlation integral statistics change when applied to rounded return series, and residuals from such series. Our findings suggest that the effects on the distributions are small, provided the return series come from realistic roundings of prices, such as those present at the Stockholm Stock Exchange. Only when we investigate heavy rounding, most likely uncommon in modern financial markets, do the null distribution of the test statistics change remarkably. This is contrary to the findings of Krämer and Runde (1997), who discovered large changes in the null distribution of the BDS test already at low roundings (tick size of 0.1), due to their highly unrealistic simulated stock prices. In chapter 5.5, we explain why these changes occur. The rounding of prices makes the corresponding return series time-variant, and introduces dependences in the series. The rounded IID returns are no longer IID. The distributions of the correlation integral statistics therefore change, because of the dependences and the nonstationarities in the rounded data.

The main conclusion is that the effects of discrete prices are small, at least for a discretization comparable to the one present at the Stockholm Stock Exchange. When investigating series with higher tick size to price ratios, such as those present in low-priced stocks, the use of statistical models and tests based on state-continuity can be questioned.

Our results are based on time series of the length of 2000. It may be the case that the rounding effects are more severe when examining shorter time series. Caution should also be taken when less traded stocks are examined. The "effective" tick size chosen by market participants can then be larger than the official tick size.

References

Amilon, H. (1999), "GARCH Estimation and Discrete Stock Prices", Working paper, Lund University, Lund.

Atchison, M., and White, M. (1996), "Disappearing Evidence of Chaos in Security Returns: A Simulation", *Quarterly Journal of Business and Economics*, 35(2), 21-37.

Brock, W., Dechert, W., Scheinkmann, J., and LeBaron, B. (1996), "A Test for Independence Based on the Correlation Dimension", *Econometric Reviews*, 15(3), 197-235.

Brock, W., Hsieh, D., and LeBaron, B. (1991), *Nonlinear Dynamics, Chaos and Instability: Theory and Evidence*, M.I.T. Press, Cambridge.

Chen, A.-S. (1997), "The Square Compass Rose: The Evidence from Taiwan", *Journal of Multinational Financial Management*, 7, 127-144.

Crack, T., F., and Ledoit, O. (1996), "Robust Structure Without Predictability: The "Compass Rose" Pattern of the Stock Market", *Journal of Finance*, June, 751-762.

Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife", *Annals of Statistics*, 7, 1-26.

Engle, R. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica* 50, 987-1007.

Grassberger, P., and Proccacia, I. (1983), "Measuring the Strangeness of Strange Attractors", *Physica D,* 9, 189-208.

Hsieh, D. (1989), "Testing for Nonlinear Dependence in Daily Foreign Exchange Rates", *Journal of Business*, vol. 62, no. 3, 339-368.

Krämer, W., and Runde, R. (1997), "Chaos and the Compass Rose", *Economics Letters*, 54, 113-118.

Savit, R. and Green, M. (1991), "Time Series and Dependent Variables", *Physica D,* 50, 95-116.

Szpiro, G,. G. (1998), "Tick Size, the Compass Rose and Market Nanostructure", *Journal of Banking and Finance*, 22*,* 1559-1569.

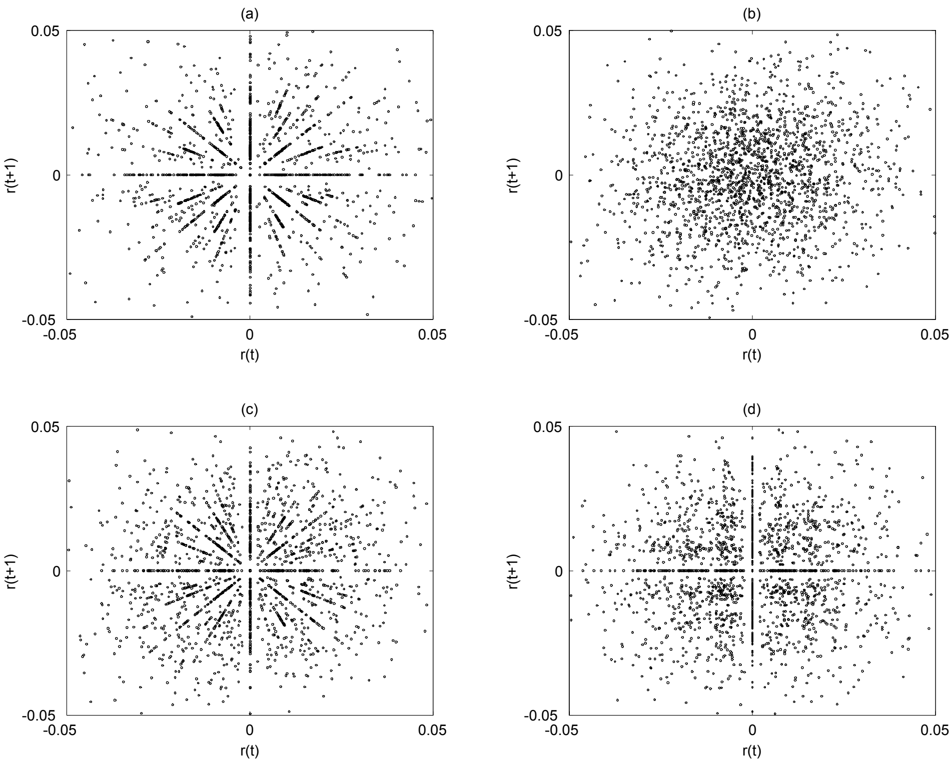

Figure 5.1: The compass rose pattern in: (a) Atlas Copco, (b) simulated returns, (c) stock exchange rounded simulated returns, and (d) scrambled stock exchange rounded simulated returns. All time series are of length 2000.

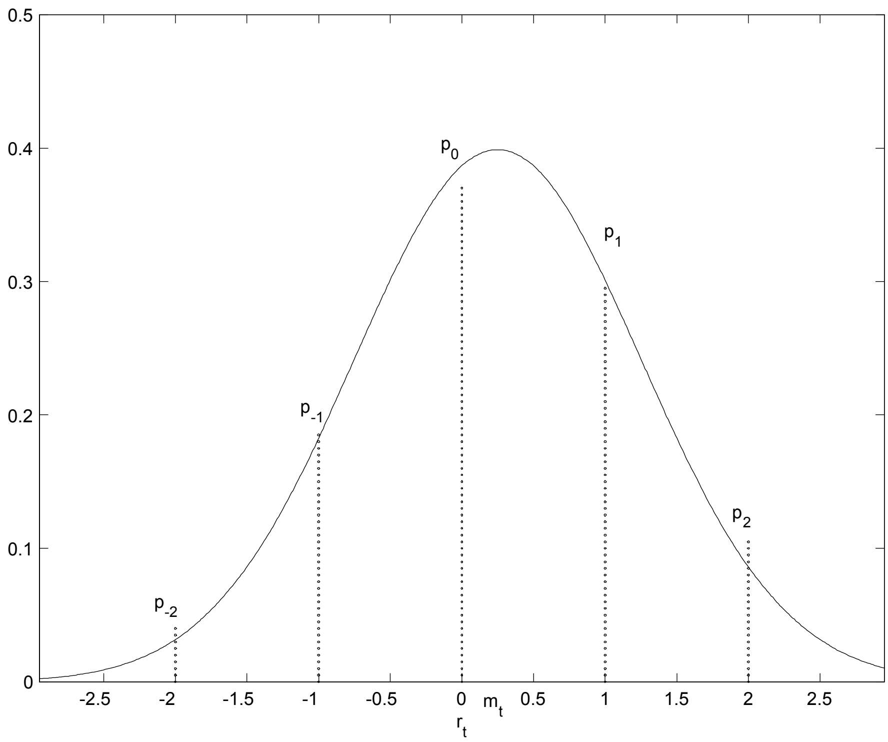

Figure 5.2: The conditional density function of r_t around m_t, and it's discrete counterparts $p_0, p_{\pm 1}$, and $p_{\pm 2}$.

