Multifractal Models, Intertrade Durations And Return Volatility

Mawuli Kouami Segnon
Department of Economics
Christian-Albrechts University of Kiel

A thesis submitted for the degree of
Dr. Sc. Pol.

Multifractal Models, Intertrade Durations And Return Volatility

Inaugural-Dissertation
zur Erlangung des akademischen Grades eines Doktors
der Wirtschafts- und Sozialwissenschaften
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Dipl. -Kfm **Mawuli Kouami Segnon**
aus Kponou, Togo

Kiel, June 17, 2015

Christian-Albrechts-Universität zu Kiel, Leibnizstraße 3, 24118 Kiel

Gedruckt mit Genehmigung der
Wirtschafts- und Sozialwissenschaften
Fakultät der Christian-Albrechts-Universität
zu Kiel

Dekan:
Prof. Dr. Achim Walter

Erstberichterstattender:
Prof. Dr. Thomas Lux

Zweitberichterstattender:
Prof. Dr. Markus Haas

Tag der Abgabe der Arbeit: 17. März 2015

Tag der mündlichen Prüfung: 10. June 2015

Supervisors

Prof. Dr. Thomas Lux
Institute for Economics, University of Kiel
Chair of Monetary Economics and
International Financial Markets
Olshausenstrasse 40, 24118 Kiel, Germany,
Tel: ++49 431 880 3661
E-mail: thomas.lux@bwl.uni-kiel.de

Prof. Dr. Markus Haas
Institute for Quantitative Business and
Economics Research, University of Kiel
Chair of Empirical Economic Research
Olshausenstrasse 40, 24118 Kiel, Germany,
Tel: ++49 431 880 5505
E-mail: markus.haas@qber.uni-kiel.de

Dedicated to

Jeanne, Claire, and the memory of my dad Françis.

Contents

List of Tables

List of Figures

Notation and Abbreviations

General Symbols

$\approx$	is distributed as
$\equiv$	equals by definition
$\overset{d}{=}$	equality in distribution
$\sim$	is distributed as
$\sum$	summation sign
$\prod$	product sign
min	minimum
max	maximum
ln	natural logarithm
exp	exponential function
$\rightarrow$	converges to or approaches
$\mathbb{N}$	natural numbers
$\mathbb{Z}$	integer numbers
$\mathbb{R}$	real numbers
$\mathbb{R}_+$	positive real numbers
$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{E}$	expectation
$\mathbb{V}ar$	variance
$\mathbb{C}ov$	covariance
L	lag operator
Δ	differencing operator
Pr	probability
$\Im$	information set available
$\xrightarrow{d}$	converges in distribution to
sup	Supremum function

KS	Kolmogorov-Smirnov
AD	Anderson Darling
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
GMM	Generalized Moments Method
H	Hurst index
ISIVaR	irregularly spaced intraday value at risk
KS	Anderson-Darling distance
L	log-likelihood
LR	likelihood ratio test
ML	maximum Likelihood
VaR	value-at-risk

Distributions

LN	Lognormal distribution
$\chi^2(p)$	χ^2-distribution with p degrees of freedom
p-value	tail probability of a statistic

Abbreviations

ACD	autoregressive conditional duration
ARMA	autoregressive moving average
ASX	Australian Stock Exchange
BACD	Burr autoregressive conditional duration
EFIACD	exponential fractionally integrated autoregressive conditional duration
FX	foreign exchange
GARCH	foreign exchange
GGACD	generalized gamma autoregressive conditional duration
LACD	Lognormal autoregressive conditional duration
Log-ACD	logarithm autoregressive conditional duration
MF	multifractal model
MSM	Markov switching multifractal
NYSE	New York Stock Exchange
PMM	Poisson multifractal model
PSE	Paris Stock Exchange
TAQ	Trade and Quotes
TAR	threshold autoregressive
TARMA	threshold autoregressive moving average
WACD	Weibull autoregressive conditional duration

Acknowledgements

First of all, I would like to thank my advisor Professor Dr. Thomas Lux for his guidance and continuous support during my doctoral studies. I gained a lot from his numerous valuable suggestions and discussions during the last three years. He drew many research questions to my attention that makes the writing of this thesis easier. I am very fortunate to work with Professor Lux and thankful for his advice and help in the scholarship and Job search process. I enjoyed very much learning his philosophy of doing research, his serenity of approaching problems. I am grateful to Professor Dr. Markus Haas who accepts to be my second advisor. I am deeply indebted to Dr. Raddant and Dipl. Math. Penner for carefully reading of the thesis. My thanks go to the Secretary Brechtefeld, to everyone at the chair of "*Monetary Economics and International Financial Markets*" who have helped me during my postgraduate studies. I also want to thank Dr. Liu for providing me the Gauss-code for the bivariate multifractal model.

I wish to thank the doctoral Programme "*Quantitative Economics*" at the Christian-Albrechts-University of Kiel and das *Evangelische Studienwerk e.v. Villigst* for providing me financial support during the last four years. I express my sincerely thanks to my mum Jeanne, my brother Mawussey, and my girlfriend Claire. This thesis could not have been written without their love and encouragement. Finally, I would also like to thank Gabriell and Tobi, Vera and Hans-Meinert and Nadou and Ayo for their encouragement and support during my study at the Christian-Albrechts-University of Kiel.

Le marché, a son insu, obéit à une loi qui le domine: la loi de la probabilité.

Louis Bachelier, Théorie de la spéculation

Depuis plus d'un siècle, les financiers et les économistes se sont efforcés d'analyser le risque dans les marchés financiers, de l'expliquer, de le quantifier et, en définitive, d'en tirer un bénéfice. Ma conviction est que la route suivie par la plupart des théoriciens est mauvaise et qu'elle conduit à une grave sous-estimation des risques de ruine financière dans une économie de marché libre et globale.

Benoît Mandelbrot, 2005

Si vous pouvez mettre en évidence certaines propriétés du marché qui demeurent constantes dans le temps ou l'espace, vous pourrez élaborer de meilleurs modèles, plus utilisables, et prendre des décisions financières plus sensées. Mon modèle multifractal n'a besoin pour fonctionner que d'un ensemble de paramètres cohérents.

Benoît Mandelbrot, 2005

Abstract

This thesis covers the application of multifractal processes in modeling financial time series. It aims to demonstrate the capacity and the robustness of the multifractal processes to better model return volatility and ultra high frequency financial data than both the generalized autoregressive conditional heteroscedasticity (GARCH)-type and autoregressive conditional duration (ACD) models currently used in research and practice. The thesis is comprised of four main parts that particularize the different procedures and the main findings.

In the first part of the thesis we first delineate the genesis of multifractal (MF) measures and processes and how one can construct a simple MF measure. We outline the generic properties of the MF processes, mention how they motivate financial time series models, and present the different tools developed for the estimation of the MF models and the forecasting of return volatilities and some empirical results. Second, we give a short overview of both autoregressive conditional duration (ACD) models and Markov switching multifractal duration (MSMD) models. We start with some theoretical microstructure literature that motivate both models. We present ACD and MSMD models and their subsequent extensions. Finally, we cite the different diagnostic tests developed in the literature for assessing their adequacy and provide some prominent empirical studies.

The second part deals with the application the Markov-switching multifractal (MSM) model and generalized autoregressive conditional heteroscedasticity (GARCH) type models in forecasting crude oil price volatility. Based on six different loss functions and by means of the superior predictive ability (SPA) test of Hansen (2005) we evaluate and compare their forecasting performance at short- and long-horizons. The results give evidence that none of our volatility models can outperform other models across all six different loss functions. However, the long memory GARCH-type models and the MSM model seem to be more appropriate in terms of fitting and forecasting oil price volatility. We also found that forecast combinations of long memory GARCH-type models and the MSM lead to an improvement in forecasting crude oil price volatility.

The third and longest part of the thesis compares the predictive ability of the Markov switching multifractal duration (MSMD) model recently introduced by Chen et al. (2013) to those of the standard ACD (cf. Engle and Russell, 1998), Log-ACD (cf. Bauwens and Giot, 2000), and fractionally integrated ACD (FIACD) (cf. Jasiak, 1998) models. We assume that innovations in the ACD and Log-ACD models follow Weibull, Burr, generalized gamma and Lognormal distributions. For FIACD we only consider the case where the innovation is standard exponentially distributed. We assess the forecasting performance of the models using density forecasts evaluation

methodologies proposed by Diebold et al. (1998) and the likelihood ratio test of Berkowitz (2001). We complement these methodologies with Kolmogorov-Smirnov and Anderson-Darling distances (cf. Rachev and Mittnik, 2000). Empirically, results are quite nice and speak for the MSMD model. In fact, the MSMD model can better capture the long memory and the fat tails observed in trade and price duration data, and therefore, outperforms both the FIACD, ACD and Log-ACD models. We also found that certain distributional assumptions for the innovations strongly enhance the forecasting performance of the ACD and Log-ACD models.

In line with the last result, we want to know to what extent different distributional assumptions for the innovation in the MSMD model may influence the model's forecasting performance. So, we assume that the innovation in the MSMD model follows generalized gamma or Burr distribution. To compare and select the model that provides better fit to the empirical data (trade, price and volume durations) we make use of the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and the likelihood ratio test. Surprisingly, both distributional assumptions for the innovation do not much affect the predictive ability of the model. It seems that the ability of the MSMD model to fit financial duration data largely stems from the multifractal processes.

Third, we generalize the univariate MSMD model to a bivariate one. The bivariate MSMD model is substantially an adaptation of the bivariate Markov switching multifractal (MSM) process proposed by Calvet et al. (2006) to high frequency financial data. We apply the bivariate MSMD model to analyze the co-movement between the bid-ask spreads of different stocks. The results indicate that bid-ask spreads of sector-specific or cross-sector stocks may be simultaneously affected by arrival of information in the market.

Fourth, we apply the standard MSMD and the generalized gamma ACD (GGACD) models to forecast irregularly spaced intra-day value-at-risk (ISIVaR) in a semi-parametric framework. We assess the performance of both models to produce accurate irregularly spaced intra-day VaR via the generalized moments method (GMM) duration-based test developed by Candelon et al. (2011). The results show that the MSMD model outperforms the GGACD model and can be used in practice to manage market risk.

The last part summarizes the main findings of the thesis and presents some outlooks for future research.

Preface

This thesis is a collection of essays on the application of multifractal processes for modeling financial markets data, especially crude oil prices, financial intertrade durations and bid-ask spread data. The essays have been assembled in book format with four parts. Part I: Review Of Multifractal And Autoregressive Conditional Duration Models contains chapter 2 and 3. Chapter 2: Multifractal Models In Finance: Their Origin, Properties, and Applications (with Thomas Lux) is prepared as a chapter for: Shu-Heng Chen and Mak Kaboudan. Forthcoming OUP Handbook on Computational Economics and Finance, Oxford University Press. All other chapters are working papers. As in the case of the chapter 1 that has been already accepted for publication, all other chapters are in course of preparation to be published in refereed journals. Note that chapter 4: Modeling and Forecasting Crude Oil Price Volatility: Evidence from Historical and Recent Data is written with Thomas Lux and Rangan Gupta and submitted to Energy Economics.

The analysis of the thesis is carried out using two programming languages, namely Matlab version 7.11 for adjusting raw data, drawing figures and for the estimation of the GARCH-type models and Gauss version 11 for the estimation of the remaining models.

1. Introduction

1.1. Motivation

The concern to understand the behavior of stock market prices and to propose a model that can reproduce their time evolution started with Bachelier's PhD thesis. Bachelier (1900) proposed a mathematical model that is now called "standard Geometric Brownian motion" and unintentionally anticipated the concept of market efficiency. In the years following his work there was a great number of empirical work on stock market prices. Examples include Cowles (1933), Working (1934), Cowles and Jones (1937), Kendall (1953), Roberts (1959), Fama (1965), among others. These research confirmed Bachelier (1900)'s findings and supported the random walk model.

However, in the earlier sixties Mandelbrot (1963) demonstrates that fluctuations of cotton prices exhibit *fat tails* and *clustering*, and thus, cannot be reproduced by Bachelier's model. Mandelbrot's findings lead to the development of the efficient market hypothesis (EMH) by Fama (1970) and triggered the discussion as to whether or not financial data exhibit such properties. Mandelbrot's work unleashed new research activities that consist in closely scrutinizing empirical financial data. These research activities have been intensified with the availability of high frequency (daily or intra-day) data and a plethora of features of financial data has been discovered and well-documented in the literature. These features, often called *universal features* or *stylized facts* in the literature, became source of inspiration for the design of many econometric models proposed in quantitative finance. Examples include among others the generalized autoregressive conditional heteroscedasticity (GARCH)-type models, the stochastic volatility (SV) family models. All these models find successful application in forecasting volatility and option pricing in empirical finance.

On the other side, the EMH has been questioned for a long time in the market microstructure literature. For instance, Grossman and Stiglitz (1980) claimed that the markets cannot be informationally efficient due to the fact that information is costly. This clearly indicates that the market is more complex than that we assumed up until now. So, it is clear that we need new models that can explain the information flow in the market, the price formation processes, the behavior of the market participants, their interactions and their decisions. These insights can help to better understand the financial market and better manage market risk. The theoretical microstructure models purport to explain the microstructure of financial markets. They are the starting point for the development of empirical models. The most prominent is the ACD model that had been developed in the literature in order to explain how information flows come in the markets, to test and confirm microstructure assertions empirically.

Although they found successful applications in empirical finance, it is well known that the ACD models cannot adequately reproduce the higher persistence of financial trade duration data. Our objective in writing this thesis is to bring to light the ability and the robustness of a new family of models, namely the multifractal models (MF), to reproduce financial data. Multifractal processes possess generic properties that are well-documented in the literature. These properties allow MF models with a few and coherent parameters to properly describe voluminous financial data. Studies by Calvet and Fisher (2001a, 2004a), Lux (2008), among others have already demonstrated and confirmed their superiority and robustness over GARCH and FIGARCH models in forecasting volatility.

The thesis extends the scope of the multifractal models to financial intertrade durations, bid-ask spreads and oil price volatility. We show that multifractal processes are convenient tools for modeling and forecasting oil price volatility, financial durations, and other trade-related variables. So, they can help to better understand intra-day price formation processes and forecast market risk.

1.2. Structure of the Thesis

This thesis is divided into four main parts. Part I comprises chapters 2 and 3, and is concerned with the review of the multifractal models and the autoregressive conditional duration models. It also serves as a general introduction and offers the readers a competent and an intimate knowledge of both models.

In chapter 2 we briefly introduce the main stylized facts of financial data, outline the first and the second generation of the multifractal models, their origin, properties and applications to finance. This chapter also presents different tools developed in the literature for the estimation of the first and second generation MF models and statistical inferences. Chapter 3 provides an overview of financial duration models, namely the autoregressive conditional duration (ACD) models and the Markov switching multifractal duration (MSMD) models introduced in the literature over the last twenty years. It presents the properties of both models, their estimation approach and diagnostic tests.

Part II is made of chapter 4. It evaluates and compares the forecasting performance of the Markov switching multifractal (MSM) and eight linear and nonlinear GARCH-type models: The generalized autoregressive conditional heteroscedasticity (GARCH), the integrated GARCH (IGARCH), the asymmetric GARCH (GJR-GARCH), the exponential GARCH (EGARCH), the asymmetric power ARCH (APARCH), the hyperbolic GARCH (HYGARCH) and the fractionally integrated APARCH (FIAPARCH) via six different loss functions and the superior predictive ability (SPA) test of Hansen (2005).

Part III includes chapters 5 through to 8. It covers the assessment of the predictive ability of the Markov switching multifractal duration model, its extension to generalized univariate and bivariate models, and its application to forecast irregularly spaced intraday value-at-risk (ISIVaR).

Chapter 5 compares the forecast performance of the MSMD model to those of the ACD, the Log-ACD and FIACD models with different distributions for the innovations via density forecasts and the likelihood ratio test. In chapter 6 we propose a generalized version of the Markov switching multifractal model in which the innovation is assumed to follow a generalized gamma or Burr distribution and has Chen et al.'s model as a special case. Chapter 7 introduces a bivariate Markov switching multifractal duration model that has been applied to analyze the covariation in the bid-ask spreads of different stocks traded on New York Stock Exchange (NYSE). In chapter 8 we apply the Chen et al.'s model to forecast irregularly spaced value-at-risk (ISIVaR) in a semi-parametric framework.

Finally, Part IV or chapter 9 presents the main findings and some outlooks for future research.

Part I.

Review Of Multifractal And Autoregressive Conditional Duration Models

2. Multifractal Models In Finance: Their Origin, Properties, and Applications

[1]

2.1. Introduction

One of the most important tasks in financial economics is the modeling and forecasting of price fluctuations of risky assets. For analysts and policy makers volatility is a key variable for understanding market fluctuations. Analysts need accurate forecasts of volatility as an indispensable input for tasks such as risk management, portfolio allocation, value-at-risk assessment, and option and futures pricing. Asset market volatility also plays an important role in monetary policy. Repercussions from the recent financial crisis on the global economy show how important it is to take into account financial market volatility in conducting effective monetary policy.

In financial markets, volatility is a measure for fluctuations of the price p of a financial instrument over time. It cannot be directly observed, but has to be estimated via appropriate measures or as a component of a stochastic asset pricing model. As an ingredient of such a model, volatility may be a latent stochastic variable itself (as it is in so-called stochastic volatility models as well as in most multifractal models) or it might be a deterministic variable at any time t (as it is the case in so-called GARCH type models). For empirical data, volatility may simply be calculated as the sample variance or sample standard deviation. Ding et al. (1993) propose using absolute returns for estimating volatility. Davidian and Carroll (1987) demonstrate that this measure is more robust against asymmetry and non-normality than others (cf. also Taylor, 1986; Ederington and Guan, 2005). Another way to measure daily volatility is to use squared returns or any other absolute power of returns. Indeed, different powers show slightly different time-series characteristics, and the multifractal model is designed to capture the complete range of behavior of absolute moments.

Recently, the concept of realized volatility (RV) has been developed by Andersen et al. (2001) as an alternative measure of the variability of asset prices (cf. also Barndorff-Nielsen and Shephard, 2002). The notion of RV means that daily volatility is estimated by summing up intra-day squared returns. This approach is based on the theory of quadratic variation which suggests that RV should provide a consistent and highly efficient non-parametric estimator of asset return volatility over a

[1] Prepared as a chapter for: Shu-Heng Chen and Mak Kaboudan. Forthcoming. OUP Handbook on Computational Economics and Finance. Oxford University Press.

given discrete interval under relatively parsimonious assumptions on the underlying data generating process. Other methods used for measuring volatility are: the maximum likelihood method developed by Ball and Torous (1984), or the high-low method proposed by Parkinson (1980). All these measures of financial market volatility show salient features which are well documented as *stylized facts*: Volatility clustering, asymmetry and mean reversion, comovements of volatilities across assets and financial markets, stronger correlation of volatility compared to that of raw returns, (semi-) heavy-tails of the distribution of returns, anomalous scaling behavior, changes in shape of the return distribution over time horizons, leverage effects, asymmetric lead-lag correlation of volatilities, strong seasonality, and some dependence of scaling exponents on market structure, cf. 2.2.

During the last decades, an immense body of theoretical and empirical studies has been devoted to formulate appropriate volatility models (cf. Andersen et al. (2006) for a recent review on volatility modeling and Poon and Granger (2003) for a review on volatility forecasting). With Mandelbrot's famous work on the fluctuations of cotton prices in the early sixties (cf. Mandelbrot, 1963), economists had already learned that the standard Geometric Brownian motion proposed by Bachelier (1900) is unable to reproduce these stylized facts. In particular, the fat tails and the strong correlation observed in volatility are in sharp contrast to the "mild", uncorrelated fluctuations implied by models with Brownian random terms. A first step toward covering time-variation of volatility had been taken with models using mixtures of distributions as proposed by Clark (1973) and Kon (1984). Econometric modeling of asset price dynamics with time-varying volatility got started with the generalized autoregressive conditional heteroscedasticity (GARCH) family and it numerous extensions (cf. Engle, 1982). The closely related class of stochastic volatility (SV) models adds randomness to the dynamic law governing the time variation of second moments (cf. Ghysels et al., 1996; Shephard, 1996, for a review on SV models and their applications).

In this chapter, the focus is on a new, alternative avenue for modeling and forecasting volatility developed in the literature over the last fifteen years or so. In contrast to the existing models the source of heterogeneity of volatility in these new models stems from the time-variation of local regularity in the price path (cf. Fisher et al., 1997). The background of these models is the theory of multifractal measures that has originally been developed by Mandelbrot (1974) in order to model turbulent flows. These multifractal processes have initiated a broad current of literature in statistical physics refining and expanding the underlying concepts and models (cf. Kahane and Peyrière, 1976; Holley and Waymire, 1992; Falconer, 1994; Arbeiter and Patzschke, 1996; Barral, 1999). The formal analysis of such measures and processes, the so-called multifractal formalism, has been developed by Frisch and Parisi (1985), Mandelbrot (1989, 1990), and Evertsz and Mandelbrot (1992), among others.

A number of early contributions have indeed pointed out certain similarities of volatility to fluid turbulence (cf. Vassilicos et al., 1994; Ghasghaie et al., 1996; Galluccio et al., 1997; Schmitt et al., 1999), while theoretical modeling in finance using the concept of multifractality started with the adaptation to an asset-pricing framework of Mandelbrot's (1974) model by Mandelbrot et al.

(1997).

Subsequent literature has moved from the more combinatorial style of the Multifractal Model of Assets Returns (MMAR) of Mandelbrot, Fisher, and Calvet (developed in the sequence of Cowles Foundation working papers authored by Calvet et al. (1997), Fisher et al. (1997), and Mandelbrot et al. (1997)) to iterative, causal models of similar design principles: The Markov-Switching Multifractal (MSM) model proposed by Calvet and Fisher (2004a) and the Multifractal Random Walk (MRW) by Bacry et al. (2001) constitute the second-generation of multifractal models that have more or less replaced the somewhat cumbersome (see below) first generation MMAR in empirical applications.

The rest of the chapter is organized as follows. Section 2.2 presents an overview over the salient stylized facts of financial data and discusses the potential of the classes of GARCH and stochastic volatility models to capture these stylized facts. In Section 2.3, we introduce the baseline concept of multifractal measures and processes and provide an overview over different specifications of multifractal volatility models. Shortcomings of the multifractal models are presented in Section 2.4. Section 2.5 introduces the different approaches to estimate MF models and to forecast future volatility. Section 2.6 reviews empirical results on the application and performance of MF models and Section 2.7 concludes.

2.2. Stylized Facts of Financial Data

With the availability of high-frequency time series for many financial markets from about the sixties, their statistical properties became a topic explored in a large strand of literature to which economists, statisticians and physicists have contributed. The two main universal features or "stylized facts" characterizing practically every series of interest at the high-end of the frequency spectrum (daily or intra-daily) are known under the catchwords "fat tails" and "volatility clustering". The use of multifractal models is motivated to some extent by both of these properties, but multifractality (or, as it is sometime also called, multi-scaling or multi-affinity) proper is a more subtle feature that gradually started to emerge as an additional stylized fact since the nineties. In the following we will provide a short review of the historical development of our knowledge and the quantification of all these features capturing in passing also some lesser known statistical properties typically found in financial returns. The data format of interest is thereby typically returns, i.e. relative price changes, $\tilde{r}_t = \frac{p_t - p_{t-1}}{p_{t-1}}$ which for high-frequency data are almost identical to log-price changes $r_t = \ln(p_t) - \ln(p_{t-1})$ with p_t the price at time t (e.g., at daily or higher frequency).

2.2.1. Fat Tails

This property relates to the shape of the unconditional distribution of a time series of returns. Historically, the first "hypothesis" on the distribution of price changes has been formulated by Bachelier (1900) who in his PhD thesis titled "Théorie de la Spéculation" assumed them to follow a Normal distribution. As is well known, many applied areas of financial economics such as option pricing theory (Black and Scholes, 1973) and portfolio theory (Markowitz, 1959) have followed this assumption, at least in their initial stages. The justification for this assumption is provided by the law of large numbers: If price changes at the smallest unit of time are independently and identically distributed random numbers (maybe driven by the stochastic flow of new information) returns over longer intervals can be seen as the sum of a large number of such *i.i.d.* observations, and irrespective of the distribution of their summands should under some weak additional assumptions converge to the Normal distribution. While this seemed plausible and the resulting Gaussian distribution would also come very handy for many applied purposes, Mandelbrot (1963) was the first to demonstrate that empirical data are distinctly non-Gaussian exhibiting excess kurtosis and higher probability mass in the center and in their tails than the Normal distribution. As can be confirmed with any sufficiently long record of stock market, foreign exchange or other financial data, the Gaussian distribution can always be rejected with statistical significance beyond all usual boundaries, and the observed largest historical price changes would be so unlikely under the Normal law that one would have to wait for horizons beyond at least the history of stock markets to observe them occur with non-negligible probability.

Mandelbrot (1963) and Fama (1963), as a consequence, proposed the so-called Lévy stable laws as an alternative for capturing these fat tails. This was motivated by the fact that in a generalized version of the central limit law dispensing with the assumption of a finite second moment, sums of *i.i.d.* random variables converge to these more general distributions (with the Normal being a special case of the Lévy stable obtained in the borderline case of a finite second moment). The desirable stability property, therefore, indicates the choice of the Lévy stable which also has a shape that -in the standard case of infinite variance- is characterized by fat tails. In a sense, the Lévy stable model remained undisputed for about three decades (although many areas of financial economics would rather continue to use the Normal as their working model), and economists indeed contributed to the advancement of statistical techniques for estimating the parameters of the Lévy distributions (Fama and Roll, 1971; McCulloch, 1986). When physicists started to explore financial time series, the Lévy stable law was discovered again (Mantegna, 1991) although new developments in empirical finance had already allowed to reject this meanwhile time-honored hypothesis.

These new insights were basically due to a different perspective: Rather than attempting to model the entire distribution, one let "speak the tails for themselves". The mathematical foundations for such an approach are provided by statistical extreme value theory (e.g., Reiss and Thomas, 1997). Its basic tenet is that the extremes and the tail regions of a sample of *i.i.d.* random variables converge in distribution to one of only three types of limiting laws. For tails, these are: Exponential decay, power-law decay and the behavior of distributions with finite endpoint of their support. Fat tails are often used as a synonym for power-law tails, so that the highest realizations of returns would obey a law like $Pr(x_t < x) \sim 1 - x^{-\alpha}$ after appropriate normalization (i.e. after some transformation $x_t = ar_t + b$). The universe of fat-tailed distributions can, then, be indexed by their tail index α with $\alpha \in (0, \infty)$. Lévy stable distributions are characterized by tail indices

α below 2 (2 characterizing the case of the Normal distribution). All other distributions with a tail index smaller than 2 would converge under summation to the Lévy stable with the same index while all distributions with an asymptotic tail behavior with $\alpha > 2$ would converge under aggregation to the Gaussian. This demarcates the range of relevance of the standard central limit law and its generalized version.