.

Table 5.1: Parameter estimates and likelihood values. $\widehat{\theta}$, $\widehat{\theta}_S$, and $\widehat{\theta}_I$ denote the parameter vector estimates of the original, the stock exchange rounded, and the integer rounded series respectively. The empirical α-percentile is denoted I_α. Small numbers are 95 percent confidence intervals.

		$\hat{\phi}_0 \cdot 10^5$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\widehat{\alpha}_0 \cdot 10^4$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\log L$
	mean	2.314	0.110	0.826	5.342	0.118	-0.054	3.502
		(2.273,2.355)	(0.109,0.111)	(0.824,0.828)	(5.103,5.582)	(0.117,0.120)	(-0.056,-0.053)	(3.500,3.505)
	$I_{0.025}$	1.258	0.075	0.757	-2.395	0.073	-0.103	3.419
$\widehat{\theta}$	$I_{0.975}$	3.815	0.147	0.882	12.819	0.165	-0.010	3.579
	minimum	0.910	0.056	0.709	-8.080	0.037	-0.135	3.367
	maximum	5.471	0.175	0.908	16.914	0.191	0.014	3.624
		$\hat{\phi}_0 \cdot 10^5$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\widehat{\alpha}_0 \cdot 10^4$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\log L$
	mean	2.370	0.107	0.828	5.372	0.108	-0.052	3.493
		(2.328,2.412)	(0.106,0.108)	(0.826,0.830)	(5.131,5.612)	(0.107,0.110)	(-0.053,-0.050)	(3.491,3.496)
	$I_{0.025}$	1.267	0.073	0.757	-2.385	0.062	-0.098	3.414
$\widehat{\theta}_S$	$I_{0.975}$	3.928	0.144	0.885	12.791	0.155	-0.008	3.568
	minimum	0.908	0.053	0.689	-8.246	0.033	-0.130	3.364
	maximum	5.495	0.179	0.908	16.881	0.178	0.023	3.607
		$\hat{\phi}_0 \cdot 10^5$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\widehat{\alpha}_0 \cdot 10^4$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\log L$
	mean	2.539	0.091	0.846	5.546	0.048	-0.042	3.433
		(2.484,2.595)	(0.090,0.093)	(0.844,0.848)	(5.289,5.802)	(0.044,0.051)	(-0.044,-0.042)	(3,429,3.437)
	$I_{0.025}$	1.181	0.042	0.769	-3.261	-0.092	-0.092	3.267
$\widehat{\theta}_I$	$I_{0.975}$	4.589	0.135	0.918	13.296	0.128	0.001	3.537
	minimum	0.621	0.047	0.681	-9.071	-0.215	-0.122	3.040
	maximum	10.022	0.162	0.959	17.885	0.152	0.027	3.571

Table 5.2: Parameter estimate and likelihood value differences. $\widetilde{\theta}_S = \widehat{\theta} - \widehat{\theta}_S$ and $\widetilde{\theta}_I = \widehat{\theta} - \widehat{\theta}_I$, where $\widehat{\theta}$, $\widehat{\theta}_S$, and $\widehat{\theta}_I$ denote the parameter vector estimates of the original, the stock exchange rounded, and the integer rounded series respectively. The empirical α-percentile is denoted I_α. Small numbers are 95 percent confidence intervals.

		$\widetilde{\phi}_0 \cdot 10^5$	$\widetilde{\phi}_1$	$\widetilde{\phi}_2$	$\widetilde{\alpha}_0 \cdot 10^4$	$\widetilde{\alpha}_1$	$\widetilde{\alpha}_2$	$\widetilde{\log L}$
$\widetilde{\theta}_S$	mean	-0.056 (-0.065,-0.048)	0.003 (0.002,0.003)	-0.002 (-0.003,-0.002)	-0.029 (-0.042,-0.016)	0.010 (0.010,0.010)	-0.003 (-0.003,-0.002)	0.009 (0.009,0.009)
	$I_{0.025}$	-0.360	-0.005	-0.017	-0.460	-0.000	-0.012	0.001
	$I_{0.975}$	0.223	0.011	0.012	0.411	0.023	0.005	0.020
		$\widetilde{\phi}_0 \cdot 10^5$	$\widetilde{\phi}_1$	$\widetilde{\phi}_2$	$\widetilde{\alpha}_0 \cdot 10^4$	$\widetilde{\alpha}_1$	$\widetilde{\alpha}_2$	$\widetilde{\log L}$
$\widetilde{\theta}_I$	mean	-0.225 (-0.261,-0.188)	0.018 (0.017,0.019)	-0.019 (-0.021,-0.018)	-0.203 (-0.249,-0.157)	0.070 (0.068,0.074)	-0.012 (-0.013,-0.011)	0.069 (0.066,0.073)
	$I_{0.025}$	-1.538	-0.003	-0.089	-1.621	0.015	-0.041	0.014
	$I_{0.975}$	0.878	0.062	0.018	1.464	0.210	0.012	0.225

Table 5.3: Means estimates and 95% confidence intervals of the out-of-sample forecast error ratios. Each statistic is computed from a test sample of length 1000. Small numbers are 95% confidence intervals of the mean estimates.

	Exchange Rounding	Integer Rounding
$\hat{E}\left[\frac{MAE_{tick\text{-}enhanced}}{MAE_{original}}\right]$	0.9951 (0.9949,0.9953)	0.9738 (0.9728,0.9748)
$I_{0.025}$	0.9876	0.9319
$I_{0.975}$	1.0014	0.9955
$\hat{E}\left[\frac{MAE_{rose\text{-}enhanced}}{MAE_{original}}\right]$	0.9993 (0.9992,0.9994)	0.9841 (0.9834,0.9848)
$I_{0.025}$	0.9945	0.9537
$I_{0.975}$	1.0036	1.0000
$E\left[\frac{RMSE_{tick\text{-}enhanced}}{RMSE_{original}}\right]$	1.0055 (1.0052,1.0058)	0.9995 (0.9989,1.0001)
$I_{0.025}$	0.9956	0.9738
$I_{0.975}$	1.0171	1.0151
$\hat{E}\left[\frac{RMSE_{rose\text{-}enhanced}}{RMSE_{original}}\right]$	1.0031 (1.0028,1.0034)	1.0012 (1.0007,1.0017)
$I_{0.025}$	0.9946	0.9781
$I_{0.975}$	1.0127	1.0144

Table 5.4: Distributional properties for the BDS statistics $W_m(\epsilon)$, and Savit and Green's dependability index $\delta_{m-1}(\epsilon)$ of the R, RS, and RI series, for $m = 2, 3, 4$, and 5, and with $\epsilon = \sigma$, the standard deviation of each time series. The α-percentile is denoted I_α. Note that the entries corresponding to the δ's are multiplied with 100, *except* for the entries corresponding to skewness and kurtosis.

		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	mean	-0.04	-0.02	-0.02	-0.03	-0.02	0.01	-0.02	-0.03
	std.dev.	0.99	0.99	1.01	1.01	0.46	0.49	0.53	0.54
	skewness	0.10	0.13	0.23	0.29	0.07	0.18	0.12	0.08
R	kurtosis	3.11	2.97	3.10	3.27	3.12	3.08	2.97	2.95
	minimum	-3.57	-3.01	-2.97	-2.82	-1.70	-1.51	-1.48	-1.56
	maximum	3.34	3.55	4.14	4.60	1.51	1.67	1.89	1.93
	$I_{0.025}$	-1.88	-1.88	-1.92	-1.87	-0.89	-0.90	-1.02	-1.06
	$I_{0.975}$	2.07	1.97	2.05	2.10	0.92	1.04	1.06	0.98
		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	mean	0.07	0.13	0.15	0.16	0.03	0.06	0.03	0.01
	std.dev.	1.00	1.00	1.02	1.04	0.46	0.50	0.53	0.56
	skewness	0.08	0.11	0.22	0.30	0.06	0.17	0.16	0.06
RS	kurtosis	3.12	2.94	2.95	3.07	3.10	2.90	2.90	3.16
	minimum	-3.20	-2.76	-2.57	-2.38	-1.43	-1.52	-1.51	-1.60
	maximum	3.32	3.46	4.04	4.41	1.48	1.63	1.82	2.30
	$I_{0.025}$	-1.88	-1.74	-1.84	-1.74	-0.85	-0.85	-0.96	-1.10
	$I_{0.975}$	2.12	2.08	2.32	2.35	0.99	1.08	1.09	1.11
		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	mean	2.42	3.06	3.48	3.84	1.47	1.14	0.91	0.81
	std.dev.	2.43	2.98	3.42	3.85	1.69	1.35	1.16	1.04
	skewness	0.96	1.13	1.27	1.42	1.41	1.42	1.39	1.17
RI	kurtosis	3.46	3.81	4.26	4.80	4.47	5.01	5.69	5.10
	minimum	-2.01	-2.01	-1.97	-1.76	-1.00	-1.23	-1.42	-1.51
	maximum	11.08	13.80	17.14	19.99	8.47	7.35	7.11	5.43
	$I_{0.025}$	-0.91	-0.69	-0.60	-0.58	-0.45	-0.56	-0.74	-0.82
	$I_{0.975}$	8.24	10.48	12.23	14.42	5.83	4.84	3.96	3.65

Table 5.5: BDS test statistics and autocorrelations in returns and in squared returns for lag 1-4, for a typical series (series no. 347) in R, RS, and RI, denoted r, rs, and ri.

					AC in returns				AC in squared returns			
	W_2	W_3	W_4	W_5	ac_1	ac_2	ac_3	ac_4	ac_1	ac_2	ac_3	ac_4
r	0.18	0.33	0.64	0.69	0.02	-0.01	0.02	-0.03	-0.01	0.01	0.03	0.01
rs	0.37	0.56	0.79	0.85	0.01	-0.01	0.02	-0.04	-0.01	0.02	0.03	0.02
ri	3.59	3.59	4.08	4.47	-0.04	-0.00	0.01	-0.02	-0.02	-0.01	0.01	0.01

Table 5.6: Distributional properties for the BDS statistics $W_m(\epsilon)$, and Savit and Green's dependability index $\delta_{m-1}(\epsilon)$ of the G, GS, and GI series, for $m = 2, 3, 4$, and 5, and with $\epsilon = \sigma$, the standard deviation of each time series. The α-percentile is denoted I_α. Note that the entries corresponding to the δ's are multiplied with 100, except for the $\text{power}_{0.05}$'s.

		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	minimum	2.99	3.28	4.23	4.69	1.39	0.16	0.25	0.06
	maximum	14.54	16.41	18.30	20.54	7.10	4.81	4.19	3.76
G	$I_{0.025}$	4.08	5.28	6.19	6.85	1.93	1.18	0.94	0.53
	$I_{0.975}$	10.81	13.23	14.96	16.74	5.22	4.07	3.33	2.96
	$\text{power}_{0.05}$	1.000	1.000	1.000	1.000	1.000	0.985	0.961	0.904
		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	minimum	2.93	3.48	4.17	4.86	1.34	0.26	0.37	-0.12
	maximum	14.14	15.98	17.30	19.24	6.92	4.60	4.09	3.75
GS	$I_{0.025}$	3.80	5.09	5.99	6.80	1.82	1.14	0.90	0.52
	$I_{0.975}$	10.48	12.84	14.99	16.48	5.10	4.02	3.35	3.00
	$\text{power}_{0.05}$	1.000	1.000	1.000	1.000	1.000	0.977	0.948	0.851
		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	minimum	2.18	3.25	4.06	4.31	1.16	0.05	0.37	-0.74
	maximum	13.70	18.16	22.28	26.37	9.81	9.66	8.46	7.62
GI	$I_{0.025}$	3.64	5.02	6.09	6.99	1.95	1.21	1.11	0.69
	$I_{0.975}$	10.63	13.62	16.30	18.67	7.19	6.07	5.31	4.64
	$\text{power}_{0.05}$	0.228	0.233	0.233	0.192	0.076	0.093	0.089	0.071

Table 5.7: Distributional properties for the BDS statistics $W_m(\epsilon)$, and Savit and Green's dependability index $\delta_{m-1}(\epsilon)$ of the S, SS, and SI series, for $m = 2, 3, 4$, and 5, and with $\epsilon = \sigma$, the standard deviation of each time series. The α-percentile is denoted I_α. Note that the entries corresponding to the δ's are multiplied with 100, *except* for the entries corresponding to skewness and kurtosis.

		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	mean	-0.02	-0.08	-0.09	-0.10	-0.01	-0.01	-0.01	-0.02
	std.dev.	0.87	0.75	0.68	0.63	0.40	0.45	0.49	0.55
	skewness	0.17	0.16	0.15	0.20	0.13	0.12	0.15	0.01
S	kurtosis	2.75	2.88	2.98	3.07	2.71	2.79	3.00	2.83
	minimum	-2.45	-2.07	-2.02	-1.82	-1.15	-1.30	-1.75	-1.78
	maximum	3.22	2.87	2.18	2.07	1.31	1.29	1.56	1.63
	$I_{0.025}$	-1.61	-1.46	-1.37	-1.30	-0.75	-0.93	-0.93	-1.08
	$I_{0.975}$	1.68	1.40	1.33	1.23	0.77	0.86	1.00	1.05
		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	mean	-0.01	-0.07	-0.09	-0.10	-0.01	-0.07	-0.02	-0.04
	std.dev.	0.88	0.77	0.69	0.64	0.41	0.45	0.50	0.54
	skewness	0.10	0.18	0.22	0.25	0.06	0.04	0.11	0.09
SS	kurtosis	2.83	3.02	3.16	3.19	2.79	2.73	3.04	2.87
	minimum	-2.31	-2.00	-2.22	-1.84	-1.14	-1.35	-1.67	-1.49
	maximum	3.38	3.31	2.72	2.49	1.41	1.17	1.55	1.77
	$I_{0.025}$	-1.71	-1.48	-1.37	-1.32	-0.79	-0.92	-0.95	-1.09
	$I_{0.975}$	1.69	1.43	1.31	1.20	0.79	0.82	1.02	1.02
		W_2	W_3	W_4	W_5	δ_1	δ_2	δ_3	δ_4
	mean	1.10	1.53	1.86	2.12	0.63	0.72	0.70	0.57
	std.dev.	1.63	1.95	2.23	2.48	0.99	1.08	1.03	0.99
	skewness	0.49	0.82	0.97	1.09	0.84	1.18	1.30	1.27
SI	kurtosis	3.33	3.44	3.62	3.87	4.72	4.53	5.08	5.74
	minimum	-4.21	-2.50	-2.21	-2.00	-3.77	-1.37	-1.36	-1.37
	maximum	6.89	8.32	10.61	12.71	4.31	5.23	4.95	5.68
	$I_{0.025}$	-1.65	-1.46	-1.17	-1.02	-0.80	-0.76	-0.69	-0.89
	$I_{0.975}$	4.85	6.24	7.25	8.29	3.11	3.43	3.44	3.05

Chapter 6

Stochastic Volatility and Pricing Bias in the Swedish OMX-Index Call Option Market

6.1 Introduction

An option is a derivative security and its value can, in principle, be determined if all underlying variables are specified. The Black-Scholes (1973) (henceforth "B-S") model is, of course, the outstanding model for this purpose. It is simple and elegant but builds on fairly restrictive assumptions, two of which, the constant stock return volatility and the constant interest rate, have been relaxed in a number of papers in the last decade.