Jansen and de Vries (1991), Koedijk et al. (1990) and Lux (1996) are examples of a literature that emerged over the nineties using semi-parametric methods of inference to estimate the tail index without assuming a particular shape of the entire distribution. The outcome of these and other studies is a tail index α in the range of 3 to 4 that now counts as a stylized fact (cf. Guillaume et al., 1997; Gopikrishnan et al., 1998). Intra-daily data nicely confirm results obtained for daily records in that they provide estimates for the tail index that are in line with the former (cf. Dacorogna et al., 2001; Lux, 2001a), and, therefore, confirm the expected stability of the tail behavior under time aggregation as predicted by extreme-value theory. The Lévy stable hypothesis, thus, can be rejected (confidence intervals of α typically exclude the possibility of $\alpha < 2$). This agrees with the evidence that the variance stabilizes with increasing sample size and does not explode. Falling into the domain of attraction of the Normal distributions, the overall shape of the return distribution would have to change, i.e. get closer to the Normal under time aggregation.[2] This is indeed the case, as has been demonstrated by Teichmoeller (1971) and many later authors. Hence, the basic finding on the unconditional distribution is that it converges toward the Gaussian, but is distinctly different from it at the daily (and higher) frequencies. *Fig.* 2.1 illustrates the very homogeneous and distinctly both non-Gaussian and non-Levy nature of stock price fluctuations. The four major South-African stocks displayed in the figure could be replaced by almost any other time series of stock markets, foreign exchange markets and a variety of other financial markets. Estimating the tail index α by a linear regression in this log-log plot would lead to numbers very close to the celebrated "cubic law".

The particular non-Normal shape then also motivates the quest for the best non-stable characterization at intermediate levels of aggregation. From a huge literature that has tried mixtures of Normals (Kon, 1984) as well as a broad range of generalized distributions (cf. Eberlein and Keller, 1995; Behr and Pötter, 2009; Fergussen and Platen, 2006) it appears that the distribution of daily returns is quite close to a Student$-t$ with three degrees of freedom. However, while a tail index between 3 and 4 is typically found for stock and foreign exchange markets, some other markets are sometimes found to have fatter tails, e.g., Koedijk et al. (1992) for black market exchange rates, and Matia et al. (2002) for commodities.

Figure 2.1 about here

[2] While, in fact, the tail behavior would remain qualitatively the same under time aggregation, the asymptotic power law would apply in a more and more remote tail region only, and would, therefore, become less and less visible for finite data samples under aggregation. There is, thus, both convergence towards the Normal distribution and stability of power-law behavior in the tail under aggregation. While the former governs the complete shape of the distribution, the latter applies further and further out in the tail only and would only be observed with a sufficiently large number of observations.

2.2.2. Volatility Clustering

The slow convergence to the Normal might be explained by dependency in the time series of returns. Indeed, while the limiting laws of extreme value theory would still apply for certain deviations from *i.i.d.* behavior, dependency could slow down convergence dramatically leading to a long regime of pre-asymptotic behavior. That returns are characterized by a particular type of dependency has also been well known for long time, and is mentioned, for instance, by Mandelbrot (1969). This dependency is most pronounced and in fact, plainly visible in absolute returns, squared returns, or any other measure of the extent of fluctuations (volatility), cf. *Fig.* 2.2. In all these measure there is long lasting, highly significant autocorrelation (cf. Ding et al., 1993). With sufficiently long time series, significant autocorrelation can be found for time lags (of daily data) up to a few years. This positive feedback is described as volatility clustering or "turbulent (tranquil) periods being more likely to be followed by still turbulent (tranquil) periods than vice versa". Whether there is (additional) dependency in the raw returns is subject to debate. Most studies do not find sufficient evidence for giving up the martingale hypothesis although a long-lasting but small effect might be hard to capture statistically. Ausloos et al. (1999) is an example of a study claiming to have identified such effects. Lo (1991) has proposed a rigorous statistical test for long term dependence that mostly does not indicate deviations from the null hypothesis of short memory for raw asset returns, but strongly significant evidence of long memory in squared or absolute returns. Similarly as for the classification of types of tail behavior, short memory comes along with exponential decay of the autocorrelation function while one speaks of long memory if the decay follows a power-law. Evidence for the later type of behavior has also accumulated over time. Documentation of hyperbolic decline in the autocorrelations of squared returns can be found in Dacorogna et al. (1993), Crato and de Lima (1994), Lux (1996) and Mills (1997). Lobato and Savin (1998) first claimed that such long-range memory in volatility measures is a universal stylized fact of financial markets while Lobato and Velasco (2000) document similar long-range dependence in trading volume. Again, particular market designs might lead to exceptions from the typical power-law behavior. Gençay (2001) as well as Ausloos and Ivanova (2000) report untypical behavior in the managed floating of European currencies during the times of the European Monetary System. Presumably due to leverage effects, stock markets also exhibit correlation between volatility and raw (i.e., signed) returns (cf. LeBaron, 1992), that is absent in foreign exchange dates.

Figure 2.2 about here

2.2.3. Benchmark Models: GARCH and Stochastic Volatility

In financial econometrics, volatility clustering has since the eighties spawned a voluminous literature on a new class of stochastic processes capturing the dependency of second moments in a phenomenological way. Engle (1982) first introduced the ARCH (autoregressive conditional heteroscedasticity model) which has been generalized to GARCH by Bollerslev (1986). It models returns as a mixture of Normals with the current variance being driven by a deterministic difference equation:

$$r_t = h_t \varepsilon_t \qquad \text{with} \quad \varepsilon_t \sim N(0,1) \tag{2.1}$$

and

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}, \;\; \alpha_0 > 0, \alpha_i, \beta_j > 0. \tag{2.2}$$

Empirical applications usually find a parsimonious GARCH(1,1) model (i.e., $p = q = 1$) sufficient, and when estimated, the sum of the parameters $\alpha_1 + \beta_1$ turns out to be close to the non-stationary case (or, expressed differently, mostly only a constraint on the parameters prevents them for exceeding 1 in their sum which would lead to non-stationary behavior). Different extensions of GARCH were developed in the literature with the objective to better capture the stylized facts. Among them there are: The Exponential GARCH (EGARCH) model proposed by Nelson (1991) that accounts for asymmetric behavior of returns, the Threshold GARCH (TGARCH) model of Rabemananjara and Zakoian (1993) which takes into account the leverage effects, the regime switching GARCH (RS-GARCH) developed by Cai (1994), and the Integrated GARCH (IGARCH) introduced by Engle and Bollerslev (1986a) that allows for capturing high persistence observed in returns time series. Itô diffusion or jump-diffusion processes can be obtained as a continuous time limit of discrete GARCH sequences (cf. Nelson, 1990; Drost and Werker, 1996).

To capture stochastic shocks to the variance process, Taylor (1986) introduced the class of stochastic volatility models whose instantaneous variance is driven by:

$$\ln(h_t) = k + \varphi \ln(h_{t-1}) + \tau \xi_t, \qquad \xi_t \sim N(0,1). \tag{2.3}$$

This approach as well has been refined and extended in many ways. The SV process is more flexible than the GARCH model and provides more mixing because of the co-existence of shocks to volatility and return innovations (cf. Gavrishchaka and Ganguli, 2003). In terms of statistical properties, one important drawback of at least the baseline formalizations (2.1) to (2.3) is their implied exponential decay of the autocorrelations of measures of volatility which is in contrast to the very long autocorrelations mentioned before. Both the elementary GARCH and the baseline SV model are characterized by only short-term rather than long-term dependence.

To capture long memory, GARCH and SV models have been expanded by allowing for an infinite number of lagged volatility terms instead of the limited number of lags appearing in (2.2) and (2.3). To obtain a compact characterization of the long memory feature a fractional differencing operator has been used in both extensions leading to the fractionally integrated GARCH (FIGARCH) model of Baillie et al. (1996) and the long-memory stochastic volatility model of Breidt et al. (1998).[3] An interesting intermediate approach is the so-called heterogenous ARCH (HARCH) model of Dacorogna et al. (1998) that considers returns at different time aggregation

[3] The "self-excited multifractal model" proposed by Filimonov and Sornette (2011) appears closer to this model rather than to models from the class of multifractal processes discussed below.

levels as determinants of the dynamic law governing current volatility. Under this model, *eq.* (2.2) would have to be replaced by

$$h_t = c_0 + \sum_{j=1}^{n} c_j r^2_{t,t-\Delta t_j}, \tag{2.4}$$

where $r_{t,t-\Delta t_j} = \ln(p_t) - \ln(p_{t-\Delta t_j})$ are returns computed over different frequencies. The development of this model was motivated by the finding that volatility on fine time scales can be explained to a larger extend by coarse-grained volatility than vice versa (cf. Müller et al., 1997). Hence, the right-hand side covers local volatility at various lower frequencies than the time step of the underlying data ($\Delta t_j = 2, 3, \ldots$). As we will see in the following, multifractal models have a closely related structure but model the hierarchy of volatility components in a multiplicative rather than additive format.

2.2.4. A New Stylized Fact: Multifractality

Both the hyperbolic decay of the unconditional pdf as well as the similarly hyperbolic decay of the autocorrelations of many measures of volatility (squared, absolute returns) would fall into the category of scaling laws in the natural sciences. The identification of such universal scaling laws in an area like finance has spawned the interest of natural scientists to further explore the behavior of financial data and to develop models to explain these characteristics (cf. Mantegna and Stanley, 1996). From this line of research, multifractality, multi-scaling or anomalous scaling emerged gradually over the nineties as a more subtle characteristic of financial data that motivated the adaptation of known generating mechanisms for multifractal processes from the natural sciences in empirical finance.

To define multifractality or multiscaling, we start with the more basic concepts of fractality or scaling. The defining property of fractality is the invariance of some characteristic under appropriate self-affine transformations. The power-law functions characterizing the pdf of returns and autocorrelations of volatility measures are scale-invariant properties, i.e., this behavior is preserved over different scales under appropriate transformations.[4] In a most general way, some property of an object or a process needs to fulfill a law like

$$x(ct) = c^H x(t) \tag{2.5}$$

in order to be classified as scale-invariant, where t is an appropriate measurement of a scale (e.g., time or distance). Strict validity of (2.5) holds for many of the objects that have been investigated in fractal geometry (Mandelbrot, 1982). In the framework of stochastic processes, such laws could only hold in distribution. In this case, Mandelbrot et al. (1997) speak of self-affine processes. An example of a well-known class of processes obeying such a scale invariance principle is fractional Brownian motion for which $x(t)$ is a series of realizations and $0 < H < 1$ is the Hurst index

[4] e.g., from the limiting power law the cdf of a process with hyperbolically decaying tails obeys $Pr(x_i < x) \approx x^{-\alpha}$ and obviously for any multiple of x the same law applies: $Pr(x_i < cx) \approx (cx)^{-\alpha} = c^{-\alpha} x^{-\alpha}$.

that determines the degree of persistence ($H > 0.5$) or anti-persistence ($H < 0.5$) of the process, $H = 0.5$ corresponding to Wiener Brownian motion with uncorrelated Gaussian increments. *Fig.* 2.2 shows the scaling behavior of different powers of returns (raw, absolute and squared returns) of a financial index as determined by a popular method for the estimation of the Hurst coefficient, H. The law (2.5) also determines the dependency structure of the increments of a process obeying such scaling behavior as well as their higher moments which show hyperbolic decline of their autocorrelations with an exponent depending linearly on H. Such linear dependence is called uni-scaling or uni-fractality. It also carries over asymptotically to processes that use a fractional process as generator for the variance dynamics, e.g. the long memory stochastic volatility model of Breidt et al. (1998).[5]

Multifractality or anomalous scaling allows for a richer variation of the behavior of a process across different scales by only imposing the more general relationship:

$$x(ct) \overset{d}{=} M(c)x(t) \equiv c^{H(c)}x(t), \tag{2.6}$$

where the scaling factor $M(c)$ is a random function with possibly different shape for different scales and $\overset{d}{=}$ denotes equality in distribution. The last equality of *eq.* (2.6) illustrates that this variability of scaling laws could be translated into variability of the index H which now is not constant anymore. One might also note the multiplicative nature of transitions between different scales: One moves from one scale to another via multiplication with a random factor $M(c)$. We will see below that multifractal measures or processes are constructed exactly in this way which implies a combinatorial, noncausal nature of these processes.

Multi-scaling in empirical data is typically identified by differences in the scaling behavior of different (absolute) moments:

$$\mathbb{E}\left[|x(t,\Delta t)|^q\right] = c(q)\Delta t^{qH(q)+1} = c(q)\Delta t^{\tau(q)+1}, \tag{2.7}$$

with $x(t,\Delta t) = x(t) - x(t-\Delta t)$, and $c(q)$ and $\tau(q)$ being deterministic functions of the order of the moment q. A similar equation could be established for uni-scaling processes, e.g. fractional Brownian motion, yielding

$$\mathbb{E}\left[|x(t,\Delta t)|^q\right] = c^H \Delta t^{qH+1}. \tag{2.8}$$

Hence, in terms of the behavior of moments, multifractality (anomalous scaling) is distinguished by a non-linear (typically concave) shape from the linear scaling of uni-fractal, self-affine processes. The standard tool to diagnose multifractality is, then, inspection of the empirical scaling behavior of an ensemble of moments. Such non-linear scaling is illustrated in *Fig.* 2.3 for three selected stock indices and a stochastic process with multifractal properties (the Markov-switching multifractal model introduced below). The traditional approach in the physics literature consists in extracting $\tau(q)$ from a chain of linear log-log fits of the behavior of various moments q for a

[5] For the somewhat degenerate FIGARCH model, the complete asymptotics have not yet been established, cf. Jach and Kokoszka (2010).

certain selection of time aggregation steps Δt. One, therefore, uses regressions to the temporal scaling of moments of powers q:

$$\ln \mathbb{E}\left[|x(t, \Delta t)|^q\right] = a_0 + a_1 \ln(\Delta t) \tag{2.9}$$

and constructs the empirical $\tau(q)$ curve (for a selection of discrete q) from the ensemble of estimated regression coefficients for all q. An alternative and perhaps even more widespread approach for identification of multifractality looks at the varying scaling coefficients $H(q)$ in *eq.* (2.7). While the unique coefficient H of *eq.* (2.8) is usually denoted the Hurst coefficient, the multiplicity of such coefficients in multifractal processes is denoted as Hölder exponents. While the unique H quantifies a global scaling property of the underlying process, the Hölder exponents can be viewed as local scaling rates that govern various patches of a time series leading to a characteristically heterogeneous (or intermittent) appearance of such series. An example is displayed in *Fig.* 2.5 (principles of construction being explained below). Focusing on the concept of Hölder exponents, multifractality then amounts to identification of the range of such exponents rather than a degenerate single H as for uni-fractal processes. The so-called spectrum of Hölder exponents (or multifractal spectrum) can be obtained by the Legendre transformation[6] of the scaling function $\tau(q)$. Define $\alpha = \frac{d\tau}{dq}$, the Legendre transform $f(\alpha)$ of the function $\tau(q)$ is given by

$$f(\alpha) = \arg\min_q [q\alpha - \tau(q)], \tag{2.10}$$

where α is the Hölder exponent (the established notation for the counterpart of the constant Hurst exponent, H) and $f(\alpha)$ the multifractal spectrum that describes the distribution of the Hölder exponents. The local Hölder exponent quantifies the local scaling properties (local divergence) of the process at a given point in time, in other words, it measures the local regularity of the price process. In traditional time series models, the distribution of Hölder exponents is degenerate converging to a single such exponent (unique Hurst exponent) while multifractal measures are characterized by a continuum of Hölder exponents whose distribution is given by the Legendre transform, *eq.* (2.10), for its particular scaling function $\tau(q)$. The characterization of a multifractal process or measure by a distribution of local Hölder exponents underlines its heterogeneous nature with alternating calm and turbulent phases.

Empirical studies allowing for such a heterogeneity of scaling relations typically identify "anomalous scaling" (curvature of the empirical scaling functions or non-singularity of the Hölder spectrum) for financial data as illustrated in *Fig.* 2.3. Historically, the first example of such an analysis is Müller et al. (1990) followed by more and more similar findings reported mostly in the emerging econophysics literature (due to the fact that the underlying concepts were well-known in physics from research on turbulent flows, but were completely alien to financial economists). Examples include Vassilicos et al. (1994), Mantegna and Stanley (1995), Ghasghaie et al. (1996), Fisher

[6] The Legendre transformation is a mathematical operation that transforms a function of a coordinate, $g(x)$, into a new function $h(y)$ whose argument is the derivative of $g(x)$ with respect to x, i.e., $y = \frac{dg}{dx}$.

et al. (1997), Schmitt et al. (1999), Fillol (2003), among others. Ureche-Rangau and de Morthays (2009) show that both volatility and volume of Chinese stocks appear to have multifractal properties, a finding one should probably be able to confirm for other markets as well given the established long-term dependence and high cross-correlation between both measures (cf. Lobato and Velasco, 2000), who among others, also report long-term dependence of volume data). While econometricians have not been looking at scaling functions and Hölder spectrums, the indication of multifractality in the mentioned studies has nevertheless some counterpart in the economics literature: The well-known finding of Ding et al. (1993) that (i) different powers of returns have different degrees of long-term dependence and that (ii) the intensity of long-term dependence varies non-monotonically with q (with a maximum obtained around $q \approx 1$) is consistent with concavity of scaling functions and provides evidence for "anomalous" behavior form a slightly different perspective.

Multifractality, thus, provides a generalization of the well established finding of long-term dependence of volatility: Different measures of volatility are characterized by different degrees of long-term dependence in a way that reflects the typical anomalous behavior of multifractal processes. Accepting such behavior as a new stylized fact, the natural next step would be to design processes that could capture this universal finding together with other well-established stylized facts of financial data. New models would be required because none of the existing ones would be consistent with this type of behavior: Baseline GARCH and SV models have only exponential decay of the autocorrelations of absolute powers of returns (short-range dependence), while their long memory counterparts (LMSV, FIGARCH) are characterized by uni-fractal scaling.[7]

One caveat is, however, in order here: Whether the scaling function and Hölder spectrum analysis provide sufficient evidence for multifractal behavior, is to some extent subject to dispute. A number of papers show that scaling in higher moments can be easily obtained in a spurious way without any underlying anomalous diffusion behavior. Lux (2004) pointed out that a non-linear shape of the empirical $\tau(q)$ function is still obtained for financial data after randomization of their temporal structure, so that the $\tau(q)$ and $f(\alpha)$ estimators are rather unreliable diagnostic instruments for the presence of multifractal structure in volatility. Apparent scaling has also been illustrated by Barndorff-Nielsen and Prause (2001) as a consequence of fat tails in the absence of true scaling. It is very likely that standard volatility models would also lead to apparent multi-scaling that could be hard to distinguish from "true" multifractality via the diagnostic tools mentioned above.[8] Formally, it will always be possible to design processes without a certain type of (multi-) scaling behavior that are locally so close to "true" (multi-)scaling that these deviations will never be detected with pertinent diagnostic tools and restricted availability of data (cf. LeBaron, 2001; Lux, 2001b).

On the other hand, one might follow Mandelbrot's frequently voiced methodological premise

[7] For FIGARCH this is so far only indicated by simulations, but given that- as for LMSV- FIGARCH consists of a uni-fractal ARFIMA process plugged into the variance equation, it seems plausible that it also has uni-fractal asymptotics.

[8] There is also a sizeable literature on spurious generation of fat tails and long-term dependence, cf. Granger and Teräsvirta (1999) or Kearns and Pagan (1997).

to model apparently generic features of data by similarly generic models rather than using "fixes" (Mandelbrot, 1997a). Introducing amendments to existing models (e.g., GARCH, SV) to adapt those to new stylized facts might lead to highly parameterized setups that lack robustness when applied to data from different markets, while simple generating mechanisms for multifractal behavior are available that could, in principle, capture the whole spectrum of time series properties highlighted above in a more parsimonious way. In addition, if one wants to account for multi-scaling proper (rather than as a spurious property) no avenue is known so far for equipping GARCH- or SV-type models with this property in a generic way. Hence, adapting in an appropriate way some known generating mechanism for multifractal behavior appears the only avenue available so far to come up with models that generically possess such features, and jointly reproduce all stylized facts of asset returns. The next section recollects the major steps in the development of multifractal models for asset-pricing applications.

Figure 2.3 about here

2.3. Multifractal Measures and Processes

In the following, we first explain the construction of a simple multifractal measure and show how one can generalize it along various dimensions. We, then, move on to multifractal processes designed as models for financial returns.

2.3.1. Multifractal Measures

Multifractal measures have a long history in physics dating back to the early seventies when Mandelbrot proposed a probabilistic approach for the distribution of energy in turbulent dissipation (e.g., Mandelbrot, 1974). Building upon earlier models of energy dissipation by Kolmogorov (1941, 1962) and Obukhov (1962), Mandelbrot proposed that energy should dissipate in a cascading process on a multifractal set from long to short scales. In this original setting, the multifractal set results from operations performed on probability measures. The construction of a multifractal "cascade" starts by assigning uniform probability to a bounded interval (e.g., the unit interval $[0, 1]$). In a first step, this interval is split up into two subintervals receiving fractions m_0 and $1 - m_0$, respectively, of the total probability mass of unity of their mother interval. In the simplest case, both subintervals have the same length (i.e., 0.5), but other choices are possible as well. In the next step, the two subintervals of the first stage of the cascade are split up again into similar subintervals (of length 0.25 each in the simplest case) receiving again fractions m_0 and $1 - m_0$ of the probability mass of their "mother" intervals (cf. *Fig.* 2.4). In principle, this procedure is repeated *ad infinitum*. With this recipe, a heterogeneous, fractal distribution of the overall probability mass results which even for the most elementary cases has a perplexing visual resemblance to time series of volatility in financial markets. This construction clearly reflects the underlying idea of dissipation of energy from the long scales (the mother intervals) to the finer scales that preserve the joint influence of all the previous hierarchical levels in the built-up of the "cascade".

Many variations of the above generating mechanism of a simple Binomial multifractal could be thought of: Instead of always assigning probability m_0 to the left-hand descendent, this assignment could as well be randomized. Furthermore, one could think of more than two subintervals to be generated in each step (leading to multinomial cascades) or of using random numbers for m_0 instead of the same constant value. A popular example of the later generalization is the Lognormal multifractal model which draws the mass assigned to new branches of the cascade from a Lognormal distribution (cf. Mandelbrot, 1974, 1990). Note that for the Binomial cascade the overall mass over the unit interval is exactly conserved at any preasymptotic stage as well as in the limit $k \to \infty$, while mass is preserved only in expectation under appropriately normalized Lognormal multipliers, or multipliers following any other continuous function. Another straightforward generalization consists in splitting each interval on level j into an integer number b of pieces of equal length at level $j + 1$. The grid-free Poisson multifractal measure developed by Calvet and Fisher (2001a) is obtained by allowing for randomness in the construction of intervals. In this setting, a bounded interval is split into separate pieces with different mass by determining a random sequence T_n of change points. Overall mass is then distributed via random multipliers across the elements of the partition defined by the T_n. A multifractal sequence of measures is generated by a geometric increase of the frequency of arrivals of change points at different levels j ($j = 1, \dots, k$) of the cascade. As in the grid-based multifractal measures, the mass within any interval after the completion of the cascade is given by the product of all k random multipliers within that segment.

Note that all the above recipes can be interpreted as implementations (or examples) of the general form (2.6) that defines multifractality from the scaling behavior across scales. The recursive construction principles are, themselves, directly responsible for the multifractal properties of the pertinent limiting measures. The resulting measures, thus, obey multifractal scaling analogous to *eq.* (2.7). Denoting by μ a measure defined on $[0, 1]$, this amounts to[9] $\mathbb{E}[\mu(t, t + \Delta t)^q] \sim c(q)(\Delta t)^{\tau(q)+1}$. Exact proofs for the convergence properties of such grid bound cascades have been provided by Kahane and Peyrière (1976). The "multifractal formalism" that had been developed after Mandelbrot's pioneering contribution consisted in the generalization and analytical penetration of various multifractal measures following the above principles of construction (cf. Tél, 1988; Evertsz and Mandelbrot, 1992; Riedi, 2002). Typical questions of interest are the determination of the scaling function $\tau(\alpha)$ and the Hölder spectrum $f(\alpha)$, as well as the existence of moments in the limit of a cascade with infinite progression.

Figure 2.4 about here

2.3.2. Multifractal Models

2.3.2.1. Univariate Continuous-Time Multifractal Models

[9] For example, for the simplest case of the Binomial cascade one gets $\tau(q) = -\ln \mathbb{E}[M^q] - 1$ with $M \in \{m_0, 1 - m_0\}$ with probability 0.5.

2.3.2.1.1. The Multifractal Model of Asset Returns

Multifractal measures have been adapted to asset-price modeling by using them as a "stochastic clock" for transformation of chronological time into business (or intrinsic) time. Formally, such a time transformation can be represented by stochastic subordination, with the time change represented by a stochastic process, say $\theta(t)$ denoted the "subordinating process", and the asset price change, $r(t)$, being given by a subordinated process (e.g. Brownian motion) measured in transformed time, $\theta(t)$. In this way, the homogenous subordinated process might be modulated in a way to give rise to realistic time series characteristics such as volatility clustering. The idea of stochastic subordination has been introduced in financial economics by Mandelbrot and Taylor (1967). A well-known later application of this principle is Clark (1973) who had used trading volume as a subordinator (cf. Ané and Geman (2000), for recent extensions of this approach).

Mandelbrot et al. (1997) seems to be the first paper that went beyond establishing phenomenological proximity of financial data to multifractal scaling. They proposed a model, termed the Multifractal Model of Asset Returns (MMAR), in which a multifractal measure as introduced in sec. 2.3.1 serves as a time transformation from chronological time to business time. While the original paper has not been published in a journal, a synopsis of this entry and two companion papers (Calvet et al., 1997; Fisher et al., 1997) has appeared as Calvet and Fisher (2002). Several other contributions by Mandelbrot (1997b, 1999, 2001a,b,c) contain graphical discussions of the construction of the time-transformed returns of the MMAR process and simulations of examples of the MMAR as a data generating process. Formally, the MMAR assumes that returns $r(t)$ follow a compound process:

$$r(t) = B_H[\theta(t)], \tag{2.11}$$

in which an incremental fractional Brownian motion with Hurst index H, $B_H[\cdot]$, is subordinate to the cumulative distribution function $\theta(t)$ of a multifractal measure constructed along the above lines. When investigating the properties of this process, the (unifractal) scaling of the fractional Brownian motion has to be distinguished from the scaling behavior of the multifractal measure. The behavior of the compound process is determined by both, but its multi-scaling in absolute moments remains in place even for $H = 0.5$, i.e. Wiener Brownian motion. Under the restriction $H = 0.5$, the Brownian motion part becomes uncorrelated Wiener Brownian motion and the MMAR shows the martingale property of most standard asset pricing models. This model shares essential regularities observed in financial time series including long tails and long memory in volatility which both originate from the multifractal measure $\theta(t)$ applied for the transition from chronological time to "business time". The heterogenous sequence of the multifractal measure, then, serves to contract or expand time and, therefore, also contracts or expands locally the homogeneous second moment of the subordinate Brownian motion.