Early studies of the Black-Scholes model and its pricing behavior include Macbeth and Merville (1979), Rubinstein (1985), and Evnine and Rudd (1985). In the case of option pricing with volatility modelled as a stochastic process, both stock and stock index options, (Hull and White (1987), Wiggins (1987), Scott (1987), Stein and Stein (1991), and Ball and Roma (1994)), and currency options (Chesney and Scott (1989), Melino and Turnball (1990), Heston (1993), and Bates (1996)) have been studied. There are also articles where the interest rate is assumed to be stochastic (Heston (1993), Amin and Ng (1993), and Saez (1995)). A common result is that an improvement in pricing (more efficient markets) follows with the inclusion of a stochastic volatility, while the impact of a stochastic interest rate seems less clear.

Several option-pricing models, with different assumptions regarding the return distribution of the underlying asset, have been developed; the vast majority of the models being based on continuous time stochastic processes and Ito calculus. When the model is specified, the option price must be solved for. Normally, this means solving a partial differential equation (PDE) and a number of methods are available. Whether direct numerical solving of the partial differential equation, Monte Carlo simulations, approximation methods, or a combination of numerical and analytical solution methods is used, depends on the kind of option to be priced as well as the processes chosen for the underlying assets. In addition, when introducing a non-traded underlying parameter like stochastic volatility, it is known from financial theory that a non-zero volatility risk premium must be introduced, which complicates the search for the option price, although not in a critical way.

With a randomly changing volatility, the option price is no longer determined by a single stochastic variable, the stock index price, but a second stochastic variable, the volatility of the stock index return, is equally important. We end up having two underlying stochastic processes, two state variables, that may be specified in different ways and may or may not be correlated. In this essay, a Geometric Brownian Motion is assumed for the stock index price and a mean-reverting Cox-Ingersoll-Ross (CIR) square-root process for the volatility (variance)

$$dS = mSdt + \sigma S dZ_1 \tag{1}$$

$$d\sigma^2 = \alpha(\theta - \sigma^2)dt + \xi\sqrt{\sigma^2}dZ_2, \tag{2}$$

where S is the stock index price, σ^2 is the volatility (variance) of the stock index return, α, θ, ξ, and m are constants, and $dZ_1$and dZ_2 are independent Wiener processes. The parameter α is the degree of mean reversion, θ is the long-run mean volatility, and ξ measures the volatility of the variance process. The choice of model for the volatility behavior is partly due to mathematical tractability where we can draw on interest rate theory and the bond pricing formula in Cox, Ingersoll and Ross (1985), and partly due to feasibility; empirically, volatility is never negative and it has a tendency to revert to a long-run average. Both these phenomena are covered by the mean-reverting square-root process.

To solve for the option price, I use the Feynman-Kac functional and the concept of risk neutrality, i.e. solving the PDE with a stochastic representation formula where the discounting is done with the risk-free rate of interest. In order to find the final stock index price distribution, I use the Fourier-Inversion method introduced by Stein and Stein (1991). They used this technique for the arithmetic Ornstein-Uhlenbeck process, and Ball and Roma (1994) modified the model for the CIR-process. In these studies, the prices given by the Fourier-Inversion model and the B-S model were compared but the pricing methods were not used to back out real-world parameters and biases. In this essay, the aim is to study the pricing bias in the Swedish OMX-Index call option market and the Fourier-Inversion model is used both to estimate volatility process parameters and to price options. While several empirical studies on stochastic volatility option pricing exist, most of these rely on Monte-Carlo methods to find the option price. By instead using the Fourier-Inversion method, I get a quicker and more flexible method. For comparison, I also calculate B-S prices in addition to stochastic volatility prices.

My choice of market is the Swedish OMX-Index option market and, to my knowledge, this is the first study applying stochastic volatility option pricing methods to this particular market. The OMX-Index option market is smaller than the bigger stock option markets in the US and the UK, but it has many interesting features. All trade is done with a computer system, the contracts are purely European style, and most important, there are no dividends in the market for the time periods studied.

To assess the stability of the results, I have looked at two separate time periods, October 1993 to February 1994 and July 1994 to December 1994. Both these periods are relatively tranquil, compared to, for instance, late 1992 when the Swedish krona was under pressure. Using data from these periods, I back out parameters for the stochastic processes and with these estimates as input, I try to judge how efficient option prices are quoted. The fast Fourier-Inversion method proves to be useful in making it possible to back out the risk neutral (Q-measure) parameters from quoted option prices (compare yield-curve inversion). This is a fairly new approach in the area of option pricing, where most authors use historical stock-return data and moment methods to estimate volatility parameters (Bakshi *et al.* (1996)). The method has the advantage of directly giving the risk neutral volatility parameters and giving possibilities to infer the sign and size of the volatility risk premium.

The bias study in this essay is divided into a static and a dynamic part. The static study is done on a daily basis and compares, out-of-sample, the market-, the B-S-, and the stochastic volatility-prices by daily updating model inputs. The dynamic efficiency test consists of a hedging scheme, where a hedged position in two call options and the underlying index is formed and daily updated with suitable Δ:s, and where risk-free arbitrage profits are calculated, *ex ante*.

The essay is organized as follows. Chapter 6.2 looks at the pricing model and the Fourier-Inversion technique. Chapter 6.3 describes the OMX-Index Option Market and estimates the volatility parameters. Chapter 6.4 contains the static bias study and the dynamic efficiency test. Finally, chapter 6.5 concludes the essay.

6.2 The Model

6.2.1 Equilibrium Pricing and the Stochastic Volatility Model

In the Black-Scholes model, the call option has a unique price. This is related to the fact that in the B-S model, every contingent claim can be replicated by a self-financed portfolio. In other words, the B-S model is complete.

In the stochastic volatility model, the situation is different. Since volatility is not spanned by assets in the economy, the volatility-risk cannot be eliminated by arbitrage methods. Instead, we must rely on equilibrium methods. It then follows that the market price of volatility risk explicitly enters the general partial differential equation for the option price:

$$\frac{\partial F}{\partial t}(t,s) + rs\frac{\partial F}{\partial s}(t,s) + \frac{1}{2}s^2\sigma^2(t,s)\frac{\partial^2 F}{\partial s^2}(t,s) + \tag{3}$$

$$\left[\alpha(\theta-\sigma^2)-\lambda\right]\frac{\partial F}{\partial(\sigma^2)}(t,s) + \frac{1}{2}\xi^2\sigma^2\frac{\partial^2 F}{\partial(\sigma^2)^2}(t,s) - rF(t,s) = 0$$

$$F(T,s) = \Phi(s),$$

where F is the option price, s is the price of the underlying asset (the OMX-index), σ^2 is the index return volatility, and finally, λ is the volatility risk premium.