As pointed out above, different powers of such a measure have different decay rates of their autocovariances. Mandelbrot et al. (1997) demonstrate that the scaling behavior of the multifractal time transformation carries over to returns from the compound process (2.11) which would obey a scaling function $\tau_r(q) = \tau_\theta(qH)$. Similarly, the shape of the spectrum carries over from the time transformation to returns in the compound process via a simple relationship: $f_r(\alpha) = f_\theta(\alpha/H)$. By writing $\theta(t) = \int_0^t d\theta(t)$, it becomes clear that the incremental multifractal random measure $d\theta(t)$ (which is the limit of $\mu[t, t + \Delta t]$ for $\Delta t \to 0$ and k (the number of hierarchical levels) $\to \infty$) can been considered as the instantaneous stochastic volatility. As a result, MMAR essentially applies the multifractal measure to capture the time-dependency and non-homogeneity of volatility. Mandelbrot et al. (1997) and Calvet and Fisher (2002) discuss estimation of the underlying parameters of the MMAR model via matching of the $f(\alpha)$ and $\tau(\alpha)$ functions, and show that the temporal behavior of various absolute moments of typical financial data squares well with the theoretical results for the multifractal model.

Any possible implementation of the underlying multifractal measure could be used for the time-transformation $\theta(t)$. All examples considered in their papers built upon a binary cascade in which the time interval of interest (in place of the unit interval in the abstract operations on a measure described in sec. 2.3.1) is split repeatedly into subintervals of equal length. The so obtained subintervals are assigned fractions of the probability mass of their mother interval drawn from different types of random distributions: Binomial, Lognormal, Poisson and Gamma distributions are discussed in Calvet and Fisher (2002) each of those leading to a particular $\tau(\alpha)$ and $f(\alpha)$ function (known from previous literature) and similar behavior of the compound process according to the relations detailed above. Lux (2001c) applies an alternative estimation procedure minimizing a Chi-square criterion for the fit of the implied unconditional distribution of the MMAR to the empirical one, and reports that one can obtain surprisingly good approximations to the empirical shape in this way. However, Lux (2004) documents that $\tau(\alpha)$ and $f(\alpha)$ functions are not very reliable as criteria for determination of the parameters of the MMAR as even after randomization of the underlying data, one still gets indication of temporal scaling structure via non-linear $\tau(\alpha)$ and $f(\alpha)$ shapes. Poor performance of such estimators is also expected on the ground of the slow convergence of their variance as demonstrated by Ossiander and Waymire (2000). One might also point out in this respect, that both functions are capturing various moments of the data, so using them for determination of parameters amounts to some sort of moment matching. It is, however, not obvious that the choice of weight of different moments implied by these functions would be statistically efficient.

While MMAR has not been pursued further in subsequent literature, estimation of alternative multifractal models has made use of efficient moment estimators as well as other more standard statistical techniques. The main drawback of the MMAR is, that despite the attractiveness of its stochastic properties, its practical applicability suffers from the combinatorial nature of the subordinator $\theta(t)$ and its non-stationarity due to the restriction of this measure to a bounded interval. These limitations have been overcome by the analogous iterative time series models introduced by Calvet and Fisher (2001a, 2004a). Leövey and Lux (2012) have also recently proposed a re-interpretation of the MMAR in which an infinite succession of multifractal cascades overcomes the limitation to a bounded interval, and the resulting overall process could be viewed as a stationary one.

It is interesting to relate the grid-bound construction of the MMAR to the "classical" formalization of stochastic processes for turbulence. Building upon previous work by Kolmogorov (1962) and Obukhov (1962) on the phenomenology of turbulence, Castaing et al. (1990) has introduced the following approach to replicate the scaling characteristics of turbulent flows:

$$x_i = \exp(\varepsilon_i)\xi_i, \tag{2.12}$$

with ξ_i and ε_i both following a Normal distribution $\xi_i \sim N(0, \sigma^2)$ and $\varepsilon_i \sim N(\ln(\sigma_0), \lambda^2)$, and ξ_i and ε_i mutually independent. This approach has been applied to various fluctuating phenomena in the natural sciences such as hadron collision (cf. Carius and Ingelman, 1990), solar wind (cf. Sorriso-Valvo et al., 1999), and human heartbeat (cf. Kiyono et al., 2004, 2005). Replacing the uniform ε_i by the sum of hierarchically organized components, the resulting structure would closely resemble that of the MMAR model. Models in this vein have been investigated in physics by Kiyono et al. (2007) and Kiyono (2009). Based on the approach exemplified in *eq.* (2.12), Ghasghaie et al. (1996) elaborate on the similarities between turbulence in physics and financial fluctuations, but do not take into account the possibility of multifractality of the data generating process.

2.3.2.1.2. The MMAR with Poisson Multifractal Time Transformation

Already in Calvet and Fisher (2001a), a new type of multifractal model has been introduced that overcomes some of the limitations of the MMAR as proposed by Mandelbrot et al. (1997) while -initially- preserving the formal structure of a subordinated process. Instead of the grid-based binary splitting of the underlying interval (or, more generally, the splitting of each mother interval into the same number of subintervals), they assume that $\theta(t)$ is obtained in a grid-free way by determining a Poisson sequence of change points for the multipliers at each hierarchical level of the cascade. Multipliers themselves might again be drawn from a Binomial, Lognormal (the standard cases), or any other distribution with positive support. Change points are determined by renewal times with exponential densities. At each change point t_n^i a new draw $M_{t_n}^i$ of cascade level i occurs from the distribution of the multipliers that is standardized in a way to ensure conservation of overall mass $\mathbb{E}[M_{t_n}^i] = 1$. In order to achieve the hierarchical nature of the cascade, the different levels i are characterized by a geometric progression of the frequencies of arrival $b^i\lambda$. Hence, the change points t_n^i follow level-specific densities $f(t_n^{(i)}; \lambda, b) = b^i\lambda \exp(-b^i\lambda t_n^i)$, for $i = 1, ..., k$. Similar grid-free constructions for multifractal measures are considered in Cioczek-Georges and Mandelbrot (1995) and Barral and Mandelbrot (2002). In the limit $k \to \infty$ the Poisson multifractal exhibits typical anomalous scaling, which again carries over from the time transformation $\theta(t)$ to the subordinate process for asset returns, $B_H[\theta(t)]$ in the way demonstrated by Mandelbrot et al. (1997).

The importance of this variation of the original grid-bound MMAR is that it provides an avenue towards constructing multifractal models (or models arbitrarily close to "true" multifractals) in a way that allows better statistical tractability. In particular, in contrast to the grid-bound MMAR, the Poisson multifractal possesses a Markov structure. Since the $t_n^{(i)}$ follow an exponential distribution, the probability of arrivals at any instant t is independent from past history. As an immediate consequence, the initial restriction upon its construction to a bounded interval in time $[0, T]$ is not really necessary, as the process can be continued when reaching the border $t = T$ in the very same way by which realizations have been generated within the interval $[0, T]$ without any disruption of its stochastic structure. This is not the case for the grid-based approach where one could, in principle, append a new cascade after $t = T$ which, however, would be completely uncorrelated with the previous one. The continuous-time Poisson multifractal has not been used itself in empirical applications, but it has motivated the development of the discrete Markov-switching multifractal model (MSM) that has become the most frequently applied version of multifractal models in empirical finance, cf. sec. 2.3.3.

2.3.2.1.3. Further Generalizations of Continuous-Time MMAR

In a foreword to the working paper version (2001) of their paper, Barral and Mandelbrot (2002) motivate the introduction of what they call "multifractal products of cylindrical pulses" by its greater flexibility compared to standard multifractals. They argue that this generalization should be useful in order to capture particularly the power-law behavior of financial returns. Again, in the construction of the cylindrical pulses the renewal times at different hierarchical levels are determined by Poisson processes whose intensities are not, however, connected via the geometric progression $b^i \lambda$ (reminiscent of the grid size distribution in the original MMAR), but are scattered randomly according to Poisson processes with frequencies of arrival depending inversely on the scale s, i.e. assuming $r_i = s_i^{-1}$ (instead of $r_i = 2^{i-k}$ at scales $s_i = 2^{k-i}$ over an interval $[0, 2^k]$ in the basic grid-bound approach for multifractal measures). Associating independent weights to the different scales one obtains a multifractal measure for this construction by taking a product of these weights over a conical[10] domain in (t, s) space. The theory of such cylindrical pulses (i.e., the pertinent multipliers $M^i_{t_n}$ that rule one hierarchical level between adjacent change points t_n and t_{n+1}) only needs the requirement of existence of $\mathbb{E}[M^i_{t_n}]$. Barral and Mandelbrot (2002) work out the "multifractal apparatus" for such more general families of hierarchical cascades pointing out that many examples of pertinent processes would be characterized by non-existing higher moments. Muzy and Bacry (2002) and Bacry and Muzy (2003) go one step further and construct a "fully continuous" class of multifractal measures in which the discreteness of the scales i is replaced by a continuum of scales.

Multiplication over the random weights is then replaced by integration over a similar conical domain in (t, s) space whose extension is given by the maximum correlation scale T (see below). Muzy and Bacry (2002) show that for this set-up, nontrivial multifractal behavior is obtained if the conical subset $C_s(t)$ of the (t, s)-half plane (note that $t \geq 0$) obeys:

$$C_s(t) = \{(t', s'),\ s' \geq s,\ -f(s')/2 \leq t' - t \leq f(s')/2\} \tag{2.13}$$

[10] The conical widening of the influence of scales being the continuous limit of the dependencies across levels in the discrete case that proceeds with, e.g., a factor 2 in the case of binary cascades.

with

$$f(s) = \begin{cases} s & \text{for } s \le T \\ T & \text{for } s > T, \end{cases} \tag{2.14}$$

i.e. a symmetrical cone around current time t with linear expansion of the included scales s up to some maximum T. The multifractal measure obtained along these lines involves a stochastic integral over the domain $C(t)$:

$$d\theta(t) = e^{\int_{(t',s)\in C(t)} d\omega(t',s)}. \tag{2.15}$$

If $d\omega(t', s)$ is a Gaussian variable, one can use this approach as an alternative way to generate a Lognormal multifractal time transformation. As demonstrated by Bacry and Muzy (2003) subordinating a Brownian motion to this process leads to a compound process that has a distribution identical to the limiting distribution of the grid-bound MMAR with Lognormal multipliers for $k \to \infty$. Discretization of the continuous-time multifractal random walk will be considered below.

2.3.3. Multifractal Models in Discrete Time[11]

2.3.3.1. Markov-Switching Multifractal Model

Together with the continuous-time Poisson multifractal, Calvet and Fisher (2001a) have also introduced a discretized version of this model, that has become the most frequently applied version of the multifractal family in the empirical financial literature. In this discretized version, the volatility dynamics can be interpreted as a discrete-time Markov-switching process with a large number of states. In their approach, returns are modeled like in *eq.* (2.1) with innovations ε_t drawn from a standard Normal distribution $N(0, 1)$ and instantaneous volatility being determined by the product of k volatility components or multipliers $M_t^{(1)}, M_t^{(2)}, \ldots, M_t^{(k)}$ and a constant scale factor σ:

$$r_t = \sigma_t \varepsilon_t \tag{2.16}$$

with

$$\sigma_t^2 = \sigma^2 \prod_{i=1}^{k} M_t^i. \tag{2.17}$$

The volatility components M_t^i are persistent, non-negative and satisfy $\mathbb{E}[M_t^i] = 1$. Furthermore, it is assumed that the volatility components $M_t^{(1)}, M_t^{(2)}, \ldots, M_t^{(k)}$ at a given time t are statistically independent. Each volatility component is renewed at time t with probability γ_i depending on its rank within the hierarchy of multipliers and remains unchanged with probability $1 - \gamma_i$. They show that with the following specification of transition probabilities between integer time steps,

[11] We note in passing that for standard discrete volatility models, the determination of the continuous-time limit is not always straightforward. For instance, for GARCH(1,1) model Nelson (1990) found a limiting "GARCH diffusion" under some assumptions while Corradi (2000) found a limiting deterministic process under a different set of assumptions. Also, while there exists a well-known class of continuous-time stochastic volatility models, these do not necessarily constitute the limit processes of their also well-known discrete counterparts.

a discretized Poisson multifractal converges to the continuous-time limit as defined above for $\Delta t \rightarrow 0$:

$$\gamma_i = 1 - (1 - \gamma_1)^{(b^{i-1})}, \tag{2.18}$$

with γ_1 the component at the lowest frequency that subsumes the Poisson intensity parameter λ, $\gamma_1 \in [0, 1]$, and $b \in (1, \infty)$. Calvet and Fisher (2004a) assume a Binomial distribution for M_t^i with parameters m_0 and $2 - m_0$ (thus, guaranteeing an expectation of unity for all M_t^i). If convergence to the limit of the Poisson multifractal is not a concern, one could also use a less parameterized form such as

$$\gamma_i = b^{-i}. \tag{2.19}$$

Here, volatility components in a lower frequency state will be renewed b times as often as those of its predecessor. An iterative discrete multifractal with such a progression of transition probabilities and otherwise identical to the model of Calvet and Fisher (2001a, 2004a) has already been proposed by Breymann et al. (2000).

For the distribution of the multipliers M_t^i, extant literature has also used the Lognormal distribution (cf. Liu et al., 2008; Lux, 2008) with parameters λ and s, i.e.

$$M_t^{(i)} \sim LN(-\lambda, s^2). \tag{2.20}$$

Setting $s^2 = 2\lambda$ guarantees $\mathbb{E}[M_t^i] = 1$. Comparison of the performance and statistical properties of MF models with Binomial and Lognormal multipliers shows typically almost identical results (Liu, di Matteo, and Lux, 2007). It, thus, appears that the Binomial choice (with 2^k different volatility regimes) has sufficient flexibility and cannot easily be outperformed via a continuous distribution of the multipliers.

In *Fig.* 2.5 the first three panels show the development of the switching behavior of Lognormal MSM process at different levels. The average duration of the second highest component is equal to 2048. As a result one expects this component to switch on average two times during the 4096 time-steps of the simulation. Similarly, for the sixth highest component displayed in the second panel renewal occurs about once within $2^5 = 32$ periods. The last panel shows the product of multipliers (displayed in the second from bottom) that plays the role of local stochastic volatility as described by *eq.* (2.17). The resulting artificial time series displays volatility clustering and outliers which stem from intermittent bursts of extreme volatility.

Due to its restriction to a finite number of cascade steps, the MSM is not characterized by asymptotic (multi-) scaling. However, its pre-asymptotic scaling regime can be arbitrarily extended by increasing the number of hierarchical components k. It is, thus, a process whose multifractal properties are spurious. However, at the same time it can be arbitrarily close to "true" multi-scaling over any finite length scale. This feature is shared by a second discretization, the multifractal random walk, whose power-law scaling over a finite correlation horizon is already manifest in its generating process.

Figure 2.5 about here

2.3.3.2. Multifractal Random Walk

In the econophysics literature, a different type of causal, iterative process has been developed more or less simultaneously, denoted the Multifractal Random Walk (MRW). Essentially, the MRW is a Gaussian process with built-in multifractal scaling via an appropriately defined correlation function. While one could use various distributions for the multipliers as the guideline for construction of different versions of MRW replicating their particular autocorrelation structures, the literature has exclusively focused on the Lognormal distribution.

Bacry et al. (2001) define the MRW as a Gaussian process with a stochastic variance as follows:

$$r_{\Delta t}(\tau) = e^{\omega_{\Delta t}(\tau)}\varepsilon_{\Delta t}(\tau), \tag{2.21}$$

with Δt a small discretization step, $\varepsilon_{\Delta t}(\cdot)$ a Gaussian variable with mean zero and variance $\sigma^2\Delta t$ and $\omega_{\Delta t}(\cdot)$ the logarithm of the stochastic variance and τ a multiple of Δt along the time axis. Assuming that $\omega_{\Delta t}(\cdot)$ also follows a Gaussian distribution, one obtains Lognormal volatility draws. For longer discretization steps (e.g. daily unit time intervals), one obtains their returns as:

$$r_{\Delta t}(t) = \sum_{i=1}^{t/\Delta t} \varepsilon_{\Delta t}(i) * e^{\omega_{\Delta t}(i)}. \tag{2.22}$$

To mimic the dependency structure of a Lognormal cascade, these are assumed to have covariances:

$$\mathbb{C}ov(\omega_{\Delta t}(t)\omega_{\Delta t}(t+h)) = \lambda^2 \ln \rho_{\Delta t}(h), \tag{2.23}$$

with

$$\rho_{\Delta t}(h) = \begin{cases} \frac{T}{(|h|+1)\Delta t}, & \text{for } |h| \leq \frac{T}{\Delta t} - 1 \\ 0, & \text{otherwise} \end{cases} \tag{2.24}$$

Hence, T is the assumed finite correlation length (a parameter to be estimated) and λ^2 is called the intermittency coefficient characterizing the strength of the correlation.

In order for the variance of $r_{\Delta t}(t)$ to converge, $\omega_{\Delta t}(\cdot)$ is assumed to obey:

$$\mathbb{E}(\omega_{\Delta t}(i)) = -\lambda^2 \ln(T/\Delta t) = -\mathbb{V}ar(\omega_{\Delta t}(i)). \tag{2.25}$$

Assuming a finite decorrelation scale (rather than a monotonic hyperbolic decay of the autocorrelation) serves to guarantee stationary of the multifractal random walk. Similar as the MSM introduced by Calvet and Fisher (2001a), the MRW model does, therefore, not obey an exact scaling function like *eq.* (2.7) in the limit $t \to \infty$ or divergence of its spectral density at zero, but is characterized by only "apparent" long-term dependence over a bounded interval. The advantage of both models is that they possess "nice" asymptotic properties that facilitate application of many standard tools of statistical inference.

As shown by Muzy and Bacry (2002) and Bacry et al. (2008) the continuous-time limit of MRW (mentioned above in 2.3.2.1.3) can also be interpreted as a time transformation of a Brownian motion subordinate to a log-normal multifractal random measure. For this purpose, the MRW can be reformulated in a similar way like the MMAR model.

$$r(t) = B\left[\theta(t)\right], \quad \text{for all} \quad t \geq 0, \tag{2.26}$$

where $\theta(t)$ is a random measure for the transformation of chronological to "business time" and $B(t)$ is a Brownian motion independent of θ_t. "Business time" θ_t is obtained along the lines of the above exposition of the MRW model as

$$\theta(t) = \lim_{\Delta \to 0} \int_0^t e^{2\omega_\Delta(u)} du. \tag{2.27}$$

Here $\omega_\Delta(u)$ is the stochastic integral of Gaussian white noise $dW(s,t)$ over a continuum of scales s truncated at the smallest and largest scales Δ and T which leads to a cone-like structure defining $\omega_\Delta(u)$ as the area delimited in time (over the correlation length) and a continuum of scales s in the (t, s) plane:

$$\omega_\Delta(u) = \int_\Delta^T \int_{u-s}^{u+s} dW(v, s). \tag{2.28}$$

To replicate the weight structure of the multipliers in discrete multifractal models, a particular correlation structure of the Gaussian elements $dW(v, s)$ needs to be imposed. Namely, the multifractal properties are obtained for the following choices of the expectation and covariances of $dW(v, s)$:

$$\mathbb{C}ov\left(dW(v, s), dW(v', s')\right) = \lambda^2 \delta(v - v')\delta(s - s')\frac{dvds}{s^2} \tag{2.29}$$

and

$$\mathbb{E}\left(dW(v, s)\right) = -\lambda^2 \frac{dvds}{s^2}. \tag{2.30}$$

Muzy and Bacry (2002) and Bacry and Muzy (2003) show that the limiting continuous-time process exists and possesses multifractal properties. Interestingly, Muzy et al. (2006) and Bacry et al. (2013) also provide results for the unconditional distribution of returns obtained from this process. They demonstrate that it is characterized by fat tails and that it becomes less heavy tailed under time aggregation. They also show that standard estimators of tail indices are ill-behaved for data from a MRW data-generating process due to the high dependency of adjacent observations. While the implied theoretical tail indices with typical estimated parameters of the MRW would be located at unrealistically large values (> 10), taking the dependency in finite samples into account one obtains biased (pseudo-)empirical estimates indicating much smaller values of the tail index that are within the order of magnitude of empirical ones. A similar mismatch between implied and empirical tail indices applies to other multifractal models as well (as far as we can see, this is not explicitly reported in extant literature, but has been mentioned repeatedly by researchers) and can be likely explained in the same way.

2.3.3.3. Asymmetric Univariate MF Models

All previous models are designed in a completely symmetric way for positive and negative returns. However, it is well known that price fluctuations in asset markets exhibit a certain degree of asymmetry due to leverage effects. The discrete-time skewed multifractal random walk (DSMRW) model proposed by Pochart and Bouchaud (2002) is an extended version of the MRW, that takes account of such asymmetries. The model is defined similarly as the MRW of *eq.* (2.21) but incorporates a direct influence of past realizations on contemporaneous volatility

$$\tilde{\omega}_{\Delta t}(i) \equiv \omega_{\Delta t}(i) - \sum_{k<i} K(k,i)\varepsilon_{\Delta t}(k), \tag{2.31}$$

where Pochart and Bouchaud propose to use $K(k,i) = \frac{K_0}{(i-k)^{\alpha}\Delta t^{\beta}}$ is a positive definite kernel for the influence of returns on subsequent volatility. Bacry et al. (2012) proposed a continuous-time skewed multifractal model that also incorporates the leverage effect.

Eisler and Kertész (2004) expand the MSM model in a similar way. They consider a refined version of the model in which asymmetry comes in via the renewal probabilities and, in addition, use a term inspired by *eq.* (2.31) to account for leverage autocorrelations.

An asymmetric MSM model has also been introduced by Calvet et al. (2013). They embed a multifractal cascade into a stochastic volatility model where the product of multipliers enters as a time-varying long-run anchor for the volatility dynamics while at the same time governing a jump component in returns that relates positive volatility shocks to negative return shocks.

2.3.3.4. Bivariate Multifractal Models

A bivariate MF model has first been introduced by Calvet et al. (2006). Consider a portfolio of two assets α and β. Let now denote r_t the vector of log-returns of the portfolio, and r_t^{α} and r_t^{β} the individual log-returns of the two assets, respectively. Following Calvet, Fisher, and Thompson the return of the portfolio is modeled as:

$$r_t = [g(M_t)]^{1/2} * \varepsilon_t, \tag{2.32}$$

where $g(M_t)$ denotes a 2×1 vector $M_{1,t} * M_{2,t} * \cdots * M_{k,t}$, $*$ denotes element by element multiplication and the column vectors $\varepsilon_t \in \mathbb{R}^2$ are *i.i.d.* Gaussian $N(\mathbf{0}, \mathbf{\Sigma})$ with covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{\alpha}^2 & \rho_{\varepsilon}\sigma_{\alpha}\sigma_{\beta} \\ \rho_{\varepsilon}\sigma_{\alpha}\sigma_{\beta} & \sigma_{\beta}^2 \end{bmatrix}. \tag{2.33}$$

ρ_ε represents the unconditional correlation between the residuals as the first source of correlation between both returns. The period t volatility state is characterized by a $2 \times k$ matrix $M_t = (M_{1,t}; M_{2,t}; \ldots; M_{k,t})$ and the vector of the components at the i^{th} frequency is $M_{i,t} = (M^\alpha_{i,t}\ M^\beta_{i,t})$. The volatility vectors $M_{i,t}$ are non-negative and satisfy $E[M_{i,t}] = \mathbf{1}$, where $\mathbf{1} = (1,1)'$. Economic intuition behind the choice of the dynamics for each vector $M_{i,t}$ is that volatility arrivals are correlated but not necessarily simultaneous across markets. For this reason Calvet, Fisher, and Thompson allow arrivals across series to be linked by a correlation coefficient λ. Consider two random variables $I^\alpha_{i,t}$ and $I^\beta_{i,t}$ which are equal to 1 if each series $c \in \{\alpha, \beta\}$ is hit by an information arrival with probability γ_i, and equal to zero otherwise. Calvet, Fisher, and Thompson specified the arrival vector to be *i.i.d.* and assumed its unconditional distribution to satisfy three conditions. First, the arrival vector is symmetrically distributed: $(I^\alpha_{i,t}, I^\beta_{i,t}) \overset{d}{=} (I^\beta_{i,t}, I^\alpha_{i,t})$. Second, the switching probabilities of both series are equal for each level i: $Pr(I^\alpha_{i,t} = 1) = Pr(I^\beta_{i,t} = 1) = \gamma_i$, with γ_i following *eq.* (2.18) as for univariate MSM. Third, there exists $\lambda \in [0, 1]$ such that

$$Pr(I^\alpha_{i,t} = 1 | I^\beta_{i,t} = 1) = (1 - \lambda)\gamma_i + \lambda.$$

These three conditions define a unique distribution of $(I^\alpha_{i,t}, I^\beta_{i,t})$ whose joint switching probabilities can be easily determined. Note that the univariate dynamics of each series coincides with a univariate MSM model. Idier (2011) proposed an extension of the bivariate MSM model by considering a time dependent covariance for the vector of residuals $\rho_\varepsilon(t)$.

Liu (2008) considered a closely related bivariate multifractal model built upon the assumption that two time series have a certain number of joint cascade levels in common, while the remaining ones are chosen independently. The returns are, then, modeled as:

$$r_{q,t} = \left[\left(\prod_{i=1}^{k} M_{i,t} \right) \left(\prod_{l=k+1}^{n} M_{l,t} \right) \right]^{1/2} * \varepsilon_t, \tag{2.34}$$

where $q = 1, 2$ refers to the two time series, both having an overall number of n levels of their volatility cascades, and they share a number k of joint cascade levels which govern the strength of their volatility correlation. Obviously, the larger k, the more correlation between the volatility dynamics of both series. After k joint multiplicators, each series has separate additional multifractal components. ε_t is defined as in *eq.* (2.32) to follow a bivariate standard Normal distribution with correlation parameter ρ_ε. This model can be seen as a special case of a slightly generalized version of Calvet et al. (2006) allowing for heterogeneity of the correlation of volatility innovations, λ_i, across hierarchical levels and choosing an extreme specification in that part of the $\lambda_i (1 \leq i \leq k)$ are equal to 1 and the remaining ones are equal to 0. Liu and Lux (2014) show that the distinction between different degrees of correlation between volatility innovations indeed improves the fit and performance of the bivariate MSM, but the extreme specification of Liu (2008) with alternation between full dependence and lack of correlation is dominated by a more flexible approach. Interestingly, whether high or low frequency components are more correlated differs between markets.