The main difference between the present situation and the B-S setting is that in the B-S model, arbitrage methods are used to find the price, while here, equilibrium arguments are used. The option price will only be unique when supply and demand in the market are equalized, and the forces of supply and demand are, in turn, determined by such phenomena as risk aversion.

One way out is to find situations where the solution to the PDE is independent of risk preferences. This is the case if (a) the volatility is a traded asset or (b) the volatility is uncorrelated with aggregate consumption (Hull and White (1987)). An alternative way is to treat the volatility as the non-traded parameter it actually is but putting the risk premium equal to zero, which is done by Scott (1987) and Hull and White (1987).

The exact form of the risk premium might not be found and one might not be comfortable with assuming a zero risk premium. Then there is the special case of a non-zero constant risk premium for the volatility that does not actually change our solution method *or* the results in any profound way. In this essay, I will assume a non-zero constant risk premium, so that a risk adjusted drift rate for the volatility can be defined in (2). The drift rate $\alpha(\theta-\sigma^2)$ changes to $\alpha(\theta'-\sigma^2)$, where the only change is a shift in the constant long-run mean[1]. λ has now disappeared from (3) and the

[1]In the PDE, all parameters are assumed to be constant. If further assuming a constant risk premium λ, then the variable sustitution $\alpha(\theta-\sigma^2)-\lambda = \alpha(\theta-\frac{\lambda}{\alpha}-\sigma^2) = \alpha(\theta'-\sigma^2)$ can be made.

new parameters, α, θ', and ξ are called risk-adjusted parameters, or Q-parameters[2].

6.2.2 The Fourier-Inversion Technique and Stochastic Volatility Option Pricing

The Fourier Inversion method as a technique to find the stock price distribution was introduced by Stein and Stein (1991), who focused on the similarity between Moment Generating Functions (MGFs) and Fourier Transforms and combined this with the averaging over time of the stock price variance. Ball and Roma (1994) continued this work by even further emphasizing the important role of the average variance and, in particular, the MGF of the average variance, also showing its importance in other solution methods. Heston (1993) developed a slightly different model and suggested the use of a square-root process, which has the advantage of always giving positive volatilities and being familiar from earlier work by Cox, Ingersoll and Ross (1985) in the different, but related, context of bond pricing.

The closed-form solution of the option price in (5), $F(t)$, is derived by following Stein and Stein (1991) and Ball and Roma (1994) and by applying the Feynman-Kac functional (risk adjusted expectation)

$$F(t,S) = \frac{1}{e^{r(T-t)}} \int_X^{\infty} (S_T - X) f(S_T) dS_T$$

to the non-lognormal stock return distribution

$$f(S_T) = \frac{e^{\frac{m(T-t)}{2}}}{2\pi S_T^{\frac{3}{2}}} \int_{-\infty}^{\infty} I[\frac{(\eta^2 + \frac{1}{4})(T-t)}{2}] cos[(ln(S_T) - m(T-t))\eta] d\eta, \tag{4}$$

where (4) is derived by using the similarity between Moment Generating Functions and Fourier-Transforms. We end up with the following expression for the call option price, where S_T is the underlying stock index value at the exercise date, X is the

[2]Under the assumption of a constant risk premium and if the ordinary parameter θ can somehow be found, an estimate of the risk premium, λ, can also be obtained.

strike price, and η merely an integration variable:

$$F(t) = \frac{1}{2\pi S_T^{\frac{3}{2}} e^{\frac{r(T-t)}{2}}} \int_X^{\infty} \int_{-\infty}^{\infty} (S_T - X)\, I[\frac{(\eta^2 + \frac{1}{4})(T-t)}{2}] cos[(ln(S_T) - r(T-t))\eta] d\eta dS_T \tag{5}$$

where

$$I[\lambda] = \exp(N + M\sigma_0^2),$$

σ_0^2 is the initial variance, and N and M are the following functions:

$$N = \frac{2\alpha\theta'}{\xi^2} \ln \left[\frac{2\gamma e^{\frac{(\alpha-\gamma)(T-t)}{2}}}{g(T-t)} \right]$$

$$M = \frac{-2(1 - e^{-\gamma(T-t)})}{g(T-t)}$$

where

$$\gamma = \sqrt{\alpha^2 + 2\left(\frac{\lambda}{(T-t)}\right)\xi^2}$$

and

$$g(T-t) = 2\gamma + (\alpha - \gamma)\left(1 - e^{\gamma(T-t)}\right).$$

What remains is an integration giving the explicit solution. Unfortunately, any attempt to find a primitive function to this integrand seems bound to fail, and consequently, we must rely on the best possible approximation found by some kind of numerical integration. To solve this problem, a combination of Simpson's method and the simple Trapetzoid method is chosen, and all programs are written in the *GAUSS* programming language. The stochastic volatility option prices in (5) are then calculated with these programs.

6.3 The Swedish OMX-Index Option Market and Parameter Estimates

6.3.1 Data

In September 1986, the Swedish exchange for options and other derivative securities (OM) introduced the OMX-index. It consists of a value-weighted combination of the 30 most actively traded stocks on the Stockholm Stock Exchange. The purpose of the introduction was for the OMX-index to serve as an underlying "security" for trading in standardized European style options and forward contracts. A unique feature of the Swedish stock-index options, at least compared to US markets, is that during a large part of the year, there are no dividends at all. The OMX-index must be adjusted for dividends only when the April-, May-, June- and July- option contracts are analyzed. This essay looks at dividend-free August to March contracts.

The OMX-index Option Market consists of European style Call- as well as Put-Options with different times to expiration. At any time throughout the year, trading is possible in at least three classes of option contracts with up to one, two and three months left to expiration, respectively. On the fourth Friday each month, when the exchange is open for trading, one class of contracts expires and a new class, with time to expiration equal to three months, is initiated. Furthermore, for options with a given time to expiration, a wide range of exercise prices is available. When options with a new expiration date are introduced, the exercise prices are chosen so that they are centered around the current value of the OMX-index.

The set of data used consists of daily closing bid and ask quotes for the two time periods, October 1993 to February 1994, and July 1994 to December 1994. The option data and index data are obtained from OM and contain both prices and volumes. Options with a time to maturity shorter than 15 days, as well as options with very low liquidity, are removed from the sample. Certain days (very few) have been removed from the data set due to errors in the data (non-feasible prices, missing bid or ask quotations, erroneous strike prices, etc.) and the total number of observations is 1694. Both the options exchange (OM) and the stock exchange (StSE) close at 4.00

p.m., minimizing the possibility of synchronization problems. Interest rates for 30, 60, and 90 days are obtained from Sveriges Riksbank, and the relevant interest rates are computed by interpolation between the two closest interest rates.

6.3.2 Parameter Estimation

The next step is to estimate the parameters in the volatility process and in the PDE. Most empirical research shows that volatility follows a mean-reverting process, like the one in (2), which gives three parameters to be determined: the reversion rate, α, the long-run mean, θ', and the volatility of the volatility, ξ. The volatility is assumed to start at its long-run level.

There are many alternative ways of estimating these parameters. An approach using the discrete time approximation of continuous time stochastic processes is the method of moments by Chesney and Scott (1989) and Hansen (1982). By looking at the moments of the stock return distribution, estimates of the discrete time process parameters as well as the continuous time parameters can be found.

If choosing to work directly with the continuous time processes, determining the distribution of stock returns as a function of the parameters in question and then applying maximum likelihood methods would be a natural approach. The problem with this approach is that stock returns are dependent over time, and the joint distribution for a sample of observations would be very difficult to derive, Scott (1987).

The estimation approach chosen here is that of calibrating the model to data. This is comparable to the way Brown and Schaefer (1994) fit the "yield curve" of bonds, and this method has the advantage of directly giving the Q- parameters (Martingale), not the objectively observed parameters.

The calibrating technique works as follows. I choose to model the stock return as a Geometric Brownian Motion and the volatility process as a mean reverting square-root process. By specifying these processes under the Q-measure, option prices can be calculated with the Fourier-Inversion method as functions of the volatility parameter set Ω, where Ω is defined as $\Omega = \{\alpha, \theta', \xi\}$. Using empirical data, I calculate the midpoint between daily quoted bid and ask OMX-Index option prices and compare these midpoint values to the modelled option prices. On any given day, I minimize the sum of the squared differences between the model prices and the empirical midpoint prices to get estimates of the parameters, Ω:

$$\min_{\Omega} SSE = \min_{\Omega} \sum_{i=1}^{N} \left(F_{\text{model,i}}(\Omega) - F_{\text{market,i}} \right)^2, \tag{6}$$

where N is the number of options on a particular day. Since the stochastic volatility option price, $F_{\text{model,i}}(\Omega)$, is highly non-linear in its parameters, the problem (6) is a non-linear least square minimization problem. Thanks to the Fourier-Inversion method, the least square minimization method of estimating the parameters is not too costly in terms of computer resources and fully compiled computer languages are not necessary to implement the routine.

The problem has been implemented in *GAUSS* and the calculations were made on a Pentium 100MHz PC. At the heart of the computation lies an integration routine and for the nonlinear least square minimization, the Gauss-Newton algorithm with numerically calculated derivatives is used. The generalized double integral is truncated to a finite region without too much loss of information. In addition, the high non-linearity creates a number of local minima, and a range of initial parameter-values has to be tried as inputs to find the global minimum.

Since I look at two separate time periods I can, to some extent, assess the stability of the estimates. Running the program each day in the sample periods gives around 250 estimates each of α, θ', and ξ. Studying how the parameters change over time reveals some time dependency but the model assumption of constant parameters does not seem very strong. Average parameter values and average asymptotic standard deviations (from the covariance matrix) are given in Table 6.1.[3]

In Table 6.1, I also include a calculation of the B-S implicit volatility. This volatility is calculated by minimizing the squared difference between the market price and

[3]The high asymptotic standard deviation for α might seem to be a subject of concern. It is due to the very low second derivative of the price function F with respect to α; a large variation in α gives only a slight variation in F. However, since the sensitivity of F with respect to α is very small, predictions of future option prices do not critically depend on the estimated value of α.

the Black-Scholes price for all options traded on a particular day, which is different from the usual approach of using at-the-money options only. Our procedure gives a worse correspondence with empirical prices at-the-money but gives a lower pricing bias in-the-money and out-of-the-money[4].

6.4 Pricing Bias

6.4.1 Static Bias

This chapter looks at the out-of-sample pricing bias in the Swedish OMX-Index call option market. I have chosen the previous day's (yesterday's) estimates of process parameters as inputs to compute the current day's (today's) model price. These model prices will then be compared to actual market prices and a possible bias is studied. This approach is supposed to replicate the behavior of practitioners and one should be aware of its tendency to favor the shortsighted B-S model; the stochastic volatility model works much better than the B-S model in the unrealistic setting of no (or not very frequent) updating of parameters. Next, the observed market price is subtracted from the model price to compute both the percentage pricing error and the absolute percentage pricing error. This is repeated for all different call options each day in the sample, both for the stochastic volatility model and the B-S model.

The percentage error is defined as

$$e = \frac{100\,(P_{\text{model}}(\Omega) - P_{\text{market}})}{P_{\text{model}}(\Omega)},$$

and moneyness is defined as

$$m = \frac{100 \cdot \left(OMX\text{-}index\ value - Strikeprice \cdot e^{-r(T-t)}\right)}{Strikeprice \cdot e^{-r(T-t)}}$$

[4] I have also tried the usual approach of only inverting options at-the-money. The results are not reported here but the two average estimates are fairly similar, even though substantial differences are present on particular days.

Far-out-of-the-money is defined as $m < -5$, out-of-the-money as $-5 \leq m < -1$, at-the-money as $-1 \leq m < 1$, in-the-money as $1 \leq m < 5$, and finally deep-in-the-money as $m \geq 5$.

If the theoretical prices are compared in a scatter plot (with empirical parameters estimated one day earlier) a smile-shaped bias structure is found between B-S prices and stochastic volatility prices, Figure 6.1.

This smile is predicted by theory and is due to the convexity properties of the B-S model (as a function of the mean variance over the life of the derivative security). Jensen's inequality says that, if F is concave, $E[F(\cdot)] < F(E[\cdot])$ where E is the expectation operator (when F is convex the opposite holds). This, together with the fact that (5) can be seen as an expectation of the B-S price over different mean variances and that the B-S price, as a function of the mean variance, is convex for large (and small) values of S/X, and concave for values of S/X close to one, where, as before, S is the stock index price and X is the strike price, gives the observed bias structure.

Figure 6.2 is a scatter plot of the bias between stochastic volatility prices and market prices as a function of moneyness. It can be seen how the options are overpriced out-of-the-money and slightly underpriced in-the-money compared to the market valuation. The plot for the B-S model shows a similar bias structure. For a more quantitative analysis of the pricing bias, we can look at the results in Tables 6.2 and 6.3. In these tables, the results for the different pricing models are reported, divided into different times to maturity as well as different levels of moneyness. The tables show how the two models demonstrate percentage errors of a similar magnitude, as well as a similar bias structure. This behavior is stable over different time periods as well as different pricing models[5]; overpricing out-of-the-money, underpricing in-the-money, and somewhat better correspondence at-the-money. For both models, the prices deviate significantly from the market price out-of-the-money and in-the-money and the percentage errors (with sign) are significantly different from zero.

[5]The results for the individual time periods are not presented, but looking at these two time periods, October 1993 to February 1994 and July 1994 to December 1994, separately, there is no significant change in patterns.

The absolute percentage errors confirm this evidence of mispricing and both models give significant pricing errors. Overall, the absolute percentage error decreases with moneyness, ranging from around 15% out-of-the-money to 3% in-the-money.

In addition to the strike price bias above, a quick, qualitative look at the time to maturity bias is also interesting. As a whole, the absolute percentage bias decreases with time to maturity. In particular, the options with the longest times to maturity have a smaller bias than the options with shorter maturity. No explanation to this behavior has been found.

For all maturities, both models systematically overprice out-of-the-money and underprice in-the-money. An explanation of this skewness might be a negative correlation between the stock return and the volatility taken into account by the market but not by the models. This would lead to an overvaluation by the models of out-of-the-money options, since the stock return distribution becomes negatively skewed and really high option prices are less likely to be achieved; when the stock price increases, volatility tends to decrease, thereby making large movements in price less likely. The opposite holds when stock prices decrease. On the other hand, several studies have also shown that the skewness is *not* significant in the equity markets.

The results are confirmed by many studies, both for the constant volatility B-S model and the stochastic volatility models. One notable finding, however, is the difference between the B-S prices in my study and the B-S prices in the study by Hansson *et al.* (1995) in the same market. Hansson *et al.* find out-of-the-money options to be better priced than at-the-money options and only slightly underpriced compared to the market price. This is somewhat surprising, since they use at-the-money options to back out the implied volatility. The only explanation for the difference in results is the different specification of the implicit volatility and the different data sets[6].