2.3.3.5. Higher Dimensional Multifractal Models

The bivariate models presented above can be generalized for more than two assets in various ways. Liu (2008)'s approach can be generalized in a straightforward way to an N-dimensional asset returns process. If one assumes that the N time series share a number of j joint cascades that govern the strength of their volatility correlation, the correlation of volatility arrivals could be generalized to the case of an arbitrary number of assets without having to add new parameters in the volatility part of the model. Additional parameters would, then, only come in via the correlation of the Gaussian innovations. If such a specification appears insufficient to capture the heterogeneity in return fluctuations across assets, one could consider a generalized framework with asset-specific multifractal parameter, m_0 or λ in the Binomial or Lognormal setting, respectively.

A generalization of the MRW in a similar vein had already been proposed by Bacry et al. (2000). They suggest to extend the MRW model to a multivariate Multifractal Random Walk (MMRW) in order to model portfolio behavior. Let $\boldsymbol{X}_t$ be a MMRW, then following Bacry, Delour, and Muzy $\boldsymbol{X}_t$ is defined as:

$$\boldsymbol{X}(t) = \lim_{t \to 0} \boldsymbol{X}_{\Delta t}(t) = \lim_{t \to 0} \sum_{k=1}^{t/\Delta t} \epsilon_{\Delta t}[k] * e^{\omega_{\Delta t}[k]}, \tag{2.35}$$

where $\epsilon_{\Delta t}$ is now a vector of Gaussians with zero mean and variance-covariance function at lag τ $\mathbb{C}ov(\epsilon_{i,\Delta t}(t), \epsilon_{j,\Delta t}(t+\tau)) = \delta(\tau)\boldsymbol{\Sigma}_{ij}\Delta t$. The magnitude process $\omega_{\Delta t}(\cdot)$ is also Gaussian with covariance $\mathbb{C}ov(\omega_{i,\Delta t}(t), \omega_{j,\Delta t}(t+\tau)) = \boldsymbol{\Lambda}_{ij}\ln(T_{ij}/(\Delta t + |\tau|))$ for $(\Delta t + |\tau| < T_{ij})$ and 0 elsewhere. The matrix $\boldsymbol{\Lambda}$, labeled "multifractal matrix", controls the non-linearity of the multifractal spectrum, and T_{ij} are different correlation lengths for the autocovariances and cross-covariances characterizing the process.

2.4. Shortcomings of MF Models

As we mention above MF models are successful in capturing simultaneously most of the stylized facts of financial data. Although this capability to properly reproduce the data, MF models suffer from some limitations which we cite here. All the extant MF models in the literature are strictly invariant under time reversal symmetry. This feature makes the MF models inapt to capture the so-called leverage effect, and the asymmetric structure of the correlations between the past and future volatilities at different time scales. Pochart and Bouchaud (2002) construct the skewed MRW model that can reproduce the leverage effect, but cannot take into account the asymmetric lead-lag correlation of volatilities. The latter stylized fact is observed to be present in all markets, even when the leverage effect is almost nonexistent. Calvet and Fisher (2001a) suggest that by relaxing the independence assumption of $B(t)$ and $\theta(t)$ in their model, one would obtain a model that can capture the leverage effects.

Although the MF models developed by Mandelbrot et al. (1997), Calvet and Fisher (2001a, 2004a) are able to display all the stylized facts observed in the financial markets (volatility clustering, fat-tails, long memory, multi-scaling,..), it is important to notice that all these models cannot properly capture leverage effects.[12] One can design the models in a way to take into account leverage effects by removing the assumption that $B(t)$ and $\theta(t)$ are independent. The relaxation of this assumption would introduce asymmetric structure in the model.

Another shortcoming of the MRW models is that the theoretical values of tail index obtained in the MRW models are higher than the empirical ones found in the range 3 to 5 for many stocks in different markets. Bacry et al. (2013) tried to design the MRW models, so that the theoretical values of the tail index approximate the empirical ones. Unfortunately, however, they find that a change of the theoretical value for the tail index will change the way the ergodicity breaking happens in the models.

In the MRW models it is assumed that the local volatility is Lognormal distributed. While a couple of studies speak in the favor of Lognormal distribution, other studies suggest to use other distributions, such as an inverse gamma distribution that fits data equally well, or even better (cf. Miccichè et al., 2002; Bouchaud and Potters, 2004). Studies of various empirical log-volatility correlation functions in the MRW models provide results that do not perfectly match with the fact the intermittency coefficient estimated from the curvature of the scaling function and the slope of the log-volatility covariance logarithmic decrease have to be equal as the models predict (cf. Arneodo et al., 1998). Another result from these studies is that the integral parameter T, i.e. the large cut-off time scale beyond which volatility correlation disappears, is on the scale of a few years (cf. Muzy et al., 2000), for instance T is larger than one year for both intraday and daily financial data (cf. Bacry and Muzy, 2010).

2.5. Estimation and Forecasting

Availability of efficient estimation procedures is essential for the application of theoretical asset-pricing models for practical purposes. The non-standard format of multifractal models has initially cast doubts on the applicability of many well-known statistical tools to this new family of volatility models. Fortunately, the members of the second generation multifractal models (MSM and MRW) seemed to be much more well-behaved (and have partially be designed to be so) in terms of asymptotic statistical behavior. Most effort has been spent so far to find stable and efficient inference methods for the discrete time MSM model with discrete or continuous distributions for multipliers or volatility components. In the following we present the estimation methods most often applied for MF models. We dispense with the traditional $f(\alpha)$ and $\tau(q)$ approach to inference which has been covered in detail in sec. 2.2.4. As it soon turned out in the pertinent literature when starting to adapt multifractal models to finance, the scaling-approach provides potentially very biased and volatile estimates in applications to financial data, and due to their fat tails, would even indicate existence of multifractal structure after randomization of such time series. The quest

[12] The leverage effect corresponds to the fact that the variation of the log-return in the past is negatively correlated with the volatility (the squared or absolute log-return) in the future

for more appropriate statistical methods has been motivated to a large extent by these deficiencies. The development of the Markov-switching multifractal model and the multifractal random walk have brought forward stochastic processes with more "convenient" asymptotic properties than their predecessors. As a consequence, they allow application of many established tools of inference. Nevertheless, their proximity to genuine long-memory might still be a concern and motivates to exert caution in empirical applications (e.g., while theoretical convergence of estimates might be trivially guaranteed, the pre-asymptotic regime might be much more extended than with other models).

2.5.1. Maximum Likelihood Estimation

Exact ML estimation has been primarily developed for the discrete-time MSM model with a discrete distribution for the volatility components or multipliers. Calvet and Fisher (2004a) introduced an ML estimation approach for the Binomial Markov-switching multifractal (BMSM) model. To show how to perform ML estimation in this context, note that the log-likelihood (L) function for a series of observations $\{r_t\}_{t=1}^{T}$ in its most general form may be expressed as:

$$L(r_1, \dots, r_T; \varphi) = \sum_{t=1}^{T} \ln g(r_t | r_1, \dots, r_{t-1}; \varphi), \tag{2.36}$$

where $g(r_t|r_1, \dots, r_{t-1}; \varphi)$ is the likelihood function of the Markov-switching multifractal model, and φ is the vector of parameters. For Markov-switching models, the likelihood function can be decomposed in the following way: $g(r_t|r_1, \dots, r_{t-1}; \varphi) = \omega_t(r_t|M_t = m^i, \varphi)(\pi_{t-1}A)$. The three components are defined as follows: $\omega_t(r_t|M_t = m^i, \varphi)$ is a vector of dimension 2^k of conditional densities of any observation r_t for volatility regimes m^i and A is the transition matrix which has components $A_{ij} = Pr(M_{t+1} = m^j|M_t = m^i)$. The last component within the likelihood function above is π_t, which is the vector of conditional probabilities of the volatility states given observations $\pi_t^i = Pr(M_t = m^i|r_1, \dots, r_t; \varphi)$. The conditional probabilities can be recursively obtained through Bayesian updating

$$\pi_t = \frac{\omega_t(r_t|M_t = m^i, \varphi) * (\pi_{t-1}A)}{\sum \omega_t(r_t|M_t = m^i, \varphi) * (\pi_{t-1}A)}. \tag{2.37}$$

Different distributional assumptions for innovations could be embedded in this framework. The parameter vector of the BMSM with Gaussian innovations would be given by $\varphi = (m_0, \sigma)'$, while the parameter vector of a BMSM with Student-t innovations would be $\varphi = (m_0, \sigma, \nu)'$ where $\nu \in (2, \infty)$ is the distributional parameter accounting for the degrees of freedom in the density function of the Student-t distribution. The Student$-t$ distribution for return innovations has been used by Lux and Morales-Arias (2010) in order to enhance out-of-sample forecasts of the MSM model because it may allow the MSM model to better distinguish between volatility dependence and fat-tailed innovations.

An advantage of the ML procedure is that, as a by-product, it allows one to obtain optimal forecasts via Bayesian updating of the conditional probabilities $\pi_t = Pr(M_t = m^i | r_1, \ldots, r_t; \varphi)$ for the unobserved volatility states m^i, $i = 1, \ldots, 2^k$. ML estimation provides good precision in finite samples (cf. Calvet and Fisher, 2004a).

Although the applicability of the ML algorithm greatly facilitates estimation of MSM models, it is restrictive in the sense that it is practically feasible only for discrete distributions of the multipliers and, therefore, is not applicable for e.g., the case of a Lognormal distribution. Due to the potentially large state space (we have to take into account transitions between 2^k distinct states), ML estimation also encounters practical bounds of its computational demands for specifications with more than about $k = 10$ volatility components in the Binomial case. For multivariate MF models, the applicability of the ML approach is even more constrained from the computational side: In the bivariate case the evaluation of its transition matrix with size $4^k \times 4^k$ becomes unfeasible for choices of about $k > 5$. There has also been a recent attempt to estimate the MRW model via a likelihood approach. Løvsletten and Rypdal (2012) develop an approximate maximum likelihood method for MRW using a Laplace approximation of the likelihood function.

2.5.2. Simulated Maximum Likelihood

This approach is more broadly applicable to both discrete and continuous distributions for multipliers. To overcome the computational and conceptional limitation of exact ML estimation, Calvet et al. (2006) developed a simulated ML approach. They propose a particle filter to numerically approximate the likelihood function. The particle filter is a recursive algorithm that generates independent draws $M_t^{(1)}, \ldots, M_t^{(N)}$ from the conditional distribution of π_t. At time $t = 0$, the algorithm is initiated by draws $M_0^{(1)}, \ldots, M_0^{(N)}$ from the ergodic distribution $\bar{\pi}$. For any $t > 0$, the particles $\{M_t^{(n)}\}_{n=1}^N$ are sampled from the new belief π_t. To this end, the formula (2.37) within the ML estimation algorithm is replaced by a Monte Carlo approximation in SML. This means that the analytical updating via the transition matrix, $\pi_{t-1}A$, is approximated via the simulated transitions of the particles. Disregarding the normalization of probabilities (i.e., the denominator), the formula (2.37) can be rewritten as

$$\pi_t^i \propto \omega_t(r_t | M_t = m^i; \varphi) \sum_{j=1}^{4^k} Pr\left(M_t = m^i | M_{t-1} = m^j\right) \pi_{t-1}^j, \tag{2.38}$$

and due to the fact that $M_t^{(1)}, \ldots, M_t^{(N)}$ are independent draws from π_{t-1}, the Monte Carlo approximation has the following format:

$$\pi_t^i \propto \omega_t(r_t | M_t = m^i; \varphi) \frac{1}{N} \sum_{n=1}^{N} \Pr\left(M_t = m^i | M_{t-1} = M_{t-1}^{(n)}\right). \tag{2.39}$$

The approximation, thus, proceeds by simulating each $M_{t-1}^{(n)}$ one step forward to obtain $\hat{M}_t^{(n)}$ given $M_{t-1}^{(n)}$. This step only uses information available at date $t-1$, and must therefore be adjusted at time step t to account for the information contained in the new return. This is achieved by drawing

N random numbers q from 1 to N with probability

$$Pr(q = n) \equiv \frac{\omega_t(r_t|M_t = \hat{M}_t^{(n)}; \varphi)}{\sum_{n'=1}^{N} \omega_t(r_t|M_t = \hat{M}_t^{(n')}; \varphi)}. \tag{2.40}$$

The distribution of particles is, thus, shifted according to their importance at time t. With simulated draws $M_t^{(n)}$ the Monte Carlo (MC) estimate of the conditional density is

$$\hat{g}(r_t|r_1, \ldots, r_{t-1}; \varphi) \equiv \frac{1}{N} \sum_{n=1}^{N} g(r_t|M_t = \hat{M}_t^{(n)}; \varphi), \tag{2.41}$$

and the log-likelihood is approximated by $\sum_{t=1}^{T} \ln \hat{g}(r_t|r_1, \ldots, r_{t-1}; \varphi)$. The simulated ML approach makes it feasible to estimate MSM models with continuous distribution of multipliers as well as univariate and multivariate Binomial models with too high a number of states for exact ML. Despite this gain in terms of different specifications of MSM models that can be estimated, the computational demands of SML are still considerable, particularly for high numbers of particles N.

2.5.3. GMM Estimation

Again, this is an approach that is, in principle, applicable for both discrete and continuous distributions for multipliers. To overcome the lack of practicability of ML estimation, Lux (2008) introduced a Generalized Method of Moments (GMM) estimator that is also universally applicable to all specifications of MSM processes (discrete or continuous distribution for multipliers, Gaussian, Student$-t$ or various other distributions for innovations). In particular, it can be used in all those cases where ML is not applicable or computationally unfeasible. Its computational demands are also lower than those of SML and independent of the specification of the model. In the GMM framework for MSM models, the vector of parameters φ is obtained by minimizing the distance of empirical moments from their theoretical counterparts, i.e.

$$\hat{\varphi}_T = \arg\min_{\varphi \in \Phi} f_T(\varphi)' A_T f_T(\varphi), \tag{2.42}$$

with Φ the parameter space, $f_T(\varphi)$ the vector of differences between sample moments and analytical moments, and A_T a positive definite and possibly random weighting matrix. Moreover, $\hat{\varphi}_T$ is consistent and asymptotically Normal if suitable "regularity conditions" are fulfilled (cf. Harris and Mátyás, 1999) which are satisfied routinely for Markov processes.

In order to account for the proximity to long memory characterizing MSM models, Lux (2008) proposed to use log differences of absolute returns together with the pertinent analytical moment conditions, i.e.

$$\xi_{t,T} = \ln|r_t| - \ln|r_{t-T}|. \tag{2.43}$$

The above variable only has nonzero auto-covariances over a limited number of lags. To exploit the temporal scaling properties of the MSM model, covariances of various moments over different time horizons are chosen as moment conditions, i.e.

$$\text{Mom}(T, q) = \mathbb{E}\left[\xi^q_{t+T,T} \cdot \xi^q_{t,T}\right], \tag{2.44}$$

for $q = 1, 2$ and different horizons T together with $\mathbb{E}[r_t^2] = \sigma^2$ for identification of σ in the MSM model with Normal innovations. In the case of the MSM-t model, Lux and Morales-Arias (2010) consider additional moment conditions in addition to (2.44), namely, $\mathbb{E}[|r_t|]$, $\mathbb{E}[|r_t^2|]$, $\mathbb{E}[|r_t^3|]$, in order to extract information on the Student$-t$'s shape parameter.

Bacry et al. (2008) and Bacry et al. (2013) also apply the GMM method for estimating the MRW parameters ($\lambda, \sigma,$ and T) using similar moments as in Lux (2008). Sattarhoff (2010) refines the GMM estimator for the MRW using a more efficient algorithm for the covariance matrix estimation. Liu (2008) adopts the GMM approach to bivariate and trivariate specifications of the MSM model. Leövey (2013) develops a simulated method of moments (SMM) estimator for the continuous-time Poisson multifractal model of Calvet and Fisher (2001a).

Related work in statistical physics has recently also considered simple moment estimators for extraction of the multifractal intermittency parameters from data of turbulent flows (cf. Kiyono et al., 2007). Leövey and Lux (2012) compare the performance of a GMM estimator for multifractal models of turbulence with various heuristic estimators proposed in the pertinent literature, and show that the GMM approach typically provides more accurate estimates due to its more systematic exploitation of information contained in various moments.

2.5.4. Forecasting

With ML and SML estimates, forecasting is straightforward: With ML estimation, conditional state probabilities can be iterated forward via the transition matrix to deliver forecasts over arbitrarily long time horizons. The conditional probabilities of future multipliers given the information set $\mathfrak{I}_t$, $\hat{\pi}_{t,n} = P(M_n|\mathfrak{I}_t)$, are given by

$$\hat{\pi}_{t,n} = \pi_t A^{n-t}, \quad \forall n \in \{t, \ldots, T\}. \tag{2.45}$$

In the case of SML, iteration of the particles provides an approximation to the predictive density. Since GMM does not provide information on conditional state probabilities, Bayesian updating is not possible and one has to supplement GMM estimation with a different forecasting algorithm. To this end, Lux (2008) proposes best linear forecasts (cf. Brockwell and Davis, 1991, chap. 5) together with the generalized Levinson-Durbin algorithm developed by Brockwell and Dahlhaus (2004). Assuming that the data of interest (e.g., squared or absolute returns) follow a stationary process $\{Y_t\}$ with mean zero, the best linear h-step forecasts are obtained as

$$\hat{Y}_{n+h} = \sum_{i=1}^{n} \phi^{(h)}_{ni} Y_{n+1-i} = \boldsymbol{\phi}^{(h)}_n \boldsymbol{Y}_n, \tag{2.46}$$

where the vectors of weights $\boldsymbol{\phi}_n^h = (\phi_{n1}^h, \phi_{n2}^h, \ldots, \phi_{nn}^h)'$ can be obtained from the analytical auto-covariances of Y_t at lags h and beyond. More precisely, $\boldsymbol{\phi}_n^{(h)}$ are any solution of $\boldsymbol{\Psi}_n \boldsymbol{\phi}_n^{(h)} = \boldsymbol{\kappa}_n^{(h)}$ where $\boldsymbol{\kappa}_n^{(h)} = (\kappa_{n1}^{(h)}, \kappa_{n2}^{(h)}, \ldots, \kappa_{nn}^{(h)})'$ denote the auto-covariances of Y_t and $\boldsymbol{\Psi}_n = [\kappa(i-j)]_{i,j=1,\ldots,n}$ is the variance-covariance matrix. In empirical applications, *eq.* (2.46) has been applied for forecasting squared returns as a proxy for volatility using analytical covariances to obtain the weights $\boldsymbol{\phi}_n^h$. Linear forecasts have also been used by Bacry et al. (2008) and Bacry et al. (2013) in connection with their GMM estimates of the parameters of the MRW model. Duchon et al. (2012) develop an alternative forecasting scheme for the MRW model in the presence of parameter uncertainty as a perturbation of the limiting case of an infinite correlation length $T \to \infty$.

2.6. Empirical Applications

Calvet and Fisher (2004a) compare the forecast performance of the MSM model to those of GARCH, MS-GARCH, and FIGARCH models across a range of in-sample and out-of-sample measures of fit. Using four long series of daily exchange rates they find that at short horizons MSM shows about the same and sometimes a better performance than its competitors. At long horizons MSM more clearly outperforms all alternative models. Lux (2008) combines the GMM approach with best linear forecasts and compares different MSM models (Binomial MSM and Lognormal MSM with various numbers of multipliers) to GARCH and FIGARCH. Although GMM is less efficient than ML, Lux (2008) confirms the tendency of superior performance of MSM models over GARCH and FIGARCH in forecasting volatility of foreign exchange rates. Similarly promising performance in forecasting volatility and value-at-risk is reported for the MRW model by Bacry et al. (2008) and Bacry et al. (2013). Bacry et al. (2008) find that linear volatility forecasts provided by the MRW model outperform GARCH(1, 1) models. Furthermore, they show that MRW forecasts of the VaR at any time-scale and time-horizon are much more reliable than GARCH(1, 1) (Normal or Student−t) forecasts for both foreign exchange rates and stock indices.

Lux and Kaizoji (2007) investigate the predictability of both volatility and volume for a large sample of Japanese stocks. Using daily data of stock prices and trading volume available over 27 years (from 01/01/1975 to 12/31/2001), they examine the potential of time series models with long memory (FIGARCH, ARFIMA, multifractal) to improve upon the forecasts derived from short-memory models (GARCH for volatlity, ARMA for volume). For both volatility and volume, they find that the MSM model provides much safer forecasts than FIGARCH and ARFIMA and does not suffer from occasional dramatic failures as is the case with the FIGARCH model. This higher degree of robustness of MSM forecasts compared to alternative models is also confirmed by Lux and Morales-Arias (2013). They estimate the typical parameters of GARCH, FIGARCH, SV, LMSV and MSM models from a large sample of stock indices and compare the empirical performance of each model when applied to simulated data of any other model with typical empirical parameters. As it turns out, the MSM almost always comes in second best (behind the true model) when forecasting future volatility and even dominates combined forecasts from many models. It, thus, appears to be relatively safe for practitioners to use the MSM even if it were misspecified and another standard model were the "true" data-generating process.

Lux and Morales-Arias (2010) introduce the MSM model with Student-*t* innovations and compare its forecast performance to those of MSM models with Gaussian innovations, and (FI)GARCH. Using country data on all-share equity indices, government bonds and real estate security indices, they find that the MSM model with Normal innovations produces forecasts that improve upon historical volatility, but are in some cases inferior to FIGARCH with Normal innovations. By adding fat tails to both MSM and FIGARCH, they obtain improvements by MSM models for forecasting volatility while the forecast performance by FIGARCH deteriorates. They find also that one can obtain more accurate volatility forecasts by combining FIGARCH and MSM.

Lux et al. (2014) apply an adapted version of the MSM model to measurements of realized volatility. Using five different stock market indices (CAC 40, DAX, FTSE 100, NYSE Composite and S&P 500), they find that the realized volatility-Lognormal MSM model (RV-LMSM) model performs better than non-RV models (FIGARCH, TGARCH, SV and MSM) in terms of mean-squared errors for most stock indices and at most forecasting horizons. They also point out that similar results are obtained in a certain number of instances when the RV-LMSM model is compared to the popular RV-ARFIMA model and forecast combinations of alternative models (non-RV and RV) could hardly improve upon forecasts of various single models.

Calvet et al. (2006) apply the bivariate model to the comovements of volatility of pairs of exchange rates. They find again that their model provides better volatility and value-at-risk (VaR) forecasts compared to the constant correlation GARCH (CC-GARCH) of Bollerslev (1990). Applying the refined bivariate MSM to stock index data, Idier (2011) confirms the results of Calvet et al. (2006). Additionally, he finds that his refined model shows significantly better performance than the baseline MSM and DCC models for horizons longer than ten days. Liu and Lux (2014) apply the bivariate model to daily data for a collection of bivariate portfolios of stock indices, foreign currencies and U.S. 1 Year and 2 Year Treasury Bonds. They find that the bivariate multifractal model generates better VaR forecasts than the CC-GARCH model, especially in the case of exchange rates, and that an extension allowing for heterogeneous dependency of volatility arrivals across levels improves upon the baseline specification both in in-sample and out-of-sample.

Chen et al. (2013) propose a Markov-switching multifractal duration (MSMD) model. In contrast to the traditional duration models inspired by GARCH-type dynamics, this new model uses the MSM process developed by Calvet and Fisher (2004a), and thus can reproduce the long memory property of durations. By applying the MSMD model to duration data of twenty stocks randomly selected from the S&P 100 index and comparing it with the autoregressive conditional duration (ACD) model both in- and out-of-sample, they find that at short horizons both models yield about the same results while at long horizons the MSMD model dominates over the ACD model.

Baruník et al. (2012) independently develop a Markov-switching multifractal duration (MSMD) model whose specification is slightly different from that proposed by Chen et al. (2013). They also use the MSM process introduced by Calvet and Fisher (2004a) as basic ingredient in the construction of the model. They apply the model to price durations of three major foreign exchange futures contracts and compare the predictive ability of the new model with those of the ACD model and long-memory stochastic duration (LMSD) model of Deo et al. (2006). They find that both LMSD and MSMD forecasts generally outperform the ACD forecasts in terms of the mean square error and mean absolute error. MSMD and LMSD models sometimes exhibit similar forecast performances, sometimes the MSMD model slightly dominates the LMSD model.

Option price applications of multifractal models have started with Pochart and Bouchaud (2002) who show that their skewed MRW model could generate smiles in option prices. Leövey (2013) proposed a "risk-neutral" MSM process in order to extract the parameters of the MSM model from option prices. As it turns out, MSM models backed out from option data add significant information to those estimated from historical return data and enhance the forecast ability of future volatility.

Calvet, Fearnley, Fisher, and Leippold (2013) propose an extension of the continuous-time MSM process which in addition to the key properties of the basic MSM process also incorporates the leverage effect and dependence between volatility states and price jumps. Their model can be conceived as an extension of a standard stochastic volatility model in which long-run volatility is driven by shocks of heterogenous frequency that also trigger jumps in the return dynamics, and, so are responsible for negative correlation between return and volatility. They also develop a particle filter that permits the estimation of the model. By applying the model to option data they find that it can closely reproduce the volatility smiles and smirks. Furthermore, they also find that the model outperforms affine jump-diffusions and asymmetric GARCH-type models in- and out-of-sample by a sizeable margin.

Calvet, Fisher, and Wu (2013) develop a class of dynamic term structure models in which the number of parameters to be estimated is independent of the number of factors selected. This parsimonious design is obtained by a cascading sequence of factors of heterogenous durations that is modeled in the spirit of multifractal models. The sequence of mean reversion rates of these factors follows a geometric progression which is responsible for the hierarchical nature of the cascade in the model. In their empirical application to a bandwidth of LIBOR and swap rates, a cascade model with 15 factors provides a very close fit to the dynamics of the term structure and outperforms random walk and autoregressive specifications in interest rate forecasting.

Taken as a whole, the empirical studies summarized above provide mounting empirical evidence of the superiority of the MF over traditional GARCH models (MS-GARCH, FIGRACH) in terms of forecasting of long-term volatility and related tasks such a VaR assessment. In addition, the model appears quite robust, and has found successful applications in modeling of financial durations, the term structure of interest rates and option pricing.

2.7. Conclusion

The motivation for studying multifractal models for asset price dynamics derives from their built-in properties: Since they generically lead to time series with fat tails, volatility clustering and different degrees of long-term dependence of power transformations of returns, they are able to capture all the universal "stylized facts" of financial markets. In the overview of extant applications above, MF-type models typically exhibit a tendency to perform somewhat better in volatility forecasting and VaR-assessment than the more traditional toolbox of GARCH-type models. Furthermore, multifractal processes appear to be relatively robust to misspecification, they seem applicable to a whole variety of variables of interest from financial markets (returns, volume, durations, interest rates) and are very directly motivated by the universal findings of fact tails, clustering of volatility and anomalous scaling. In fact, multifractal processes constitute the only known class of models in which anomalous scaling is generic while all traditional asset-pricing models have a limiting uni-scaling behavior. Capturing this stylized fact may, therefore, well make a difference - even if one can never be certain that multiscaling is not spuriously caused by an asymptotically unifractal model and although those multifractal models that have become the workhorse in empirical applications (MSM, MRW) are characterized themselves by only preasymptotic multiscaling.

Obviously, the introduction of multifractal models in finance did not unleash as much research activity as that of the GARCH or SV families of volatility models in the decades before. The overall number of contributions in this area is still relatively small and comes from a relatively small group of active researchers only. The reason for this abstinence might be that the first generation of multifractal models might have appeared clumsy and unfamiliar to financial economists. Their non-causal principles of construction along the dimension of different scales of a hierarchical structure of dependencies might have appeared too different from known iterative time series models hitherto applied. In addition, the underlying multifractal formalism (including scaling functions and distribution of Hölder exponents) had been unknown in economics and finance, and application of standard statistical methods of inference to multifractal processes appeared cumbersome or impossible. However, all these obstacles have been overcome with the advent of the second generation of multifractal models (MSM and MRW) that are statistically well-behaved and of an iterative, causal nature. Besides their promising performance in various empirical applications they even provide the additional advantage of having clearly defined continuous-time asymptotics so that applications in dicrete- and in continuous-time can be embedded in a consistent framework.

While the relatively short history of multifractal models in finance has already brought about a variety of specifications and different methodologies for statistical inference, some areas can be identified in which additional work should be particularly welcome and useful. These include: Multivariate MF models, applications of the MF approach beyond the realm of volatility models such as the MF duration model, and its use in the area of derivative pricing.

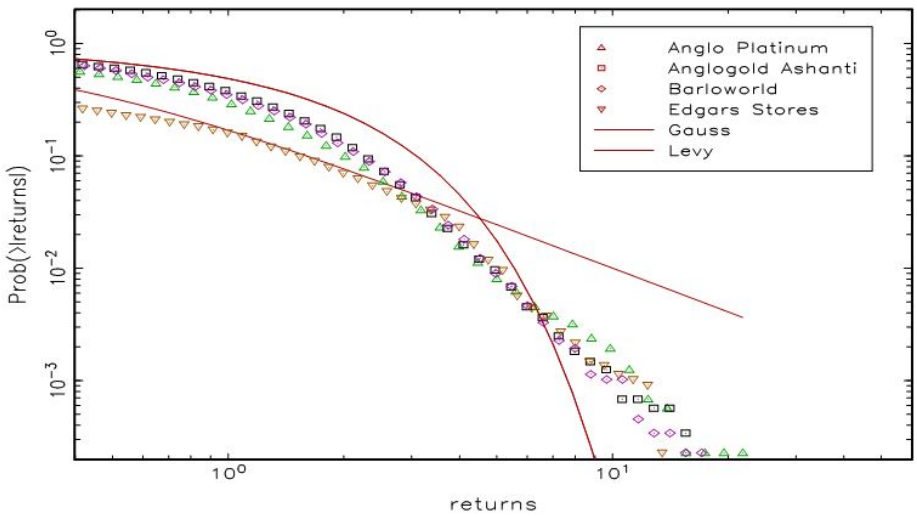

Figure 2.1.: Cumulative distribution for daily returns of four South African stocks (from 1973 until 2006). The solid lines correspond to the Gaussian and Levy distributions. The tail behavior of all stocks is different from that of both the Gaussian and Levy distribution (for the latter, a characteristic exponent $\alpha = 1.7$ has been chosen that is a typical outcome of estimating the parameters of this family of distributions for financial data).

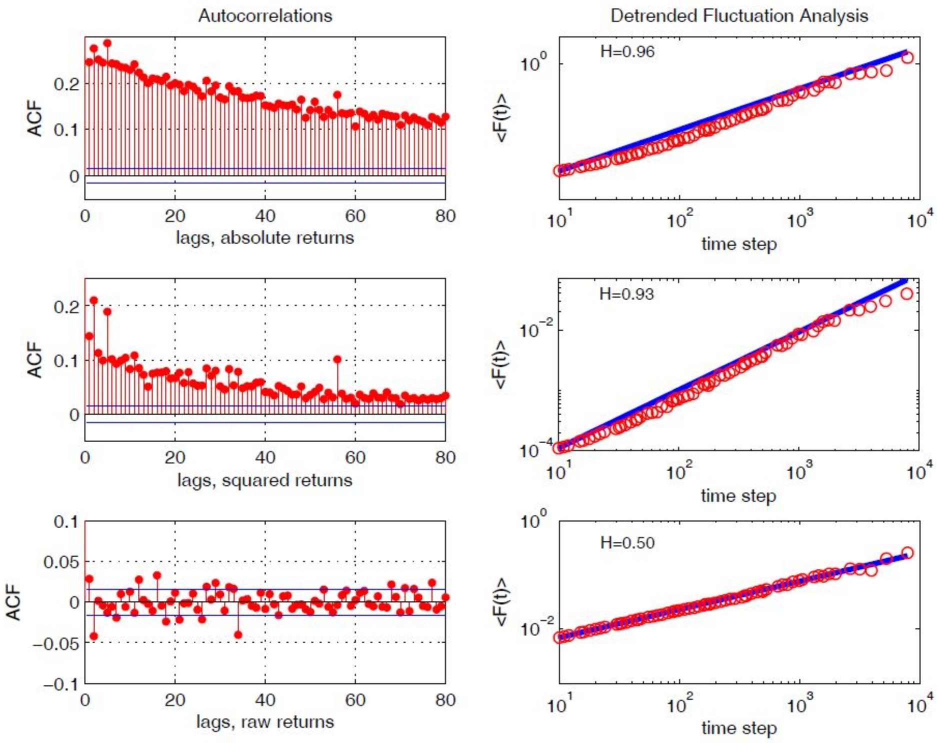

Figure 2.2.: Illustration of the long-term dependence observed in the absolute and squared returns of the Standard & Poor's 500 index (S&P 500) (left upper and central panel). In contrast, raw returns (lower left panel) are almost uncorrelated. The determination of the corresponding Hurst exponent H via the so-called Detrended Fluctuation Analysis (DFA, cf. Chen et al. (2002)) is displayed in the right-hand panels. Note that we obtain the following scaling of the fluctuations (volatility): $< F(t) >\sim t^H$. $H = 0.5$ corresponds to absence of long-term dependency while $H > 0.5$ indicates a hyperbolical decay of the ACF, i.e. long-lasting autoregressive dependency.

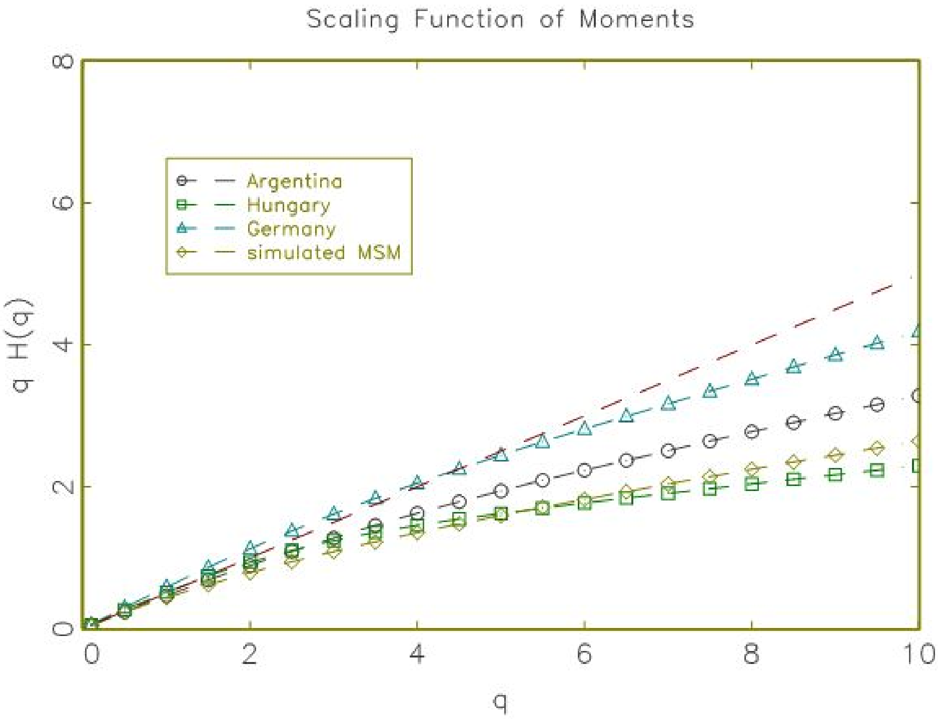

Figure 2.3.: Scaling exponents of moments for three selected financial time series and an example of simulated returns from an MSM process. The empirical samples run from 1998 to 2007, and the simulated series is the one depicted in the lower panel of *Fig.* 2.5. The broken line gives the expected scaling $H(q) = q/2$ under Brownian motion. No fit has been attempted of the simulated to one of the empirical series.

Figure 2.4.: An illustration of the baseline Binomial multifractal cascade. Displayed are the resulting products of multipliers at steps 1, 4, 8 and 12. By moving to higher levels of cascade steps one observes a more and more heterogeneous distribution of the mass over the interval $[0,1]$.

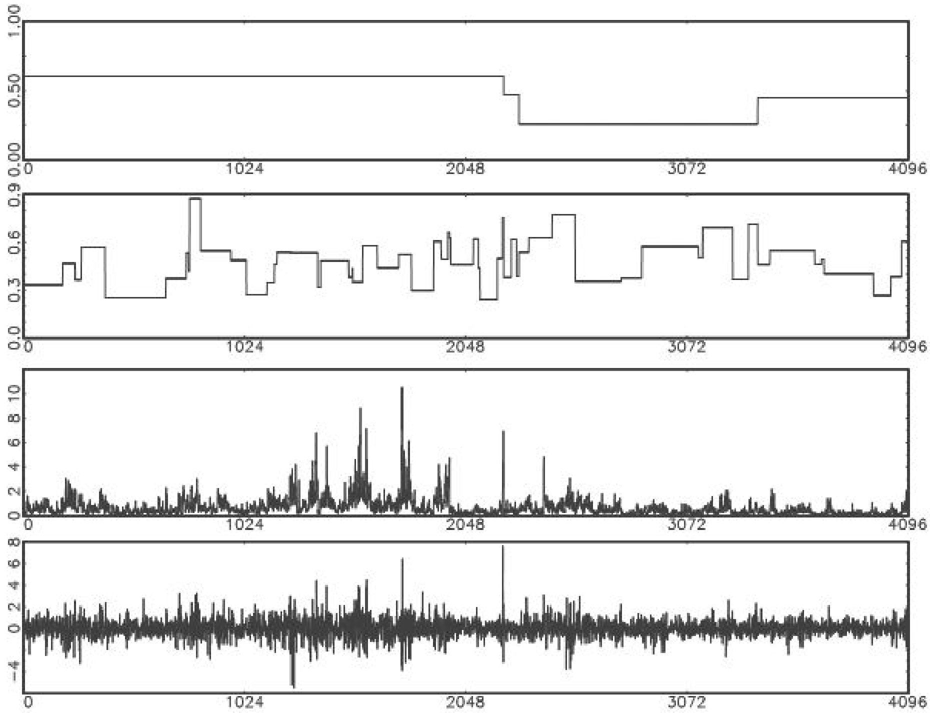

Figure 2.5.: Simulation of a Markov-switching multifractal model (MSM) with Lognormal distribution of the multipliers and $k = 13$ hierarchical levels. The location parameter of the Lognormal distribution has been chosen as $\lambda = 1.05$. The first panel illustrates the development of the second multiplier (with average replacement probability of 2^{-11}), the second panel shows the sixth level, while the third panel shows the product of all 13 multipliers. Returns in the lowest panel are simply obtained by multiplying multifractal local volatility by Normally distributed increments.

3. Financial Duration Models: A Survey

3.1. Introduction

Modeling of high-frequency data attracts a lot of attention in empirical finance because it first allows to understand market microstructure and all issues related to the price adjustment process. Second, it permits researchers to test and corroborate theoretical models (cf. Garman, 1976; Ho and Stoll, 1981; Glosten and Milgrom, 1985; Easley and O'Hara, 1992) developed in the market microstructure literature. The rapid development of information technology (IT) in the early nineties facilitates to store data of all market transactions (trades, quotes, etc...) for every security and triggers the advent of an intensive empirical analysis of high-frequency data. The principal issue related to this kind of data is that they are irregularly spaced. This peculiar feature renders the analysis of the data with existing econometric models such as GARCH[1] (cf. Engle, 1982) unfeasible.

Research by Diamond and Verrecchia (1987), and Easley and O'Hara (1992) presage the information content of the time between transaction events. Empirical investigations of the relationship between security trades and bid-ask quote revisions for stocks traded on the New York Stock Exchange (NYSE) also pointed out that trade durations have information content (cf. Hasbrouck, 1988, 1991). In order to model the irregular spacing of the data and properly gauge their information content Engle and Russell (1998) proposes in their seminal paper an econometric model, termed autoregressive conditional duration (ACD). Engle and Russell (1998) combined the results of transition analysis and the autoregressive structure of GARCH models that achieve a lot of success in modeling time-varying volatility of returns in empirical finance. The autoregressive structure allows the ACD models to capture the information flow that arrives in cluster in the market.

Recent empirical investigations of financial duration data brought new facts to light, e.g., that financial durations exhibit long memory (cf. Jasiak, 1998; Bauwens et al., 2004), asymmetric features (cf. Feng et al., 2004), and fat tailedness (cf. Engle and Russell, 1998; Bauwens and Giot, 2001; Bauwens et al., 2004) which all cannot be captured by the standard ACD model of Engle and Russell (1998). Another feature of the data that has been observed and reported by Engle and Russell (1998) is that high-frequency financial durations show a strong seasonality. This imposes an adjustment of the data before[2] any estimation in order to avoid spurious inferences. So, various

[1] GARCH models are designed for regular spacing data and their use for modeling high frequency data will lead to a loss of primary information.

[2] Some authors execute adjustment and estimation simultaneously (cf. Veredas et al., 2001).

extensions of the standard ACD model have been developed with the aim to better model the above-mentioned features of high-frequency duration data: Fractionally integrated ACD (FIACD) by Jasiak (1998), Log-ACD by Bauwens and Giot (2000), threshold ACD (TACD) by Zhang et al. (2001), stochastic conditional duration (SCD) by Bauwens and Veredas (2004), stochastic volatility duration (SVD) by Ghysels et al. (2004), augmented ACD (AACD) by Fernandes and Grammig (2006) and mixture ACD (MACD) by Hujer and Vuletić (2007).

During the last decade a great number of empirical studies in econophysics has documented and reported the presence of scaling[3] in the intertrade duration distribution of different U.S. stocks (cf. Ivanov et al., 2004; Politis and Scalas, 2008) and Chinese Stocks (cf. Jiang et al., 2008). Sun et al. (2008) computed a Hurst index[4] for 18 Dow Jones index component stocks and found evidence of a fractal structure in intertrade duration data. A Paper by Chen et al. (2013) also provides evidence that the clustering observed in the intertrade durations exhibits self-similarity properties, i.e., it looks similar at different time scales. The presence of self-similarity in the data suggests that the information flow arrives in the markets not only in clusters, but also in cascades. This is in harmony with the conjecture of heterogeneous market participants who act at different time scales, and have limited attention. The effects of a limited investor attention in the financial markets have recently been investigated in detail in the literature (cf. Huberman, 2001; Peng and Xiong, 2006; Barber and Odean, 2008; Corwin and Coughenour, 2008). All these authors find that limited attention significantly influences market participants' decisions, and therefore, trading processes. Corwin and Coughenour (2008), for instance, find that due to limited attention specialists have to allocate effort across securities in their portfolio during busy time periods, and this heavily affects liquidity provision in securities markets. In sum, the trading activity of a stock is not only influenced by information, but also by the attention that is paid to it.

In order to capture long memory observed in the data and to take all these new facts into account Chen et al. (2013) introduced the Markov switching multifractal duration (MSMD) model. While Chen et al. (2013) proposed a mixture of exponential representation for intertrade durations, Baruník et al. (2012) independently introduced a multiplicative error form MSMD model where durations are defined as product of the mean intensity and the innovation. Both models are designed based on the Markov switching multifractal process developed by Calvet and Fisher (2001a, 2004a) that found great acceptance in empirical finance due to its ability to reproduce the scaling law, fat tails and long memory properties (cf. Calvet and Fisher, 2004a; Lux, 2008).

In this chapter we present both classes of models, namely the standard ACD and its subsequent extensions and the MSMD models. We briefly give an overview of the theoretical models that motivate the development of both models in the literature. We mention different diagnostic tests used for testing the adequacy of both types of models and some relevant empirical results gained from their application to financial durations. Until now the only one paper published on a review for ACD models has been accomplished by Pacurar (2008). With these new competitive models, it

[3] The presence of scaling in the intertrade duration distribution has been criticized by Eisler and Kertész (2006).

[4] In addition to its ability to model long memory (cf. (Hurst, 1951, 1955), Hurst index is also a measure for self-similarity scaling.

would be expedient to briefly review both classes of models, their strengths and deficiencies, and outline some future avenues of research.

The rest of the chapter is organized as follows. Section 3.2 reviews the theoretical microstructure models. Section 3.3 presents different ACD models. The MSMD models are illustrated in Section 3.4. Section 3.5 delineates the extant diagnostic tests. Some empirical application results and studies are presented in Section 3.6. Section 3.7 concludes.

3.2. Theoretical Microstructure Models

The theoretical models developed in the microstructure literature try to explain the determinants of the behavior of prices, the way new information is incorporated into prices, and how the market structure can influence the efficiency of the stock market. The underlying idea in these models is that market participants trade with one another for either information or liquidity-motivated reasons. Accordingly, the theoretical microstructure models can be differentiated into two groups: Information- and inventory-based models.

The basic idea in the information-based models was built up by Bagehot (1971), and has then been formalized, developed and extended by Copeland and Galai (1983), Glosten and Milgrom (1985), Kyle (1985), Admati and Pfleiderer (1988), and Easley and O'Hara (1987). Bagehot (1971) considers a scenario in which the market participants are heterogenous on the basis of information they have at hand. He distinguishes between market-makers, informed and uninformed traders. The competitive and risk-neutral market maker who only has public information available does not know whether he is trading with informed or uninformed traders. The adverse selection or asymmetric information problem that the market maker faces here is due to the presence of the informed traders who have superior information (in addition to the public information they also have private information). To safeguard himself from losses he incurs through trading with informed traders in the market, the market-maker has to maintain the spread between ask and bid prices wide enough. Informed traders for their part want to exploit their informational advantage and to maximize their profits. In Kyle (1985)'s model informed traders only know the asset's terminal value. So, their strategy will consist in proportionally trading to the difference between the asset's terminal value and the market clearing price set by the risk neutral market maker. By Glosten and Milgrom (1985) the informed traders will trade intensively whenever they have opportunity to trade in order to immediately benefit from their informational advantages.

The inventory models have in detail been studied by Stoll (1978), Ho and Stoll (1983), and Amihud and Mendelson (1980). The role of inventory control by the market maker has first been discussed in Garman (1976). The basic idea in the inventory control models is that a risk-averse market maker has to adjust the price level if a discrepancy between his actual and desired positions occurs during the trading day. Inventory control leads to price adjustment, and thus, to the existence of the spread between ask and bid prices.

The most contributions mentioned above consider time as an exogenous variable that does not affect the price adjustment process. This role of time has been changed by Easley and O'Hara (1992) who in their model[5] give time a prominent informational role. Easley and O'Hara (1992) argue that uninformed traders trade for liquidity reasons, and not for informational reasons. This is not the case for informed traders whose decisions to trade depend on the quality of new information arriving in the market. Long durations between financial events or low trading intensity would indicate that information arrival in the market is not price relevant and the probability to deal with uninformed traders is high. Accordingly the market maker will decrease the spread between ask and bid prices. This is a proof that time plays an essential role in the price adjustment process. This new role of time heavily influences the market maker's decision to adjust the price. The theoretical model of Easley and O'Hara (1992) opened a new research field in financial econometrics and motivated the development of financial durations models in the empirical financial literature.

3.3. ACD Models

The general form of the ACD models can be expressed as

$$\begin{aligned} x_t &= \Psi_t \xi_t, \\ \Psi_t &= h(x_{t-1}, \ldots, x_{t-p}, \Psi_{t-1}, \ldots, \Psi_{t-q}; \theta), \end{aligned} \tag{3.1}$$

where x_t and Ψ_t are the duration and conditional expected duration at time t, respectively, ξ_t denotes the innovation in the models. The conditional expected duration Ψ_t is a function in p past durations, and q past expected durations. θ is a parameter vector in the models and $h(\cdot)$ can be a linear or a nonlinear function.

3.3.1. The Standard ACD Model

The conditional expected duration of basic ACD(p, q) model proposed by Engle and Russell (1998) has a linear functional form and can be formalized as

$$\begin{aligned} \Psi_t &= \omega + \sum_{j=1}^{p} \beta_j x_{t-j} + \sum_{j=1}^{q} \delta_j \Psi_{t-j}, \\ &= \omega + \beta(L) x_t + \delta(L) \Psi_t, \end{aligned} \tag{3.2}$$

where L denotes the lag operator, $\beta(L) = \beta_1 L + \beta_2 L^2 + \cdots + \beta_p L^p$, and $\delta(L) = \delta_1 L + \delta_2 L^2 + \cdots + \delta_p L^p$ are polynomials, and $\omega > 0$, $\beta_j > 0$, $\delta_j \geq 0$ in order to ensure the positivity of the conditional expectation of the duration, and thus, of the duration. Engle and Russell (1998) construct the model in such a way that the intertemporal correlation in the durations can be condensed in their conditional expectations so that x_t/Ψ_t is independent and identically distributed. They find that

[5] The model by Easley and O'Hara (1992) is an extension of the Glosten and Milgrom's model.

the *eq.* (3.2) nicely captures the clustering of transactions as outguessed by the models of Kyle (1985), Admati and Pfleiderer (1988), and Easley and O'Hara (1992).

Note that for $p = q = 1$, we obtain the ACD(1,1) model which is the most popular and often used model in empirical analysis due to its ability to reproducing the temporal dependence in financial duration data in most cases. The first and second moments of ACD(1,1) model can be found in Engle and Russell (1998) and its autocorrelation function in Bauwens and Giot (2000). The specification of the ACD model imposes an exponential decrease of the autocorrelation function which does not match the hyperbolical decay of the empirical autocorrelation function.

In their seminal paper Engle and Russell used two distributional assumptions, namely the standard exponential and Weibull[6] distributions for the innovation in the model. The exponential distribution is an asset for the estimation of the model, because it provides consistent quasi-maximum likelihood estimators (cf. Drost and Werker, 2004), but inadequate for the modeling of the data due to the fact that it leads to constant conditional hazard function, and thus, not in conformity with the empirical conditional hazard function. This is overcome by using a Weibull distribution whose conditional hazard function is increasing when the shape parameter is larger than 1 and decreasing when the shape parameter is less than 1. However, this flexibility obtained by using a Weibull distribution is not enough for a good modeling of financial duration data.

In order to obtain an appropriate conditional hazard function, flexible distributions for the innovation have been proposed in the literature. Grammig and Maurer (2000) used a Burr distribution for the innovation. This distribution includes Weibull, log-logistic, and exponential as special cases. Note that the Burr distribution requires some parametric restrictions in order to ensure that the first, second moments and higher moments exist. These restrictions sometimes lead to poor results when the Burr ACD (BACD) model is applied for modeling high unconditional moments of financial durations (cf. Bauwens et al., 2008). Another more attractive and often used distribution in all recent empirical studies is the generalized gamma which encompasses gamma, Weibull, and exponential as particular cases (cf. Lunde, 1999). The generalized gamma distribution offers a flexible hazard function[7] which is increasing for small durations and decreasing for long durations. The generalized F distribution[8] which encompasses the Burr-type 12, the Lomax, the Fish, and the folded t distributions as particular cases has been proposed in Hautsch (2001) to analyze excess volume durations. The Birnbaum-Saunders distribution for financial durations has recently been proposed by Bhatti (2010).

For forecasting purposes *eq.* (3.2) is not convenient and one has to rewrite the ACD(p,q) process as an ARMA($\max(p, q), q$) process for durations. This can be obtained as follows.

Let $e_t = x_t - \Psi_t$ be the innovation associated with the duration process or the martingale differ-

[6] The Weibull distribution reduces to exponential one if the shape parameter is set to one.
[7] The hazard function of the generalized gamma distribution can be found in Glaser (1980).
[8] The hazard function of the generalized F has been studied in McDonald and Richards (1987).

ence. By inserting e_t in the *eq.* (3.2), and rearranging terms we obtain

$$x_t = \omega + \sum_{j=1}^{\max(p,q)} (\beta_j + \delta_j)x_{t-j} - \sum_{j=1}^{q} \delta_j e_{t-j} + e_t, \qquad (3.3)$$

or equivalently

$$[1 - \beta(L) - \delta(L)]x_t = \omega + [1 - \delta(L)]e_t. \qquad (3.4)$$

A well-defined duration process in *eq.* (3.3) imposes that the following condition muss be satisfied: $\sum_{j=1}^{p} \beta_j + \sum_{j=1}^{q} \delta_j < 1$. As shown by Nelson and Cao (1992) for GARCH processes it is clear that the stationarity and invertibility conditions for duration processes in *eq.* (3.4) require that the roots of $[1 - \beta(L) - \delta(L)]$ and $[1 - \delta(L)]$, respectively, lie outside the unit circle.

A natural way to extend the ACD model is to include some exogenous economic variables such as the bid-ask spread, the unexpected trading volume[9] in the conditional duration equation, cf. *eq.* (3.2). This has been done in many papers to improve the forecast performance of the model. However, some authors find this extension of the ACD model to be incomplete, because it does not care about the information revealed by the price process that is primary for forecasts. Engle (2000) proposed an ACD-GARCH model that is a combination of a marginal ACD model for durations and a GARCH model for returns. The ACD-GARCH model has also been studied by Grammig and Wellner (2002). Models developed in the literature with the same objective can be found in Meddahi et al. (2006), Hafner (2005), Darolles et al. (2000), and Russell and Engle (2005).