To summarize, the difference between the theoretical B-S price and the stochastic volatility price is in accordance with theory and shows the expected smile-shaped bias

[6]Even if not presented in my paper, it is interesting to notice that the B-S model with at-the-money-estimated implicit volatility shows substantially worse pricing behavior than the other two models out-of-money and in-the-money. This is not very surprising, considering the well-known volatility smile existing in implicit B-S volatilities.

structure. The difference between market prices and model prices is larger, though, and both models overprice out-of-the-money and underprice in-the-money. To fully evaluate the pricing performance of the models, we turn to a dynamic test measuring riskfree profits over time by holding, and each day rebalancing, risk-free portfolios designed with each of the models.

6.4.2 Dynamic Efficiency Test

If the B-S price is closer to the "correct price" than is the market price, then it should be possible to make riskfree arbitrage profits by trading with the B-S model. Further, since the variance is observed to vary randomly, a trader using a random variance model may be even more efficient in identifying mispriced options. To test this hypothesis, I compute *ex ante* net gains from a hedged position of options and the underlying variables[7]. In the B-S case, I use a standard Δ-hedge with a call option and a hedged position in the OMX-Index (or the stocks making up the index). In the random variance case, both the OMX-Index and the volatility must be hedged, which is accomplished by taking positions in two call options as well as the underlying index. In this way, both the random sources in the option pricing model are hedged (Chesney and Scott (1989)).

The dynamic Δ-neutral hedge for the stochastic volatility model is created in the following way. Each day, I take a position in the option that is most mispriced. If the model price is higher than the midpoint of the bid-ask spread, then I buy the option, if it is lower, I sell it. The position in the second option must be the opposite in order to hedge the volatility risk, and if I need to sell the second option, I choose one with a model price below the midpoint price. Finally, a position is taken in the OMX-Index.

The hedged position is

$$F(S, \sigma, t, K_1) + w_1 S_t + w_2 F(S, \sigma, t, K_2),$$

[7] *Ex ante* means that model prices at time t are calculated by using parameters estimated at time $t-1$.

where w_1 and w_2 are

$$\begin{aligned} w_1 &= -\frac{\partial F(S,\sigma,t,K_1)}{\partial S} - w_2 \frac{\partial F(S,\sigma,t,K_2)}{\partial S} \\ w_2 &= -\frac{\frac{\partial F\left(S,\sigma,t,K_1\right)}{\partial \sigma^2}}{\frac{\partial F\left(S,\sigma,t,K_2\right)}{\partial \sigma^2}}. \end{aligned}$$

Each day the net gain on this hedge is calculated:

$$[F(t_2,K_1) - F(t_1 - K_1)] + w_1 \left[S_{t_2} - S_{t_1}\right] + w_2 \left[F(t_2,K_2) - F(t_1,K_2)\right] - r_{t_2 - t_1}\left[F(t_1,K_1) + w_1 S_{t_1} + w_2 F(t_1,K_2)\right].$$

Every transaction is made at the midpoint price in an attempt to exclude transaction costs, due to the bid/ask spread. For both the B-S model and the stochastic volatility model, positive *ex ante* average profits from using the trading rules are found.

Table 6.4 shows means, medians, and standard deviations for the B-S model and the stochastic volatility model for the two different time periods. For both time periods, the stochastic volatility model gives higher profits than the B-S model, whose profits are not significant. The existence of these profits indicates mispricing in the OMX-Index market, even though it is important to notice that my hedging scheme assumes that all trade can be done within the bid-ask spread and without transaction costs. In practice, this is the case for large traders and market-makers only.

6.5 Conclusions

The standard Black-Scholes model for pricing European call options assumes a lognormal probability distribution for the underlying stock-index price and a constant stock-index return volatility. Considering empirical evidence, a more plausible hypothesis is that volatility changes randomly.

In this essay, I specify the volatility process as a mean-reverting square-root process and calculate theoretical option prices with the Fourier-Inversion Technique. The option pricing equation contains a preference term, the volatility risk premium, as volatility is a non-traded asset. It is not obvious, a priori, what constitutes a reasonable value for the price of volatility risk and I have chosen to treat the risk-premium as a constant. In this way, the Fourier-Inversion method can be used.

I quote actual prices on dividend-free European call options from the Swedish OMX-Index Option Market, and from these market prices, I estimate daily volatility process parameters by a non-linear least square minimization of the difference between market and model prices. This procedure has the advantage of directly giving the risk-neutral parameters.

From a static point of view, I find a smile-shaped bias between the Black-Scholes prices and the stochastic volatility prices. Both models give prices showing a similar bias compared to actual prices quoted in the market and both models price options in-the-money and at-the-money more accurately than out-of-the-money. The absolute percentage bias decreases with time to maturity for both models.

The dynamic hedging test reveals riskfree arbitrage profit possibilities, at least for the stochastic volatility model, which supports the existence of mispricing in the OMX-Index Option Market, when transaction costs are not considered.

My conclusion is that the stochastic volatility model dominates the standard Black-Scholes model and produces a more efficient market. Considering the easy implementation of the stochastic volatility pricing model, this model is seen as an alternative to the established Black-Scholes model in actual pricing.

For further research, I suggest a comparative study on stock- and currency options and exchanges versus OTC-trading, and a more thorough study of the risk premium, for instance in the context of different macroeconomics situations (the 1992 currency crisis etc.). The application of the stochastic volatility model in related areas might also prove to be useful.

References

Amin, K.I., and Ng, V.K. (1993), "Option Valuation with Systematic Stochastic Volatility", *The Journal of Finance*, Vol.48, No.3, 881-906.

Bakshi, G., Cao, C., and Chen, Z. (1996), "Option Pricing and Hedging Performance under Stochastic Volatility and Stochastic Interest Rates", *Working Paper, Department of Economics, University of Maryland.*

Ball, C.A., and Roma, A. (1994), "Stochastic Volatility Option Pricing", *Journal of Financial and Quantitative Analysis*, 29, 589-60.

Bates, D.S. (1996), "Jumps and Stochastic Volatility: Exchange Rate Processes Implied in Deutsche Mark Options", *The Review of Financial Studies*, 9, No. 1, 69-107.

Black, F., and Scholes M. (1973), "The Pricing of Options and Corporate Liabilities", *Journal of Political Economics,* 81, 637-659.

Brown, R.H., and Schaefer S.M. (1994), "The Term Structure of Real Interest Rates and the Cox, Ingersoll, and Ross Model", *Journal of Financial Economics,* 35, 3-42.

Chesney, M., and Scott L. (1989), "Pricing European Currency Options: A Comparison of the Modified B-S Model and a Random Variance Model", *Journal of Financial and Quantitative Analysis*, 24, 267-284.

Cox, J., Ingersoll Jr. J., and Ross S. (1985), "A Theory of the Term Structure of Interest Rates", *Econometrica,* 53, 385-407.

Evnine, J., and Rudd, A. (1985), "Index Options: The Early Evidence", *Journal of Finance,* 40, 743-763.

Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50, 1029-1054.

Hansson, B., Hordahl, P., and Nordén, L. (1995), "Empirical Evidence of Biases in the Black-Scholes Option Pricing Formula", *Working Paper, Department of Economics, Lund University.*

Heston, S. L. (1993), "A Closed-form Solution for Option Prices with Stochastic Volatility with Applications to Bond and Currency Options", *Review of Financial Studies,* 6, 327-343.