Note that the estimation of the standard ACD model with different distributional assumptions for innovations can be easily performed by the maximum likelihood approach. Except for exponential distribution, other distributional assumptions do not provide asymptotically consistent estimators under model misspecification by quasi maximum likelihood estimation (QMLE), so that their inference depends on the quality of the model fit. By exploiting the results of Lee and Hansen (1994) Engle and Russell (1998) furnished asymptotic properties of the ACD(1,1).

3.3.2. The Logarithmic ACD (Log-ACD) Model

In addition to the fact that the ACD model is well suited to the analysis of the time elapsed between consecutive transactions, we are also interested in testing some market microstructure hypotheses. This requires to add other economic variables than lagged durations in the standard ACD model, and therefore, more restrictions on the parameters to ensure positivity of the conditional expected duration. To avoid the non-negativity constraints on the parameters, Bauwens and Giot (2000) proposed the logarithmic version of ACD model. The Log-ACD model imposes a nonlinear relationship between the conditional expected duration and their lagged. The generalized equation

[9] The unexpected trading volume is defined as the deviation of the actual trading volume from the time-of-the-day adjusted volume.

form for the conditional expected duration in the Log-ACD model is given by

$$\psi_t = \omega + \sum_{j=1}^{p} \beta_j f(\xi_{t-j}) + \sum_{j=1}^{q} \delta_j \psi_{t-j}, \tag{3.5}$$

where different functional forms for $f(\xi_{t-j})$ can be used. ψ_t is the logarithm of Ψ_t. Bauwens and Giot (2000) proposed two choices:

1. The first one is $f(\xi_{t-j}) = \ln(\xi_{t-j}) = \ln\left(x_{t-j}/\Psi_{t-j}\right)$ and *eq.* (3.5) becomes

$$\psi_t = \omega + \sum_{j=1}^{p} \beta_j \ln(x_{t-j}) + \sum_{j=1}^{q} \left(\delta_j - \beta_j\right) \psi_{t-j}, \tag{3.6}$$

   and this model specification is called Log-ACD$_1$ in the original paper of Bauwens and Giot (2000). Note that for covariance stationarity the following condition has to be satisfied $\left|\sum_{j=1}^{p} \beta + \sum_{j=1}^{q} \delta\right| < 1$.

2. The second one is $f(\xi_{t-j}) = \xi_{t-j} = x_{t-j}/\Psi_{t-j}$ and *eq.* (3.5) becomes

$$\psi_t = \omega + \sum_{j=1}^{p} \beta_j \left[x_{t-j}/\exp(\psi_{t-j})\right] + \sum_{j=1}^{q} \delta_j \psi_{t-j}. \tag{3.7}$$

   • This model specification is termed Log-ACD$_2$, the necessary condition for covariance stationarity is $\left|\sum_{j=1}^{q} \delta\right| < 1$. The Log-ACD$_2$ is preferred in practice and in empirical analysis due to its ability to better fit financial duration data than the Log-ACD$_1$. All the above-mentioned distributional assumptions for the innovation in the standard ACD model can also be used in the Log-ACD models. The moments of Log-ACD models with any distribution with positive support are provided in Bauwens et al. (2008) and statistical properties of the Log-ACD models with Burr and generalized F distributions for innovations have been investigated in Karanasos (2008). Recently, Allen et al. (2008) proposed the Lognormal distribution for innovations in the Log-ACD models and proved the consistency and asymptotic normality of quasi-maximum likelihood estimators that are essential for a valid inference and diagnostic tests. The estimation of the Log-ACD models can also be performed via exact maximum likelihood method.

Bauwens and Giot (2003) argue that the sign of changes in ask and bid prices would affect the duration for the next price movement, and therefore, have to be taken into account when modeling duration. To incorporate this primary information in the Log-ACD model, they combined a two-state transition model with a Log-ACD model to obtain an asymmetric Log-ACD model that can jointly model the duration process and the information on the direction of price movement. Following Bauwens and Giot (2003)'s idea, Allen et al. (2008) developed two new asymmetric Log-ACD models. The first model is related to the model of Glosten et al. (1993) and is obtained by introducing an indicator function I_t in the second term on the right-hand side of *eq.* (3.6) that takes 0 if the significant change in the mid-price is positive and 1 if it is negative. The second one is the Log-ACD model that permits including some exogenous variables in the *eq.* (3.6).

3.3.3. The Augmented ACD (AACD) Model

Fernandes and Grammig (2006) generalized the standard ACD model using a Box-Cox transformation with parameter $\lambda \geq 0$. The motivation is to find a class of models that will avoid an over-prediction after either very long or very short durations as stressed in Engle and Russell (1998). The ACD model after the transformation can be expressed as

$$\frac{\Psi_t^\lambda - 1}{\lambda} = \omega^* + \beta^* \Psi_{t-1}^\lambda [\xi_{t-1} - b - c(\xi_{t-1} - b)]^\nu + \delta \frac{\Psi_{t-1}^\lambda - 1}{\lambda}. \tag{3.8}$$

The augmented ACD (AACD) model is obtained by rewriting *eq.* (3.8) as

$$\Psi_t^\lambda = \omega + \beta \Psi_{t-1}^\lambda [|\xi_{t-1} - b| - c(\xi_{t-1} - b)]^\nu + \delta \Psi_{t-1}^\lambda, \tag{3.9}$$

where $\omega = \lambda\omega^* - \delta + 1$ and $\beta = \lambda\beta^*$. Note that the Box-Cox transformation is concave if $\lambda \leq 1$ and convex if $\lambda \geq 1$. The shocks impact curve[10] $g(\xi_t) = [|\xi_{t-1} - b| - c(\xi_{t-1} - b)]^\nu$ allows the conditional duration process to capture asymmetric effects through the shift and rotation parameters b and c, respectively. The asymmetric responses implied by the shocks impact curve are identified with the shift parameter b. The parameter c reveals information on the type of rotation. $c < 0$ indicates a clockwise rotation and $c > 0$ a counterclockwise. The shape parameter ν determines whether the shocks impact curve is concave ($\nu \leq 1$) or convex ($\nu \geq 1$). The AACD model includes various ACD models such the Box-Cox ACD($\lambda \to 0, b = c = 0$) model proposed by Dufour and Engle (2000a), the standard ACD ($\lambda = \nu = 1$ and $b = c = 0$) model, Log-ACD$_1$ ($\lambda \to 0, \nu = 1$ and $b = c = 0$) model and Log-ACD$_2$ ($\lambda, \nu \to 0$ and $b = c = 0$) model. Sufficient conditions that guarantee finite higher-order moments for conditional duration processes, strict stationarity, geometric ergodicity and β-missing property with exponential decay can be found in Fernandes and Grammig (2006). The parameters of the AACD models can easily be estimated using the maximum likelihood approach.

3.3.4. Long Memory ACD Models

The basic ACD model belongs to the class of ARMA-type models, and thus, can just account for short serial dependence in conditional duration. However, empirical intertrade data exhibits long memory features, i.e. the autocorrelation functions of empirical data display a slow, hyperbolic rate of decay. Inspired by fractionally integrated GARCH (FIGARCH model) proposed by Baillie et al. (1996), Jasiak (1998) developed a fractionally integrated version of ACD, termed FIACD

[10] cf. Fernandes and Grammig (2006) for illustration of the shocks impact curve for different parameter values for b, c, and ν.

model that can reproduce the long memory properties observed in the empirical data. By introducing the fractional differencing operator $(1-L)^d$, with $d \in [0,1]$ in the ARMA representation of the ACD(p,q) in *eq.* (3.4) one obtains the FIACD(p,d,q) model that can be expressed as

$$\begin{aligned}[1-\delta(L)]\Psi_t &= \omega^* + [1-\delta(L) - [1-\beta(L)-\delta(L)](1-L)^d]x_t, \\ &= \omega^* + A(L)x_t,\end{aligned} \tag{3.10}$$

where $A(L) = a_1 L + a_2 L^2 + \ldots$ and $\delta(L) = \delta_1 L + \delta_2 L^2 + \ldots$ are polynomials with $a_k \geq 0$ and $\delta_k \geq 0$, for $k = 1, 2\ldots$, and $\omega^* > 0$ in order to guarantee the positivity of the conditional duration. The fractional differencing operator $(1-L)^d$ is given by

$$(1-L)^d = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} L^j, \tag{3.11}$$

where $\Gamma(\cdot)$ is the gamma function. For $d \in [0,1]$ the FIACD model is strictly stationary and ergodic (cf. Jasiak, 1998). The FIACD model can display long memory properties of financial duration data. If $0 < d < 0.5$, the FIACD process is a long memory process and the autocorrelation function decays hyperbolically. For $d = 0$, the FIACD model becomes an ACD model, and for $d = 1$ the FIACD reduces to an integrated ACD (IACD) model. However, the principal issue related to the FIACD process is that it is not covariance stationary, in other words, it does not possess finite first and second unconditional moments. Consequently, the FIACD model does not have long memory in the usual sense. Furthermore, the asymptotic properties of the model estimator are until now not well-documented. Recently, alternative specifications have been made available by Koulikov (2003) and Karanasos (2004). Both proved that their processes possess finite first and second moments under certain conditions that are not mentioned here. We refer the reader to Koulikov (2003) and Karanasos (2004). Jasiak (1998) suggested that the asymptotic properties of QML estimators of the FIACD (p,d,q) model with $d \in (0,1)$ can be obtained by extending the results found by Lee and Hansen (1994) for the IGARCH(1,1) process with a Gaussian misspecified pdf.

Another model for capturing the long memory properties of financial durations has been developed by Deo et al. (2010). They proposed a long memory stochastic duration (LMSD) model that is an extension of the stochastic volatility duration (SCD) model developed by Bauwens and Veredas (2004), cf. 3.3.6. The model can be formalized as

$$\begin{aligned}\Psi_t &= \exp(\psi_t) \\ \psi_t &= \omega + (1-L)^d e_t,\end{aligned} \tag{3.12}$$

where $\omega \in \mathbb{R}$, e_t is a zero-mean Gaussian stationary short memory series, L is the lag operator, and $d \in [0, 0.5]$. ξ_t in *eq.* (3.1) are *i.i.d.*, independent of e_t. Note that here it is difficult to implement the MLE due to the fact that the variable ψ_t in the model is latent, i.e. is unobservable and has to be integrated out. To circumvent this difficulty, Deo et al. (2010) make use of the Whittle's approximation to implement a quasi-maximum likelihood estimator for the parameters.

3.3.5. Regime-Switching ACD Models

Threshold ACD Model:

It is well-documented that there is a nonlinear dependence between conditional expectations of durations and past information set available (cf. Engle and Russell, 1998; Zhang et al., 2001; Meitz and Teräsvirta, 2006). In order to model different dynamics observed in fast and slow trading period in the market Zhang, Russell, and Tsay proposed the threshold ACD (TACD) model that is wedded to the threshold autoregressive (TAR) model and the more general threshold autoregressive moving average (TARMA) model. Defining $R_i = [r_{i-1}, r_i)$, $i = 1, 2, \ldots, I$, for a positive integer I, where $-\infty = r_0 < r_1 < \cdots < r_I = \infty$ are the threshold values. A I-regime threshold ACD(p,q) model can be formalized as

$$\Psi_t = \omega^{(i)} + \sum_{j=1}^{p} \beta_j^{(i)} x_{t-j} + \sum_{j=1}^{q} \delta_j^{(i)} \Psi_{t-j}, \quad \text{if } l_{t-d} \in R_i. \tag{3.13}$$

l_{t-d} is the threshold variable that determines the regime boundaries. The delay parameter d is a positive integer. Here it is important to know that the parameter of the innovation distribution in the TACD model varies across I-regimes, allowing for different shapes for the hazard function in different trading regimes. Zhang et al. (2001) intensively study the TACD(1,1) model and provide conditions for geometric ergodicity and existence of moments which can be easily generalized for higher order models.

Inspired by the smooth transition GARCH models (cf. Lee and Degennaro, 2000; Lundbergh and Teräsvirta, 2002) Meitz and Teräsvirta (2006) introduced the smooth transition ACD (STACD) model. The model is closely related to the TACD model and can help avoiding the overprediction of the expected durations often observed by the linear ACD model after either very long or very short durations. Meitz and Teräsvirta (2006) also proposed a time-varying ACD (TVACD) model that in contrast to the standard ADC model allows for changing parameters over the sample period. The idea to incorporate time-varying parameters in the standard ACD model seems to be more realistic due to the fact that the economic environment is often affected by negative or positive shocks, and thus, can also affect the structure of the trading process. The TVACD model represents the ideal tool for testing the constancy of parameters.

Markov Switching ACD Model:

With the objective to find a model that can capture a broad range of different dynamics observed in financial duration data, Hujer et al. (2002) proposed a Markov switching ACD (MSACD) model. The idea is to introduce in the conditional mean function an unobserved random regime variable s_t whose evolution over time follows a Markov chain process. The MSACD model can be formalized as

$$\Psi_t = \sum_{i=1}^{k} Pr(s_t = i | \Im_{t-1}; \theta) \Psi_t^{(i)}, \tag{3.14}$$

where $\Psi_t^i = \mathbb{E}(x_t|s_t = i, \Im_{t-1}; \theta)$ is the regime specific conditional mean and may have a linear or nonlinear autoregressive specification according to the dynamics of a standard ACD model. $Pr(s_t = i|\Im_{t-1}; \theta)$ represents the probability that s_t is in state i given the information set $\Im_{t-1}$ available at time $t-1$. How one has to specify the conditional mean function $\Psi_t^{(i)}$ and the stationarity conditions are discussed in detail in Hujer et al. (2002). In their paper Hujer et al. (2002) used Burr family distributions for each regime specific distribution. De Luca and Zuccolotto (2006) studied the regime switching Pareto ACD models and found that this specification permits capturing better the dynamic of the duration process.

Discrete Mixture ACD Model:

Recently, Hujer and Vuletić (2007) proposed a discrete mixture ACD (DMACD) model that is designed by introducing a discrete-valued latent regime variable in the ACD process. By doing so, Hujer and Vuletić (2007) transform the observable duration process to a latent stochastic duration process. This new representation for durations encompasses various ACD models such as the MSACD and the standard ACD models. The DMACD model is a weak form of the ACD model, because innovations in this modeling framework are serially independent with known discrete mixture distribution that can be specified as

$$h(\xi_t; \theta) = \sum_{i=1}^{N} \alpha_i h(\xi_t|s_t = i; \theta), \tag{3.15}$$

where $\alpha_i \in [0, 1]$ is the probability for prevailing state i, s_t is latent regime variable with countable support $\mathbb{J} = \{i|1 \leq i \leq I\}$, $I \in \mathbb{N}$.

The idea of mixture distributions for modeling financial durations is not new and mixture ACD models for durations have been proposed by De Luca and Zuccolotto (2003) and De Luca and Gallo (2004) before. They found that mixture distributions are convenient to model the presence of heterogeneous traders in the market. Hujer and Vuletić (2007) gained insights from the empirical application that by assuming constant regime probabilities all along the trading time one obtains a static MACD (SMACD) model that parsimoniously models the high persistence of intraday durations. They recommended to use this static representation to overcome the distributional problem of the duration in De Luca and Gallo (2004). Though the mixture models can reproduce high-frequency duration data, they exhibit poor forecasting performance. Hujer and Vuletić (2007) argue that this is due to the fact that the mixture models cannot properly classify future regimes. However, these models help to better understand the trade behavior of the market participants.

The estimation of regime switching can be performed by the maximum likelihood method. The maximum likelihood method is simple, however not so appropriate for switching models due to the fact that their likelihood functions may have more than one local maximum and these may be located in boundary regions of the parameter space (cf. Hujer et al., 2002). Hujer et al. (2002) proposed to use the Expectation-Maximization (EM) algorithm developed by Dempster et al. (1977) that in contrast to the standard algorithms can solve the problem of multiple local maximums.

3.3.6. Stochastic Conditional Duration Model

Bauwens and Veredas (2004) proposed a stochastic conditional duration (SCD) model whose construction is based on the assumption that durations are generated by a dynamic stochastic latent variable. The latent conditional duration is formalized as

$$\begin{aligned} \Psi_t &= \exp(\psi_t), \\ \psi_t &= \omega + \beta\psi_{t-1} + e_t, \end{aligned} \tag{3.16}$$

where $|\beta| < 1$ and e_t denotes the innovation in the model.

This specification for the conditional duration permits the SCD model to provide a flexible structure for the dynamics of the duration process, but however is not adequate for capturing the asymmetric behavior in duration data. To tackle this shortcoming, Feng et al. (2004) extended the SCD model by introducing an intertemporal error term in the latent conditional duration process, allowing the SCD model more flexibility. Feng et al. (2004) formalized the asymmetric SCD model as

$$\begin{aligned} \ln(x_t) &= \mu + \Psi_t + \epsilon_t, \\ \Psi_t &= \beta\Psi_{t-1} + \varrho\epsilon_{t-1} + v_t, \end{aligned} \tag{3.17}$$

where $|\beta| < 1$, ϵ_t and v_t are *i.i.d.* innovations and are mutually independent. They assumed that v_t follows Gaussian $N(0, \sigma_v^2)$ and consider three distributions for ϵ_t, namely log-Weibull, log-gamma and log standard exponential.

The estimation of the SCD model is difficult, because the likelihood function involves a multi-dimensional integral due to the presence of the unobservable variable Ψ_t. Bauwens and Veredas (2004) proposed an attractive QML method based on the Gaussianity assumption of the log of the innovations and the use of the Kalman filter in a linear space state model (cf. Harvey et al., 1994). The shortcomings of this estimation method are that it does not provide efficient estimates of the parameters. However, the estimators are asymptotically consistent and the estimation is time parsimonious. Recently, Bauwens and Galli (2009) applied the efficient importance sampling methodologies developed by Liesenfeld and Richard (2003) to estimate the SCD model and obtained a significant gain in forecasting exercises. Furthermore, the empirical characteristic function and the GMM methods can also be used to perform the estimation of the SCD model (cf. Knight and Ning, 2008). Feng et al. (2004) used the Monte Carlo maximum-likelihood (MCML) approach proposed by Durbin and Koopman (1997) to estimate the asymmetric SCD model.

3.3.7. Stochastic Volatility Duration Models

Ghysels et al. (2004) developed a stochastic volatility duration (SVD) model with the aim to capture different patterns of temporal dependence observed in the conditional mean and variance of financial durations. In the SVD model durations are formalized as

$$x_t = \frac{U_t}{aV_t}, \tag{3.18}$$

where U_t, V_t are independent, U_t follows the standard exponential distribution ($U_t \sim \text{Exp}(1)$, or gamma$(1, 1)$), and V_t follows a gamma distribution with positive parameter b ($V_t \sim \text{gamma}(b, b)$). This specification for durations can be remodeled through suitable transformations in a two factors model where the factors are Gaussian:

$$x_t = \frac{G(1, \Phi(F_{1t}))}{aG(b, \Phi(F_{2t}))} = \frac{H(1, F_{1t})}{aH(b, F_{2t})}, \tag{3.19}$$

where F_{1t}, F_{2t} are *i.i.d.* standard Normal variables, $G(b, .)$ is the quantile function of the gamma(b, b) distribution, and Φ is the cdf of the standard Normal. Ghysels et al. (2004) generalized the model to a class of SVD models by utilizing a bivariate vector autoregressive (VAR) time series representation for the process $F_t = (F_{1t}, F_{2t})'$, where the marginal distribution of F_t is constrained to be $N(0, I)$ in order to guarantee that the marginal distribution of x_t is Pareto.

The model can be expressed in its generalized form by

$$F_t = \sum_{i}^{I} \Omega_i F_{t-i} + \varepsilon_t, \tag{3.20}$$

where Ω_i is a matrix of autoregressive VAR parameters, and ε_t is a vector of Gaussian white noise random variables with variance-covariance matrix $\Sigma(\Omega)$ such that $\mathbb{V}ar(F_t) = Id$.

Since its introduction, the SVD model did not achieve success in empirical application due to the fact that its estimation causes enormous problems. Indeed, the likelihood function involves a multidimensional integral due to the presence of latent factors. It is clear that a simulated maximum likelihood method can be used to perform the estimation (cf. Shephard and Pitt, 1997). However, this estimation approach is computationally intensive and time consuming. To make the estimation easier without any additional assumptions on the model parameters, Ghysels et al. (2004) proposed estimation procedures which consist in first estimating the parameters a and b using QML method, and then, second making use of the method of simulated moment (cf. McFadden, 1989; Gouriéroux and Monfort, 1996, chap. 2) after replacing the parameters a and b by their respective estimates $\hat{a}$ and $\hat{b}$. In addition to the estimation difficulties, the SVD model is found by Bauwens et al. (2004) to exhibit poor forecast performance compared to the standard ACD or Log-ACD model.

3.4. Markov Switching Multifractal Duration Models

Here we present a new class of financial duration models. This class includes two duration models recently developed in the literature. Both models are independently proposed by Chen et al. (2013)

and Baruník et al. (2012) who used the Markov switching multifractal process proposed in Calvet and Fisher (2001a, 2004a) as basic ingredients.

3.4.1. Chen/Diebold/Schorfheide Model

Chen et al. (2013) proposed a mixture-of-exponentials representation for intertrade durations. Their model can be formalized as

$$x_t = \frac{\xi_t}{\lambda(M_t)}, \tag{3.21}$$

with

$$\lambda(M_t) = \bar{\lambda} \prod_{i=1}^{k} M_t^{(i)}, \tag{3.22}$$

where x_t represents the time elapsed between two consecutive financial events, ξ_t is *i.i.d.* standard exponential distributed, $\bar{\lambda}$ is the unconditional mean intensity ($\bar{\lambda} > 0$) that controls the overall intensity level, $k \in \mathbb{N}$ and $M_t = \left(M_t^{(1)}, M_t^{(2)}, \ldots, M_t^{(k)}\right)$ is the trading intensity state vector at time t. The latent intensity components $M_t^{(1)}, M_t^{(2)}, \ldots, M_t^{(k)}$ are drawn from a Binomial distribution taking values m_0 and $2 - m_0$; $m_0 \in (0, 2]$, with equal probability so that $\mathbb{E}\left[M_t^i\right] = 1$ is guaranteed. Each intensity component, $M_t^{(i)}$, is renewed at time t with probability γ_i depending on its rank within the hierarchy of multipliers and remains unchanged with probability $1 - \gamma_i$. The transition probabilities are specified as

$$\gamma_i = 1 - (1 - \gamma_1)^{(b^{i-1})}, \quad i = 1, \ldots, k, \tag{3.23}$$

with parameters $\gamma_1 \in (0, 1)$ and $b \in (1, \infty)$. The transition matrix related to each intensity component has the following form:

$$P_i = \begin{pmatrix} 1 - \frac{1}{2}\gamma_i & \frac{1}{2}\gamma_i \\ \frac{1}{2}\gamma_i & 1 - \frac{1}{2}\gamma_i \end{pmatrix}. \tag{3.24}$$

From *eq.* (3.23) it is clear that the renewal probabilities grow approximately at a geometric rate b (cf. Calvet and Fisher, 2004a), creating intensity components in descending frequency order, i.e., from low-frequency to high frequency components. In sum, the value of γ_i determines the average lifetime or persistence of a M_t^i shock. This means that the smaller γ_i is, the longer average lifetime the M_t^i shock will have, and conversely. Note that the duration process is stationary, ergodic because first the processes M_t^i are strictly stationary and ergodic (due to the transition matrix in *eq.* (3.24)), and second they are also independent across k and independent of ξ_t.

Binomial distribution for multipliers implies a finite number of states of the hidden Markov process, and this permits the estimation of the model using the exact maximum likelihood via Bayesian updating (cf. Calvet and Fisher, 2004a). The issues related to the maximum likelihood approach is that it becomes unfeasible if the multipliers have a continuous probability distribution[11] or if the number of the multiplier components is greater or equal to ten ($k \geq 10$). As stressed

[11] A continuous probability distribution for multipliers implies an infinite state space of the hidden Markov chain.

in Chen et al. (2013) the MSMD model is afflicted with the following identification problems: the parameter b is non-identifiable if $\gamma_i = 1$, but becomes weakly identifiable if γ_i approaches its upper bound that is set to 0.999. In the next subsection we present an alternative MSMD model introduced by Baruník et al. (2012). Other identification problems related to $b = 1$ and $m_0 = 1$ do not have any relevance for the empirical application.

3.4.2. Baruník/Shenai/Žikeš Model

Baruník et al. (2012) proposed a multiplicative error form where the adjusted duration, x_t, is the product of the Markov switching multifractal process of Calvet and Fisher and an *i.i.d.* unit-mean innovation. The model is defined as

$$x_t = \lambda(M_t)\varepsilon_t, \tag{3.25}$$

where $\lambda(M_t)$ is defined as in *eq.* (3.22). Any positive distribution with positive support can be assumed for the unit-mean innovation ξ_t in the model.

In the paper by Baruník et al. (2012) the latent intensity components or multipliers are drawn from a Binomial distribution taking m_0 and $2 - m_0$; $m_0 \in (1, 2)$, with equal probability in order to guarantee the unit-mean of $M_t^{(i)}$. In addition to Binomial distribution Baruník et al. (2012) also considered continuous distribution for the multipliers, namely Lognormal distribution (cf. Lux, 2008). In this case multipliers are determined by the random draws from a Lognormal distribution with parameter μ, i.e.

$$M_t^i \sim LN(-\mu, 2\mu). \tag{3.26}$$

Additionally to the exact maximum likelihood method proposed by Calvet and Fisher (2004a), which can only be used for the estimation of the model when the multipliers follow discrete distributions, e.g., the Binomial distribution, they also propose the Whittle estimator for the parameters in the MSMD model. The advantage of the latter is that it is applicable to the models with a discrete or continuous distribution for multipliers. The Whittle estimator is obtained by minimizing the negative Whittle log-likelihood function. For more details of the estimation procedures, we refer the reader to Baruník et al. (2012). For forecasting purposes optimal (cf. Calvet and Fisher, 2004a) or linear (cf. Lux, 2008) forecasting methodologies can be performed.

3.5. Diagnostic Tests

One important question when modeling high-frequency financial durations remains how to test the adequacy of models used. Engle and Russell (1998) proposed in their seminal paper to examine the estimated residuals ($\hat{\xi}_t = x_t/\hat{\Psi}_t$) and the squared estimated residuals ($\hat{\xi}_t^2$). The idea is that a correct specification of the model would imply that ξ_t are *i.i.d.* which means that the model can capture the intertemporal dependence. They applied the well-known Ljung-Box test to $\hat{\xi}_t$ and $\hat{\xi}_t^2$ to check the independency hypothesis. This common way of examining the dynamical properties of the estimated residuals has often been used in a great number of papers. The issue related to the Ljung-Box Q-statistic is that its asymptotical behavior is not documented for the ACD models. This doubtfulness about the asymptotic distribution of the test statistic has been reinforced by the work of Li and Mak (1994), who find that when applying the Ljung-Box test to the estimated standardized residuals in GARCH framework the test statistics do not have the usual asymptotic chi-squared (χ^2) distribution under the null hypothesis. Li and Yu (2003) developed a Portmanteau test for the goodness-of-fit under the assumption that the innovations are exponential distributed[12]. Although the Portmanteau test of Li and Yu (2003) is adapted to the ACD framework, its application to assess the goodness-of-fit of ACD models remains scarce. Perhaps, this is due to the fact that exponential or Weibull distribution for durations is not in conformity with the empirical distribution function of durations. Additionally to Ljung-Box or Portmanteau test some authors prefer to visualize the autocorrelation function of estimated residuals (cf. Jasiak, 1998; Bauwens and Giot, 2000) or compare the marginal density of durations obtained from the model with the empirical marginal density of the observed durations (cf. Bauwens and Veredas, 2004; Ghysels et al., 2004). Chen et al. (2013) used the information matrix (IM) test developed by White (1982) to test *i.i.d.* data and later extended by White (1994) to time series models to assess whether the MSMD model is well specified. In other papers QQ-plots (cf. De Luca and Gallo, 2004) or Bartlett identity tests (cf. Prigent et al., 2001) have been used for assessing the adequacy of the ACD models. Duchesne and Pacurar (2008) proposed a class of tests for testing the adequacy of ACD models. The test procedures are based on Hong (1996, 1997)'s approach that consists in utilizing the kernel-based spectral density estimator of the standardized residuals.

Some authors prefer to concentrate their efforts on verifying the distributional assumptions for innovations in the ACD models. Engle and Russell (1998) developed an overdispersion test that can help checking whether the distributional assumptions (exponential, Weibull) for innovations are suited. Fernandes and Grammig (2005) found that the overdispersion test exhibits poor performance. Dufour and Engle (2000a) developed a new Lagrange multiplier test that can be used to evaluate the accuracy of density forecasts. The test helps to assess whether the ACD models are well specified. Bauwens et al. (2004) employed density forecast evaluation methods of Diebold et al. (1998) to assess the specification of duration models. Their methodology is simple and consists in testing the null hypothesis that the sequence of probability transforms of the one-step-ahead forecasts of the conditional densities of durations are $i.i.d. U(0, 1)$ distribution. The null hypothesis cannot be rejected if the one-step-ahead forecasts of the conditional densities of the durations are in conformity with the true densities of durations. Allen et al. (2009) combine density and interval forecast methodologies to test the adequacy of the ACD models. Baruník et al. (2012) check the goodness-of-fit of their model using the specification test proposed by Chen and Deo (2004). The idea of Chen and Deo (2004) is that under the null hypothesis of correct model specification the difference between the estimated model's spectral density and the smoothed periodogram of the data has to be zero. Fernandes and Grammig (2005) also proposed new procedures to test

[12] In the case the innovations are Weibull distributed one just needs to make a change of variable to obtain an appropriate portmanteau test for goodness-of-fit (cf. Li and Yu, 2003).

the ACD models' specification. This consists in gauging the distance between the parametric and nonparametric estimates of the density and hazard rate functions of the residuals obtained from QML estimation.

For other authors the functional form of the conditional mean duration can also be a source of misspecification. So, other test procedures have been developed for testing the specification of the conditional mean duration function. The crucial condition for obtaining QML estimators in the ACD models is the correct specification of the conditional mean function. This basic intuition has been exploited to develop different test procedures. Meitz and Teräsvirta (2006) proposed a Lagrange Multiplier (LM) test for testing the functional form of the conditional mean duration. They also proposed a more general LM test that can be used for testing the adequacy of different forms (linear, nonlinear, or higher-order ACD models) of specification of the conditional mean duration function, and the constancy of the parameters over the sample period. In the same line, Hautsch (2006) developed LM tests against sign bias alternative and nonlinearities in the news impact function (cf. 3.3.3). He also used various conditional moment (CM) tests and integrated conditional moment (ICM) tests which are obtained as a conversion of conditional moment test in to a chi-square test (cf. Bierens, 1990). With the objective to generalize the LM tests that possess optimal power against local alternative Hautsch (2008) proposed a conditional moment test which is the robust form of Newey (1985)'s conditional moment test adapted to the ACD framework. Hautsch (2012) recommended to use CM tests as complements to LM tests in real application. Chen and Hsieh (2010) also proposed generalized moment tests. Their test allows for testing the conditional mean function, the *i.i.d.*ness, and the distributional misspecification.

3.6. Some Empirical Results

The basic ACD model and its extensions have successfully been applied for testing market microstructure hypotheses and for managing market risk in empirical finance (cf. Bauwens and Giot, 2000; Prigent et al., 2001; Giot, 2005; Dionne et al., 2009). We distinguish three different financial durations in the literature, namely the trade duration that is defined as the time elapsed between two consecutive trades, the price duration which is the time required to observe a change in the mid-price not less than a given threshold (ι_p), and the volume duration that is the time in want of trading certain amount of securities not less than a given threshold (ι_v). As we mention before, financial durations are characterized by high persistence and dispersion, but there exist some differences between them. Empirical data clearly show that persistence in trade durations is higher than that observed in price and volume durations. This can be observed when plotting the autocorrelation functions of adjusted empirical data which hyperbolical decay is more pronounced for trade durations than that for price or volume durations. While trade and price durations exhibit overdispersion, underdispersion is often observed in volume durations. In empirical applications two major problems are encountered when using financial duration data. The first one is how to properly remove seasonal patterns from the data without destroying their dynamics properties.

The second is how to deal with zero durations present in the data.

Engle and Russell (1998) report the presence of seasonality in high-frequency duration data. They observe high trading activity at the opening and closing time in the market and a slowdown of trading intensity around noon, which corresponds to lunchtime. These observations clearly show that data has stochastic and deterministic components. The deterministic component can in turn be categorized in a *day-of-the-week* effect and *time-of-the-day* effect (cf. Bauwens and Giot, 2000) and has to be removed before any estimation (cf. Engle and Russell, 1998). The presence of the deterministic component in the raw data can be explained by the behavior of market participants (traders, market-maker) and the institutional features of trading places. For instance, at the beginning of each trading day each trader or market-maker wants to benefit from the overnight macroeconomic news and this leads to higher trading intensity, and thus, short waiting time between transactions, and at the end of the trading day some traders precipitate to close their positions causing an increase of the trading activity. There exist two methods that are often used when computing the *time-of-the-day* function, namely the spline smoothing and the kernel smoothing. In Engle and Russell (1998) the *time-of-the-day* function is obtained as follows: each trading day is split in 13 intervals of thirty minutes, for each interval an expected duration is computed, and then cubic splines are used to smooth the *time-of-the-day* function on the thirty minutes intervals. The second method proposed by Veredas et al. (2001) consists in regressing the raw duration on the *time-of-the-day* in a non-parametric framework using a gamma kernel with the Nadaraya-Watson estimator.

Zero trade durations have first been reported by Engle and Russell (1998) when studying International Business Machines (IBM) intertrade durations (about two-thirds of the intertrade IBM data are zero durations). Engle and Russell (1998) dumped the zero durations before modeling the data. As by Engle and Russell (1998) zero trade durations are discarded by many authors. The authors justify their treatment approach for zero trade durations by arguing that simultaneous observations may arise from split-transactions, and therefore, are not crucial for the analysis. Recently, Veredas et al. (2001) claim that zero durations may have information content and their remove will affect the dynamic properties of the data. Their assertion has been confirmed by Bauwens (2006) who observed an increase of $Q-$statistics and residual autocorrelation after zero durations have been removed from the data. However, the results obtained by taking into consideration zero durations when analyzing trade durations (cf. Zhang et al., 2001; Bauwens, 2006) are not satisfactory.

Engle and Russell (1998) find clustering effects in financial durations, i.e. short durations tend to be followed by short durations and long durations by long durations. This result confirms the theoretical models developed by Kyle (1985), Admati and Pfleiderer (1988), and Easley and O'Hara (1992). They also find that neither exponential nor Weibull versions of ACD model are appropriate for modeling trade duration data. Bauwens and Giot (2000) find a significant negative impact of the trading intensity, the average volume per trade, and the average spread on the bid-ask quote process using the Log-ACD model. This finding is in conformity with the Easley and

O'Hara (1992) model. Engle and Lange (2001) proposed a new statistic, VNET, that can be used for quantifying realized market depth for a specific price deterioration, and find that market depth varies with volume, transactions and volatility. Hautsch (2003) finds that mid-quote changes, bid-ask spreads, and trading volumes are crucial for forecasting the intensity of liquidity demand.

Research by Bauwens et al. (2004) reveals that complicated ACD models such as TACD, SCD and SVD do not exhibit superior forecasting performances compared to the standard ACD or Log-ACD with flexible distributions (generalized gamma, Burr) for innovations. Allen et al. (2009) confirm the ability of the basic ACD model with flexible distributions (Burr, generalized gamma) to provide a good modeling of financial duration data. They also find that Lognormal distribution performs as well as generalized gamma distribution, and thus, is a convenient candidate for modeling duration data.

Chen et al. (2013) find that the MSMD model can reproduce the clustering effects, the non-linearities, and long memory features observed in financial data. Chen et al. (2013)'s model dominates the ACD model with exponential distribution for innovation in forecasting trade durations over different horizons. Baruník et al. (2012) also find that their MSMD model can properly reproduce most stylized facts of the price duration data. In empirical application they find that the model exhibits similar forecasting performance as the LMSD model of Deo et al. (2010). We finish by summarizing 27 studies on the financial duration data using different ACD and MSMD models with different distributional assumptions on innovations in Tables 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, and some diagnostic tests in Table 3.8. For each study we provide the author, the model used, the distributional assumptions about innovations, the type of financial durations, the stocks and the market places where they are traded.

3.7. Conclusion

This chapter has briefly presented various ACD and MSMD models developed in the empirical finance literature for modeling financial durations. It cannot be denied that the standard ACD model and its extensions achieve a lot of success in empirical finance. This success is overshadowed by the inability of the ACD models to properly capture the most *stylized facts* (scaling behavior, self-similarity, fat tails, long memory) of financial durations and trade-related variables such as bid-ask spread, trading volume recently reported in the literature. The applicability of the multifractal processes for modeling high-frequency duration data is in the early stages of development, but there is by now evidence that the MSMD models can provide better fit to financial durations, especially the intertrade durations than the traditional ACD models. We finish by identifying some research avenues that can be tracked in the future in order to provide a deep understanding of the intraday price process in the markets. More investigations are needed about zero durations in order to really know how the dynamic properties of the financial durations are influenced by zero durations. The MSMD models can be applied to analyze the impact of the market depth, the bid-ask spread and dealer costs on the trading process. They can also find applications in forecasting intraday market risk, liquidity risk and in pricing options. Multivariate MSMD models can be introduced to investigate the interdependence between price and duration processes. We are confident that these research can substantially help to better understand financial markets.

Table 3.1.: Different financial duration studies(1)

Authors	Models	Error distribution(s)	Type of duration	Stocks	Exchange
Engle and Russell (1998)	ACD	Exponential, Weibull	Trade: data from TORQ data set by Joel Hasbrouck (1992) 1/11/1990-31/01/1991	International Business Machines (IBM)	NYSE
Jasiak (1998)	FIACD		trade	Alcatel, IBM	SBF Paris Bourse NYSE
Lunde (1999)	ACD	exponential, Weibull generalized gamma	Trade, price: data from TAQ database 4/08/1997-30/09/1997	Disney Walt (DIS) Federal National Mortgage Ass. (FNM) General Motors (GM) Bank of America (BAC) McDonald's (MCD) Monsanto Company (MTC) Schlumberger Limited (SLB)	NYSE
Grammig and Maurer (2000)	ACD	exponential, Weibull	Price: data from TAQ database Sept. to Nov. 1996	Boeing (BA), IBM Coca-Cola (KO), Disney (DIS) Exxon (XON)	NYSE
Bauwens and Giot (2000)	ACD, Log-ACD	Exponential, Weibull	Price: data from TAQ database Sept. to Nov. 1996	Boeing (BA), IBM Disney (DIS)	NYSE

Table 3.2.: Different financial duration studies(2)

Authors	Models	Error distribution(s)	Type of duration	Stocks	Exchange
Dufour and Engle (2000a)	ACD, Log-ACD Box-Cox-ACD (BCACD)	exponential, Weibull generalized gamma	Trade: data from TORQ data set by Joel Hasbrouck (1992) 01/11/1990-31/01/1991	Boeing (BA), Calfed (CAL) Colgate (CL), IBM CPC International (CPC) Dresser Industries (DI) FPL Group (FPL), Exxon (XON) Federal Express (FDX) General Electric (GE) Glaxo Holdings PLC ADR (GLX) Hanson PLC ADR (HAN) Nipsco Industries (NI) Philip Morris Companies (MO) Federal National Mortgage Ass. (FNM), Potomac Electric Power Company (POM) American Telephone and Telegraph Co. (T), Schlumberger LTD (SLB)	NYSE
Zhang et al. (2001)	TACD, ACD	generalized gamma Weibull	Trade: data from TORQ data set by Joel Hasbrouck and NYSE 1/11/1990-31/01/1991	International Business Machines (IBM)	NYSE
Engle and Lange (2001)	ACD	Weibull	Market depth: data from NYSE TORQ data set 01/11/1990-31/01/1991	BA, CAL, CL, CPC DI, FDX, FNM, FPL GE, GLX, HAN, IBM MO, POM, SLB, T, XON	NYSE

Table 3.3.: Different financial duration studies(3)

Authors	Models	Error distribution(s)	Type of duration	Stocks	Exchange
Hujer et al. (2002)	Log-ACD, MSACD	Burr	Trade: data from TAQ database 1/11/1996-27/11/1996	Boeing (BA)	NYSE
Hautsch (2003)	ACD, Log-ACD BCACD	Generalized F	Excess volume: data from TAQ database 2/01/2001-31/05/2001	AOL, AT&T (T) Boeing (BA), Disney (DIS) IBM, JP Morgan (JPM) Philip Morris (PM)	NYSE
Bauwens and Giot (2003)	Asymmetric Log-ACD	Weibull	Price: data from TAQ database From Sept. to Nov. 1996	Walt Disney, IBM	NYSE
Feng et al. (2004)	SCD	Log exponential	Trade: data from Engle's website (IBM) 1/11/1990-21/12/1990 TAQ database (BA and KO) Feb. to March 2002	IBM, Boeing (BA) Coca Cola (KO)	NYSE
Bauwens and Veredas (2004)	SCD, ACD Log-ACD	Weibull, gamma	Trade, price, volume: data from TAQ database From Sept. to Nov. 1996	Boeing, Coca-Cola Disney, Exxon	NYSE
Bauwens et al. (2004)	ACD, Log-ACD SCD, SVD, TACD	Exponential, Weibull Burr, generalized gamma	Trade, price, volume: data from TAQ database From Sept. to Nov. 1996	Boeing, Coca-Cola Disney, Exxon	NYSE

Table 3.4.: Different financial duration studies(4)

Authors	Models	Error distribution(s)	Type of duration	Stocks	Exchange
Ghysels et al. (2004)	SVD, ACD	exponential	Trade (July, 1996)	Alcatel	PSE
Fernandes and Grammig (2005)	ACD	exponential, Weibull Burr, generalized gamma	Price: data from TAQ database Sept. to Nov. 1996	Exxon	NYSE
Fernandes and Grammig (2006)	AACD, ACD, BCACD Asymmetric Log-ACD EXACD, PACD Asymmetric PACD	Burr	Price: data from TAQ database Sept. to Nov. 1996	IBM	NYSE
Bauwens (2006)	ACD, Log-ACD	exponential, Weibull Burr, generalized gamma	Trade: data from Bloomberg From March until Juli 2003	Nippon Steel Corp. (NPS) Sony Corp. (SON) Toyota Motor Corp. (TOY) Tokyo Electric Power Co., Inc. (TKE)	TSE
Meitz and Teräsvirta (2006)	ACD, Log-ACD TACD, STACD TVACD	exponential	Trade: data from TAQ database 1/7/2002-31/12/2002	IBM	NYSE

Table 3.5.: Different financial duration studies(5)

Authors	Models	Error distribution(s)	Type of duration	Stocks	Exchange
Allen et al. (2008)	Log-ACD, ALACDI	Lognormal, Weibull	Price: data from SIRCA	Australia and New Zealand	ASX
	ALACDX	exponential, generalized gamma	1/7/2003-1/10/2003	Bank (ANZ), Coles Myer	
				Ltd (CML), Telstra (TLS)	
				Woolworths (WOW)	
				BHP Billiton (BHP)	
				Woodside Pretroleum (WPL)	
				QANTAS Airways (QAN)	
Bauwens et al. (2008)	Log-ACD, ACD	exponential, Weibull	Trade, Price, Volume:	Boeing, Coca-Cola	NYSE
	SCD	Burr, gamma	from TAQ database	Disney, Exxon	
		generalized gamma	Sept. 1996 to April 1997	IBM	
Sun et al. (2008)	ACD, Log-ACD	Lognormal, exponential	Trade: data from SIRCA	Alcoa Inc. (AA), Caterpillar (CAT)	NYSE
	TACD, SCD	stable distribution, Weibull	4/01/2003-31/12/2003	America Express (AXP), IBM	
		fractional Gaussian noise		E.I. Dupont de Nemours	
		fractional stable noise		Walt Disney (DIS)	
				General Electric (GE), 3M Co. (MMM)	
				General Motor (GM), McDonalds (MCD)	
				Eastman Kodak Co. (EKDKQ)	
				Int.Paper Company	
				Coca-Cola (KO), Altria Group (MO)	
				Merk & Co. (MRK), AT &T Inc.	
				United Technologies (UTX)	

Table 3.6.: Different financial duration studies(6)

Authors	Models	Error distribution(s)	Type of duration	Stocks	Exchange
Allen et al. (2009)	ACD, Log-ACD GV-ACD	exponential, Weibull	Price: the sampling periods for NAB, WBC, and TLS 2/01/2004-31/03/2004 2/04/2002-22/04/2002 2/01/2003-23/01/2003, respectively	National Australia Bank (NAB), Westpack Banking Corp (WBC) Telstra (TLS)	ASX
Deo et al. (2010)	LMSD, ACD	exponential, Weibull	Trade, price: data from TAQ database July to December 2002	American International Group (AIG), American Express (AXP), Boeing (BA) Coca-Cola (KO), IBM Beverly Enterprizes (BEV) Bally Total Fitness (BFT) CBL Associates Properties (CBL) Commercial Federal Corporation (CFB), Sony (SNE)	NYSE
Bhatti (2010)	ACD	Birnbaum-Saunders generalized gamma	Trade: data from TAQ database 01/01/2002-28/02/2002	General Motors (GM), IBM Johnson and Johnson Company (JNJ) McDonald (MCD), Proctor and Gamble Company (PG), Schlumberger Limited (SLB)	NYSE

Table 3.7.: Different financial duration Studies(7)

Authors	Models	Error distribution(s)	Type of duration	Stocks	Exchange
Chen et al. (2013)	MSMD	Exponential	Trade: data from	Alcoa (AA), American Express (AXP)	NYSE
			TAQ database	Abbott Laboratories (ABT), Dell (DELL)	
			1/2/1993-26/2/1993	Bank of America (BAC), Ford Motor (F)	
				Cisco Systems (CSCO), Microsoft (MSFT)	
				General Electric (GE), IBM	
				Intel Corporation (INTC), Coca-Cola (KO)	
				Texas Instruments (TXN), Wal-Mart (WMT)	
				Xerox Corp (XRX), Merck & Co (MRK)	
				McDonald's Corp (MCD), Wells Fargo (WFC)	
				Johnson & Johnson Inc (JNJ), Intel Corp (INTC)	
Baruník et al. (2012)	MSMD	Exponential, Weibull	Price: data from	Swiss Franc (CHF), Euro EUR, Japanese Yen (JPY)	CME
			TickData, Inc.		
			9/11/2009-29/01/2010		

Table 3.8.: Different diagnostic tests

Authors	Diagnostic tests
Engle and Russell (1998)	Ljung-Box test, overdispersion test
	Test for detecting nonlinear dependence
Jasiak (1998)	Ljung-Box test
Bauwens and Giot (2000)	Ljung-Box test
Dufour and Engle (2000a)	LM-test
Prigent et al. (2001)	Bartlett identity test
Ghysels et al. (2004)	Ljung-Box test
Bauwens and Veredas (2004)	QQ-plots
Bauwens et al. (2004)	Density forecasts
Fernandes and Grammig (2005)	D-test, H-test
Meitz and Teräsvirta (2006)	LM-test, general battery of tests of LM
Hautsch (2006)	LM-tests, CM-tests, ICM-tests
Duchesne and Pacurar (2008)	Generalized Box-Pierce/Ljung-Box test
Chen and Hsieh (2010)	Generalized moments tests
Chen et al. (2013)	White's information-matrix test

Part II.

Forecasting Return Volatility: An Application To Crude Oil Prices

4. Modeling and Forecasting Crude Oil Price Volatility: Evidence from Historical and Recent Data

4.1. Introduction

The recent literature shows a growing interest in modeling and forecasting oil price volatility due to its impact on the global and regional economies (cf. Wang et al., 2012; Rahman and Serletis, 2012). How oil price shocks may affect economic growth is well-documented in a large body of research. Different transmission mechanisms were developed in the literature. Examples include Rotemberg and Woodford (1996) and Finn (2000), among others. Papers by Hamilton (1983) Davis and Haltiwanger (2001), and Lee and Ni (2002) clearly demonstrated that positive oil price shocks induce a slow-down in aggregate measures of growth or employment and that negative oil price shocks lead to an increase in aggregate measures of growth or employment. Recently, Elder and Serletis (2010) found that increased uncertainty about oil price changes causes a significant drop in real output and heavily affects measures of durable consumption and fixed investment in the United States. Their finding is also confirmed by Rahman and Serletis (2012) for the Canadian economy. In his seminal paper, Hamilton (2003) confirmed the existence of a strong relationship between oil price changes and GDP growth and showed that this relationship is of a nonlinear nature. Jones and Kaul (1996) and Sadorsky (1999) showed that oil price shocks have direct or indirect influence on financial markets. According to Backus and Crucini (2000) they may be responsible for fluctuations in the international terms of trade. Oil price volatility also represents an important input for macro-econometric models (cf. Ferderer, 1996), pricing of derivatives (cf. Wang et al., 2008) and portfolio selection models (cf. Geman and Kharoubi, 2008). So, it is of primary importance for firms, financial market participants and policy makers to have models available that can properly reproduce the *stylized facts* of oil price volatility and provide accurate forecasts.

The widespread tool used in the literature to analyze oil price volatility consists in GARCH-type models (cf. Kang et al., 2009; Cheong, 2009; Mohammadi and Su, 2010; Wei et al., 2010). All these papers have attempted to find the most appropriate GARCH-type models, linear or nonlinear, that can properly reproduce the stylized facts of oil price volatility, and thus, produce accurate forecasts. While some results speak in favor of fractionally integrated GARCH (FIGARCH) models (cf. Kang et al., 2009), others provide evidence that the standard GARCH and FIAPARCH (cf.

Cheong, 2009), and the APARCH models (cf. Mohammadi and Su, 2010) could be more appropriate. In contrast to the previous papers, Wei et al. (2010) consider nine GARCH-type models and compare their forecasting performance based on six different loss functions. They found that none of these models can consistently outperform each other, despite the fact that the nonlinear models can properly capture long memory volatility and/or the asymmetric leverage effect in volatility.

This chapter extends the work of Wei et al. (2010) in two important respects: (i) we add to the set of GARCH models used in Wei et al. (2010) a new type of volatility model, namely the Markov switching multifractal (MSM) model, (ii) we consider a large data set that contains oil price observations of the pre- and post-1900 eras. Our objective is to compare the forecasting performance of the MSM model with that of GARCH models. Availability of daily data for a twenty-year period within the 19th century provides the valuable opportunity to compare the statistical features of the modern oil market with those of a much earlier phase of the same market. The multifractal[1] model provides a completely new approach to the modeling of financial volatility which it conceives as a multiplicative, hierarchically structured process. Via its particular principles of construction, it allows to estimate a Markov-switching model with a high number of states without falling victim to the curse of dimensionality. This structure gives it an intermediate nature between "true" long-memory processes and simple regime-switching processes allowing to modulate the temporal dependency via its parameters and the number of hierarchical components. The flexible regime-switching nature makes it attractive for time series that show pronounced differences between highly volatile and more tranquil periods (as oil prices do). Research on stock and foreign exchange markets has documented superior forecasting capabilities of MSM against traditional GARCH models (Calvet and Fisher, 2004b; Lux and Kaizoji, 2007; Lux et al., 2014). It seems interesting to explore in how far these findings can be confirmed with important commodities such as oil. As in Wei et al. (2010), we also use six different loss functions as criteria for comparison, and then apply the predictive ability test of Hansen (2005) in order to infer whether one particular model is outperformed by others or not. Here we prefer the predictive ability test of Hansen (2005) to other powerful evaluation techniques existing in the literature (cf. Diebold and Mariano, 1995; West, 1996; White, 2000) due to its robustness, and the fact that it allows to compare a benchmark (possibly nested) model for a whole set of competitors.

The remainder of the chapter is organized as follows. Section 4.2 presents the descriptive statistics of our data sets. Section 4.3 introduces the different volatility models. The forecasting evaluation methodologies are presented in Section 4.4 and results are provided in Section 4.5. Finally, Section 4.6 concludes.

[1] The term multifractal refers to the fractal structure of the resulting volatility process. The MSM has actually been adapted from very similar models that have first been developed for turbulent flows (cf. Mandelbrot, 1974). Fractality is also a concept that plays an important role in geophysical research and petroleum geology (cf. Barton and La Pointe, 1995), but it seems unlikely that the two aspects - fractality of oil fields and fractality of oil price volatility - are materially related to each other.

4.2. Data

We use daily closing oil prices (in US dollars per barrel) of West Texas Intermediate (WTI) over two different sample periods. The first one covers the period from January 02, 1875 to December 31, 1895 and the second one runs from January 03, 1977 to March 24, 2014. For the more recent era, we also split the sample into two different parts. This will help us to better observe the time evolution of oil prices. The samples are driven purely by availability of daily data at the time of writing this paper, with the data being sourced from the Global Financial Database, https://www.globalfinancialdata.com. We compute the percent continuously compounded returns r_t as

$$r_t = 100 * [\ln(p_t) - \ln(p_{t-1})], \tag{4.1}$$

where p_t denotes the oil price at the end of period t and p_{t-1} is the oil price on the previous day.