Hull, J., and White, A (1987), "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance*, 42, 281-300.

Hull, J., and White A. (1988), "An Analysis of the Bias in Option Pricing Caused by a Stochastic Volatility", *Advances in Futures and Options Research*, 3, 29-61.

Macbeth, J., and Merville L. (1979), "An Empirical Examination of the Black-Scholes Call Option Pricing Model", *Journal of Finance*, 34, 1173-1186.

Melino, A., and Turnbull M. (1990), "Pricing Foreign Currency Options With Stochastic Volatility", *Journal of Econometrics,* 45, 239-265.

Rubinstein, M. (1985), "Nonparametric Tests of Alternative Option Pricing models Using all Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31", *Journal of Finance*, 40, 455-480.

Saez, M. (1995), "Option Pricing under Stochastic Volatility and Stochastic Interest Rate in the Spanish Case", *Economics Working Paper, Universitat Pompeu Fabra.*

Scott, L.O. (1987), "Option Pricing When the Variance Changes Randomly: Theory, Estimation, and an Application", *Journal of Financial and Quantitative Analysis*, 22, 419-438.

Stein, E.M., and Stein J.C. (1991), "Stock Price Distributions with Stochastic Volatility: An Analytical Approach", *Review of Financial Studies*, 4, 727-752.

Wiggins, J. B. (1987), "Option Values under Stochastic Volatility", *Journal of Financial Economics*, 19, 351-372.

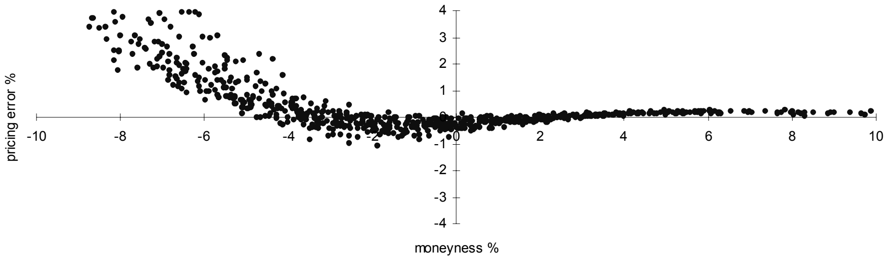

Figure 6.1: % Bias—Stochastic Volatility Prices minus B-S Prices. Both time periods (1649 obs.).

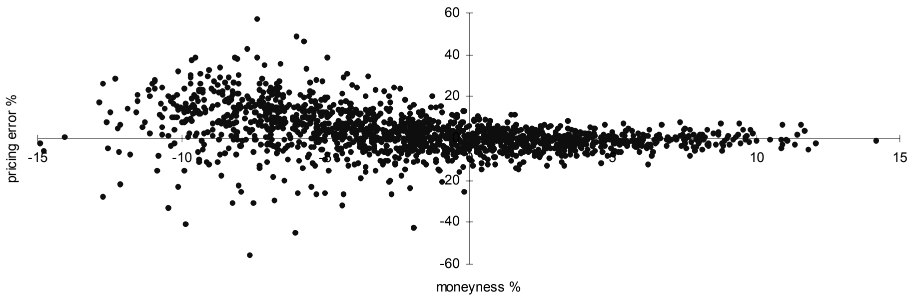

Figure 6.2: % Bias——Stochastic Volatility Prices minus Market Prices. Both time periods (1649 obs.).

.

Table 6.1: Average volatility parameters.

	Oct. 93 to Feb. 94		Jul. 94 to Dec. 94		Both Periods	
	Parameter	Std.	Parameter	Std.	Parameter	Std.
θ'	0.058	0.002	0.037	0.001	0.048	0.001
α	3.96	24.71	4.13	18.39	4.04	21.55
ξ	0.51	0.46	0.39	0.26	0.45	0.36
B-S Impl. Vol.	0.076	0.001	0.036	0.001	0.056	0.001

Parameter = average parameter value over the time period. Std. = average value over the time period of the standard deviation coming from the asymptotic covariance matrix for the parameter estimates.

Table 6.2: Stochastic Volatility average pricing errors. Both time periods (the total number of options is 1649).

	T-t	15-30 days	31-45 days	46-60 days	all days
	Average % Error	7.17	9.25	8.84	8.81
		2.23	0.95	0.93	0.68
Far Out	Average Abs. % Error	14.73	14.02	12.39	13.61
		1.68	0.67	0.66	0.48
	No. of Options	89	256	156	501
	Average % Error	3.51	2.67	1.10	2.08
		2.15	0.61	0.61	0.47
Out	Average Abs. % Error	11.73	7.30	5.93	7.29
		1.54	0.42	0.37	0.33
	No. of Options	75	258	128	461
	Average % Error	1.02	-0.42	-1.90	-0.75
		1.51	0.60	0.61	0.44
At	Average Abs. % Error	6.38	5.18	4.22	5.01
		1.01	0.36	0.38	0.27
	No. of Options	36	123	62	221
	Average % Error	-0.33	-1.58	-2.30	-1.60
		0.75	0.34	0.43	0.26
In	Average Abs. % Error	3.96	3.64	3.48	3.61
		0.50	0.23	0.30	0.17
	No. of Options	454	176	73	303
	Average % Error	-1.39	-1.12	-0.68	-1.12
		0.53	0.31	0.39	0.24
Deep In	Average Abs. % Error	2.91	2.93	1.59	2.73
		0.35	0.23	0.23	0.17
	No. of Options	44	99	20	163

Small numbers are standard deviations.

Table 6.3: Black-Scholes average pricing errors. Both time periods (the total number of options is 1649).

	T-t	15-30 days	31-45 days	46-60 days	all days
	Average % Error	-8.86	5.14	5.83	1.16
		3.56	1.09	1.02	0.99
Far Out	Average Abs. % Error	23.57	13.73	11.26	15.55
		2.72	0.74	0.90	0.73
	No. of Options	89	256	156	501
	Average % Error	3.86	2.82	0.58	2.29
		1.73	0.58	0.61	0.47
Out	Average Abs. % Error	11.27	7.33	5.60	7.42
		1.21	0.40	0.38	0.33
	No. of Options	75	258	128	461
	Average % Error	2.15	-0.12	-1.68	-0.38
		1.18	0.59	0.61	0.42
At	Average Abs. % Error	6.08	5.20	3.27	4.97
		0.79	0.35	0.37	0.26
	No. of Options	36	123	62	221
	Average % Error	0.10	-1.54	2.31	-1.51
		0.64	0.34	0.43	0.25
In	Average Abs. % Error	3.77	3.65	3.47	3.59
		0.43	0.23	0.30	0.17
	No. of Options	454	176	73	303
	Average % Error	-1.47	-1.29	-1.41	-1.31
		0.46	0.31	0.66	0.29
Deep In	Average Abs. % Error	2.73	3.04	2.05	2.79
		0.33	0.23	0.31	0.19
	No. of Options	44	99	20	163

Small numbers are standard deviations.

Table 6.4: Risk-free daily profits (SEK per option and day) from using a dynamic trading rule.

			Oct. 93 to Feb. 94	July 94 to Dec. 94	Both Periods
		Mean	0.43	0.39	0.41
	Stoc. Volatility	Median	0.18	0.37	0.29
		Mean Std.	0.21	0.33	0.20
Profit					
		Mean	0.054	0.14	0.090
	Black-Scholes	Median	0.023	0.19	0.16
		Mean Std.	0.37	0.10	0.16