To get some first impression of our data sets we first plot the oil prices, their log-returns and squared log-returns (cf. *Figs.* 4.1 through 4.8). Their descriptive statistics are reported in Tables 4.1, 4.2, 4.3 and 4.4. The data sets exhibit high variability, in other words the standard deviations are very high compared to the sample means. We observe positive skewness for the data set of pre-1900 and a negative one for the data set of post-1900. Both data sets exhibit excess kurtosis. These results show that the computed log-returns do not follow a Normal distribution. This observation is confirmed by the Jarque-Bera test, which rejects the null hypothesis of Normally distributed log-returns at any level of significance. We also apply the augmented Dickey-Fuller (ADF) unit-root test of Dickey and Fuller (1979) to oil returns and the results clearly speak for the stationarity of both data sets. The Hurst indices reported in Tables 4.1, 4.2, 4.3 and 4.4 are computed via Detrended Fluctuation Analysis (DFA) (cf. Weron, 2002). The Hurst index values for log-returns are close to 0.5 and not significantly different from this value at the 95% confidence level, implying absence of long memory features in oil price returns. For absolute and squared returns the Hurst index values are significantly above 0.5, indicating the presence of long memory in oil price volatility. Finally, in order to show the decay of the unconditional distribution of oil price returns in its extremal region, we compute the so-called Hill estimator for the tail index (cf. Hill, 1975b). We find that the estimates for the tail indices are in the vicinity of 3 and these results are in harmony with typical findings for other commodities and financial assets, cf. Tables 4.1, 4.2, 4.3 and 4.4.

Figs. 4.2, 4.4, 4.6 and 4.8 depict the autocorrelation functions of log-returns, absolute and squared log-returns. We observe that the absolute and squared log-returns are highly correlated and this observation is in conformity with the Ljung-Box statistics, $Q(10)$ and $Q(20)$. The Ljung-Box tests also reject the null hypothesis of no serial correlation for raw log-returns at the 5% significance level. This indicates the presence of some serial dependence in the oil price log-returns. The higher statistics of the Ljung-Box statistics for the raw returns in the 19th century might indicate a lower degree of "financialisation" of this commodity at earlier times.

4.3. Model Framework

In this section we briefly present the volatility models used for our forecasting exercises. In general, financial returns in these models are formalized as

$$r_t = \mu_t + \sigma_t e_t, \tag{4.2}$$

where $r_t = 100 * [\ln(P_t) - \ln(P_{t-1})]$, $\ln(P_t)$ is the log asset price, $\mu_t = \mathbb{E}_{t-1}[r_t]$ is the conditional mean of the return series, σ_t is the volatility process and e_t is standard Normally distributed. Defining $x_t = r_t - \mu_t$, the *centered* returns are given by

$$x_t = \sigma_t e_t. \tag{4.3}$$

In this chapter we assume that μ_t follows an AR(1) process and consider two different types of volatility models for describing σ_t, namely the linear and nonlinear GARCH-type models and the Markov switching multifractal (MSM) model.

4.3.1. GARCH-type Models

The underlying idea of the autoregressive conditional heteroskedasticity (ARCH) model was developed by Engle (1982) in his seminal paper. The ARCH model and its subsequent generalized versions are well known in the literature for their ability to capture the most important *stylized facts* (e.g. clustering effects, long-memory and short-memory effects, asymmetric leverage effects) observed in all measures of volatility (e.g. absolute log-returns, squared log-returns, etc...). In the following we list the eight different GARCH models used in this study.

4.3.1.1. The GARCH and IGARCH Models

Introduced by Bollerslev (1986) the linear GARCH model is the most popular volatility model in the literature. In the simple, but effective GARCH(1,1) (cf. Bollerslev et al., 1994) the conditional variance is modeled as

$$\sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{4.4}$$

where $\omega > 0, \alpha > 0, \beta > 0$ and $\alpha + \beta < 1$. The nonnegativity constraints on ω, α and β guarantee the positivity of σ_t^2.

h-step ahead forecasts from GARCH(1,1) are obtained recursively as

$$\begin{aligned}\hat{\sigma}_{t+h}^2 &= \omega + (\alpha + \beta)\,\hat{\sigma}_{t+h-1}^2 \\ &= \bar{\sigma}^2 + (\alpha + \beta)\left(\hat{\sigma}_{t+h-1}^2 - \bar{\sigma}^2\right) \\ &= \bar{\sigma}^2 + (\alpha + \beta)^{h-1}\left(\hat{\sigma}_{t+1}^2 - \bar{\sigma}^2\right).\end{aligned} \tag{4.5}$$

where $\bar{\sigma}^2 = \omega(1-\alpha-\beta)^{-1}$ is the unconditional variance. As $h \to \infty$, it is clear that the volatility forecast in *eq.* (4.5) approaches the unconditional variance $\bar{\sigma}^2$ and $(\alpha+\beta)$ dictates the speed of the mean reversion.

If $\alpha+\beta = 1$, the GARCH(1,1) reduces to the IGARCH(1,1) model proposed by Engle and Bollerslev (1986b) in order to account for infinite persistence in the conditional variance. The h-step ahead forecast representation becomes

$$\begin{aligned} \hat{\sigma}^2_{t+h} &= \hat{\omega} + \hat{\sigma}^2_{t+h-1} \\ \hat{\sigma}^2_{t+h} &= \hat{\omega}h + \hat{\sigma}^2_t. \end{aligned} \tag{4.6}$$

4.3.1.2. The Exponential GARCH Model

The exponential GARCH (EGARCH) model was proposed by Nelson (1991) with the aim to capture the asymmetric relation between stock returns and volatility changes noted by Black (1976). The conditional variance in the EGARCH (1,1) model is given by

$$\ln\left(\sigma^2_t\right) = \omega + \alpha e_{t-1} + \gamma\left(|e_{t-1}| - \mathbb{E}\left[|e_{t-1}|\right]\right) + \beta \ln\left(\sigma^2_{t-1}\right), \tag{4.7}$$

where γ represents the asymmetric leverage parameter that quantifies the degree of the volatility leverage effect in the model and α the magnitude. As in *eq.* (4.2), $e_t \sim N(0,1)$ with $\mathbb{E}\left[|e_{t-1}|\right] = \sqrt{2/\pi}$. The model parameters are free from nonnegativity constraints.

Following the same procedures as with GARCH(1,1), the h-step ahead forecast formula of the EGARCH(1,1) can be expressed as

$$\ln\hat{\sigma}^2_{t+h} = \bar{\sigma}^2 + \beta^{h-1}\left(\ln\hat{\sigma}^2_{t+1} - \bar{\sigma}^2\right), \tag{4.8}$$

where $\bar{\sigma}^2 = (\omega - \gamma/\sqrt{2/\pi})/(1-\beta)$.

4.3.1.3. The Glosten/Jagannathan/Runkle GARCH Model

The GJR-GARCH model developed by Glosten et al. (1993) is designed in a way that allows the model to account for the potential larger impact of negative shocks on return volatility. The conditional variance in the GJR-GARCH(1,1) can be formalized as

$$\sigma^2_t = \omega + \left[\alpha + \gamma D(x_{t-1} < 0)\right] x^2_{t-1} + \beta\sigma^2_{t-1}, \tag{4.9}$$

where $D(.)$ is an indicator function that takes the value 1 if $x_{t-1} < 0$ (bad news), and 0 (good news) otherwise. The parameter γ quantifies the magnitude of the asymmetric leverage effect. The h-step ahead forecast representation of the GJR-GARCH(1,1) can be formalized as

$$\hat{\sigma}^2_{t+h} = \bar{\sigma}^2 + \left(\alpha+\beta+\frac{\gamma}{2}\right)^{h-1}\left(\hat{\sigma}^2_{t+1} - \bar{\sigma}^2\right), \tag{4.10}$$

where $\bar{\sigma}^2 = \omega/(1-\alpha-\beta-\gamma/2)$ is the unconditional or long run variance.

4.3.1.4. The Asymmetric Power ARCH Model

The asymmetric power ARCH (APARCH) model introduced by Ding et al. (1993) aims to reproduce both leverage and the *Taylor* effect, named after Taylor (1986) who first documented the fact that the sample autocorrelation of absolute returns was usually larger than that of squared returns. The conditional variance in the APARCH(1,1) model is given by

$$\sigma_t^\delta = \omega + \alpha\left(|x_{t-1}| - \gamma x_{t-1}\right)^\delta + \beta\sigma_{t-1}^\delta, \tag{4.11}$$

where $\delta > 0$ and γ is the leverage coefficient. The APARCH(1,1) model reduces to GARCH(1,1) when $\delta = 2$ and $\gamma = 0$.

The h-step ahead forecast formula of the APARCH(1,1) is given by

$$\begin{aligned}\hat{\sigma}_{t+h}^\delta &= \omega + \left(\alpha\mathbb{E}_t\left[\left(|e_{t+h-1}| - \gamma e_{t+h-1}\right)^\delta\right] + \beta\right)\hat{\sigma}_{t+h-1}^\delta \\ &= \kappa + (\alpha c + \beta)^{h-1}\left(\hat{\sigma}_{t+1}^\delta - \kappa\right),\end{aligned} \tag{4.12}$$

where $\kappa = \omega(1-\alpha c-\beta)^{-1}$ is the long run variance to the power δ and $c = \mathbb{E}_t\left[\left(|e_{t+h-1}| - \gamma e_{t+h-1}\right)^\delta\right]$ is given by

$$c = \frac{1}{\sqrt{2\pi}}\left[(1+\gamma)^\delta + (1-\gamma)^\delta\right]2^{\frac{\delta-1}{2}}\Gamma\left(\frac{\delta+1}{2}\right).$$

4.3.1.5. The Fractionally Integrated GARCH Model

By introducing fractional differences in the GARCH process Baillie et al. (1996) obtained the FIGARCH model that can reproduce the long memory property of financial returns volatility. The FIGARCH(1,d,1) model volatility can be expressed as

$$\sigma_t^2 = \omega + \left[1 - \beta(L) - \phi(L)(1-L)^d\right]x_t^2 + \beta\sigma_{t-1}^2, \tag{4.13}$$

where $\omega > 0$, $\phi < 1$, $\beta < 1$, $0 \leq d \leq 1$. L denotes the lag operator and d is the parameter of fractional differentiation. The parameters have to fulfill the following conditions:

$$\beta - d \leq \phi \leq \frac{(2-d)}{3} \tag{4.14}$$

and

$$d\left[\phi - \frac{(1-d)}{2}\right] \leq \beta(d - \beta + \phi). \tag{4.15}$$

We can rewrite *eq.* (4.13) as follows

$$\begin{aligned}\sigma_t^2 &= \omega(1-\beta)^{-1} + \left[1-(1-\beta)\phi(L)(1-L)^d\right]x_t^2 \\ &= \omega(1-\beta)^{-1} + \eta(L)x_t^2,\end{aligned} \tag{4.16}$$

where $\eta(L) = \eta_1 L + \eta_2 L^2 + \ldots, \eta_j \geq 0$ for $j = 1, 2, \ldots$.
$\eta(L)$ can be computed from the recursions:

$$\begin{cases}\eta_1 = & \hat{\phi} - \hat{\beta} + \hat{d}, \\ \vdots & \vdots \\ \eta_j = & \hat{\beta}\eta_{j-1} + \left[\left(j-1-\hat{d}\right)j^{-1} - \hat{\phi}\right]\pi_{j-1}\end{cases} \tag{4.17}$$

where $\pi_j \equiv \pi_{j-1}\left(j-1-\hat{d}\right)j^{-1}$ are the coefficients in the MacLaurin series expansion of the fractional differencing operator $(1-L)^d$. As in previous research, we set the truncation order of the infinite series $(1-L)^d$ to 1000 lags.

The FIGARCH model reduces to the GARCH model when $d = 0$ and the IGARCH model when $d = 1$.

From *eq.* (4.16) one can easily derive the one-step ahead forecast of σ_t^2

$$\hat{\sigma}_{t+1}^2 = \omega(1-\beta)^{-1} + \eta_1 x_t^2 + \eta_2 x_{t-1}^2 + \ldots \tag{4.18}$$

Using recursive substitution described above the h-step ahead forecasts of the FIGARCH(1,d,1) are obtained as

$$\hat{\sigma}_{t+h}^2 = \omega\,(1-\beta)^{-1} + \sum_{i=1}^{h-1}\eta_i\hat{\sigma}_{t+h-i}^2 + \sum_{j=0}^{\infty}\eta_{h+j}x_{t-j}^2. \tag{4.19}$$

4.3.1.6. The Hyperbolic GARCH Model

Recently developed by Davidson (2004), the hyperbolic GARCH (HYGARCH) model is constructed in a way that allows the model not only to reproduce long memory features in volatility of many financial time series, but also (unlike FIGARCH) to be covariance stationary. The HYGARCH(1,d,1) process models the conditional variance as

$$\begin{aligned}\sigma_t^2 &= \omega + \left\{1-\beta(L)-\phi(L)\left[(1-\tau)+\tau(1-L)^d\right]\right\}x_t^2 + \beta\sigma_{t-1}^2 \\ &= \omega(1-\beta)^{-1} + \lambda(L)x_t^2\end{aligned} \tag{4.20}$$

where $\lambda(L) = \left\{1-(1-\beta(L))\phi(L)\left[(1-\tau)+\tau(1-L)^d\right]\right\}$, $\omega > 0$, $\phi < 1, \beta < 1$, $0 \leq d \leq 1$ and $\tau \geq 0$. $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \ldots, \lambda_j \geq 0$ for $j = 1, 2, \ldots$. L is the lag operator and the HYGARCH model reduces to FIGARCH and IGARCH when $\tau = 1$ and $\tau = 0$, respectively. *Eqs.* (4.14) and (4.15) become

$$\beta - \tau d \leq \phi \leq \frac{(2-d)}{3} \tag{4.21}$$

and

$$\tau d\left[\phi - \frac{(1-d)}{2}\right] \leq \beta(\tau d - \beta + \phi). \tag{4.22}$$

We refer the reader to Conrad (2010) for more details on the non-negativity conditions for the HYGARCH model and for the proof for the covariance stationarity of the process. The *h*-step ahead forecasts of the HYGARCH(1,d,1) are easily obtained by following the same procedures used for FIGARCH(1,d,1).

4.3.1.7. The Fractionally Integrated APARCH Model

Inspired by the FIGARCH model Tse (1998) incorporates fractional differences into the asymmetric power ARCH model of Ding et al. (1993) to obtain the fractionally intergrated APARCH model. The FIAPARCH(1,d,1) model is defined as

$$\sigma_t^\delta = \omega + \left[1 - \beta(L) - \phi(L)(1-L)^d\right](|x_{t-1}| - \gamma x_{t-1})^\delta + \beta\sigma_{t-1}^\delta, \tag{4.23}$$

where $\omega > 0, \phi < 1, \beta < 1, 0 \leq d \leq 1$ and $-1 < \gamma < 1$.

The FIAPARCH process seems to be a promising model due to the fact that it is able to simultaneously capture long memory and asymmetric leverage effects in the data. The FIAPARCH model encompasses the FIGARCH model for $\gamma = 0$ and $\delta = 2$. Following the same procedures described above the forecasts for future variance can be easily obtained.

Note that the parameters in all formulas for forecasting future volatility have to be replaced by their corresponding estimates. All GARCH-type models are estimated via (quasi-) maximum likelihood as it is customary in the literature.

4.3.2. The Markov-Switching Multifractal Model

The recently introduced Markov-switching multifractal models are characterized by a multiplicative rather than additive structure of the volatility process. In the MSM framework instantaneous volatility is modeled as a product of k volatility components or multipliers $M_t^1, M_t^2, \ldots, M_t^k$ and a positive scale factor σ^2 (cf. Calvet and Fisher, 2001b, 2004b; Lux, 2008). Formally, we have

$$\sigma_t^2 = \sigma^2 \prod_{i=1}^{k} M_t^{(i)}. \tag{4.24}$$

The multipliers or volatility components are assumed to be independent of each other at any time and satisfy $\mathbb{E}\left[M_t^i\right] = 1$. Each multiplier M_t^i is renewed at time t with probability γ_i depending on its rank within the hierarchy of multipliers and remains unchanged with probability $1 - \gamma_i$. In their seminal paper Calvet and Fisher (2001b) derived a formalization for the transition probabilities, γ_i, that guarantee the convergence of the discrete-time MSM to a Poisson multifractal process in the continuous-time limit. Here we are not interested in the continuous-time process, and therefore, we prefer to use the pre-specified transition probabilities proposed by Lux (2008) that are given by

$$\gamma_i = 2^{i-k}. \tag{4.25}$$

To fully specify the MSM model we assume that the random multipliers follow a Lognormal[2] distribution with parameters λ and ν, i.e.,

$$M_t^i \sim LN(-\lambda, \nu). \tag{4.26}$$

We normalize the distribution of the multipliers to guarantee $\mathbb{E}\left[M_t^i\right] = 1$ which leads to

$$\exp\left(-\lambda + \frac{1}{2}\nu^2\right) = 1. \tag{4.27}$$

From *eq.* (4.27) it is obvious that the shape parameter ν can be expressed as: $\nu = \sqrt{2\lambda}$. With this restriction the Lognormal distribution of multipliers is fully defined by the scale parameter λ. So, the parameters to be estimated in the Lognormal MSM (LMSM) are only λ and σ. We carry out their estimation for all specifications $k = 2, \ldots, 20$ using the GMM approach proposed by Lux (2008). We then choose the specification with the lowest GMM criterion as our preferred model for the subsequent forecasting exercise. Note that higher k increases the number of regimes (which is 2^k), and generates proximity to long memory over a larger number of lags, but comes at no additional computational cost in our approach. The pertinent moments used for the estimation can be found in Lux (2008). Note that maximum likelihood would be possible only for MSM models with a finite, discrete support of the multipliers, and computationally feasible only for a limited number of hierarchical components up to about 8.

We perform the out-of-sample forecasting on the base of the LMSM model using the standard approach for best linear forecasts outlined in Brockwell and Davis (1991) together with the generalized Levinson-Durbin algorithm proposed by Brockwell and Dahlhaus (2004). The forecasting procedure is performed in two steps.

1. In the first step: We compute the following zero-mean time series

$$Z_t = x_t^2 - \mathbb{E}\left[x_t^2\right] = x_t^2 - \sigma^2, \tag{4.28}$$

where $\hat{\sigma}$ is the estimate of the scale factor σ.

[2] Other distributional assumptions such as Binomial, Gamma can be used as well, but have been found to make little difference in previous literature, cf. Liu et al. (2007), Lux (2008).

2. In the second step: Assuming that the oil price volatility data follow the stationary process $\{Z_t\}$ defined in the first step, h-step best linear forecasts are given by

$$\hat{Z}_{n+h} = \sum_{i=1}^{n} \psi_{ni}^{(h)} Z_{n+1-i} = \Psi_n^{(h)} \mathbf{Z}_n, \tag{4.29}$$

where the vectors of weights $\Psi_n^{(h)} = \left(\psi_{n1}^{(h)}, \psi_{n2}^{(h)}, \ldots, \psi_{nn}^{(h)}\right)'$ are solutions of

$$\Gamma \Psi_n^{(h)} = \gamma_n^h, \tag{4.30}$$

with $\gamma_n^h = (\gamma(h), \gamma(h+1), \ldots, \gamma(n+h-1))'$ being the auto-covariances for the data generating process of Z_t at lags h and beyond, and $\Gamma_n = [\gamma(i-j)]_{i,j=1,\ldots,n}$ the pertinent variance-covariance matrix. The pertinent auto-covariances for the multifractal model can be found in Lux (2008).

In sum, our portfolio of volatility models includes two linear GARCH models (GARCH, IGARCH), six nonlinear GARCH models (EGARCH, GJR-GARCH, APARCH, FIGARCH, HYGARCH, FI-APARCH) and one multifractal model (LMSM).

4.4. Forecast Evaluation Methodologies

To obtain our forecasts we proceed as follows: We first split the pre-1900 data set containing oil price observations from January 3, 1875 to December 31, 1895 into two subgroups. The first one covers the period from January 3, 1875 to December 31, 1892 and is used as in-sample data for model estimation. The second one contains oil prices of the last three years, i.e., from January 3, 1893 to December 31, 1895 and serves as out-of-sample data that we use for evaluation purposes. The estimation period is rolled forward by adding one observation and removing one day by day, so that the size of the data set used for the estimation remains fixed over the out-of-sample period. Forecasts are computed for horizons of various lengths: 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 days.

Second, we take a portion of the post-1900 data that contains oil price observations from January 6, 1992 to December 31, 2009 and split the oil price observations into in-sample data for volatility estimation covering the period from January 6, 1992 to December 29, 2006 and out-of-sample data stretching over the period from January 2, 2007 to December 31, 2009, which is in line with Wei et al. (2010). The great recession of 2008-2009 after the global financial crisis of 2007-2008 caused a demand contraction of oil and oil prices fluctuated from USD 145.31 (July 03, 2008) to USD 30.28 per barrel (December 23, 2008). Therefore, we find that this period should be interesting for testing the performance of our volatility models.

Third, we consider the extended data set covering the period from January 06, 1992 to March 24, 2014. This period of time does not cover only the great recession of 2008-2009, but also the subsequent recovery of the world economy. During this period the oil price stabilized at about USD 100 per barrel. We use oil price observations from January 6, 1992 until December 31, 2009 as in-sample data and the remaining observations, i.e., oil prices from January 4, 2010 to March 24, 2014 as of-out-sample data.

Finally, we also take the whole post-1900 data, i.e., from January 3, 1977 to March 24, 2014 to evaluate the contribution of a longer in-sample set. We use oil price observations from January 3, 1977 until December 31, 2009 as in-sample data and the remaining observations, i.e., oil prices from January 4, 2010 to March 24, 2014 as of-out-sample data. Note that forecasts in the second, third and fourth forecasting experiments are computed as previously done in the first one.

4.4.1. Forecasting Evaluation Criteria

We evaluate the forecasting ability of our volatility models in all four forecasting experiments by means of the following six different loss functions:

$$\text{MSE} = T^{-1}\sum_{i=1}^{T}\left(\sigma^2_{f,t} - \sigma^2_{a,t}\right)^2, \tag{4.31}$$

$$\text{MAE} = T^{-1}\sum_{i=1}^{T}\left|\sigma^2_{f,t} - \sigma^2_{a,t}\right|, \tag{4.32}$$

$$\text{HMSE} = T^{-1}\sum_{i=1}^{T}\left(1 - \frac{\sigma^2_{a,t}}{\sigma^2_{f,t}}\right)^2, \tag{4.33}$$

$$\text{HMAE} = T^{-1}\sum_{i=1}^{T}\left|1 - \frac{\sigma^2_{a,t}}{\sigma^2_{f,t}}\right|, \tag{4.34}$$

$$\text{QLIKE} = T^{-1}\sum_{i=1}^{T}\left[\ln\left(\sigma^2_{f,t}\right) + \frac{\sigma^2_{a,t}}{\sigma^2_{f,t}}\right], \tag{4.35}$$

$$\text{RLOG} = T^{-1}\sum_{i=1}^{T}\left[\ln\left(\frac{\sigma^2_{a,t}}{\sigma^2_{f,t}}\right)\right]^2, \tag{4.36}$$

where $\sigma^2_{f,t}$ denotes the volatility forecast obtained using a GARCH-type model or MSM model, $\sigma^2_{a,t}$ is the daily actual volatility that is computed using the daily squared returns, and T denotes the number of out-of-sample observations. MSE and MAE are the mean square error and mean absolute error, respectively, and HMSE and HMAE are their corresponding heteroscedasticity adjusted statistics. QLIKE quantifies the loss implied by a Gaussian likelihood and RLOG puts more weight on small observations (cf. Bollerslev et al., 1994).

All the above-mentioned loss functions are well known in the literature and each of them can be used depending on the contexts and the objective of the users. However, based only on these loss function criteria, it is difficult to conclude that the forecasting performance of one model dominates that of the other one. To draw such conclusions, we need statistical tests that can provide more reliable information. In the next section, we briefly describe the superior predictive ability (SPA) test of Hansen (2005).

4.4.2. Superior Predictive Ability Test

The superior predictive ability (SPA) test of Hansen (2005) sheds light on the relative performance of a particular model in comparison with its competitors. In other words, it answers the question whether any of the alternative models are better than the particular benchmark model in terms of expected loss. The null hypothesis that the benchmark model is not dominated by any of the other competitive models is postulated as follows

$$H_0 : \max_{i=1,\ldots,K} \mathbb{E}\left[d_t\right] \leq 0, \tag{4.37}$$

where $d_t = (d_{i,t}, \ldots, d_{K,t})'$ is a vector of relative performances, $d_{i,t}$, that are computed as $d_{i,t} = L^{(0)}_{t,h} - L^{(i)}_{t,h}$. K is the number of the competitive models, h denotes the forecasting horizon and $L^{(0)}_{t,h}$ and $L^{(i)}_{t,h}$ are the loss functions at time t for a benchmark model M_0 and for its competitor models, $M_{i_{(i=1,\ldots,K)}}$, respectively.

The associated test statistic is given by

$$\mathrm{SPA} = \max_{i=1,\ldots,K} \frac{\sqrt{T}\bar{d}_i}{\sqrt{\lim\limits_{T\to\infty} \mathbb{V}ar(\sqrt{T}\bar{d}_i)}}, \tag{4.38}$$

where $\bar{d} = T^{-1}\sum d_t$. We use a stationary bootstrap procedure to obtain the p-values of the SPA. A high p-value indicates non-rejection of the null hypothesis that a particular model is not outperformed by its competitors. We refer the reader to Hansen (2005) for more details on technical issues.

4.5. Empirical Results

4.5.1. Estimation Results

We estimate the GARCH models via the ML approach and the results are reported in Tables 4.5, 4.6, 4.7 and 4.8. Overall the estimates of β in GARCH, IGARCH, EGARCH, GJR-GARCH and APARCH models are close to 1 and significant at the 1% level. While the asymmetric leverage parameters are significant at the 1% level in the EGARCH model in Tables 4.5, 4.6 and 4.7 and not significant at any level in Table 4.8, they are insignificant at any level in the GJR-GARCH and APARCH models.

With the pre-1900 oil price data, the estimate of τ in the HYGARCH model is quite close to 1 and significant at the 1% level. The estimates of δ are 1.748 in the APARCH model and 1.315 in the FIAPARCH model. In contrast to the APARCH model the asymmetric leverage parameter in the FIAPARCH model is significant at the 1% level. The estimates of d in FIGARCH, HYGARCH and FIAPARCH models are significant at the 1% level and give evidence of the presence of long memory effects in oil price volatility.

With the post-1900 oil price data, we first estimate the GARCH models using oil price observations from January 6, 1992 to December 31, 2009. Here the estimate of τ in the HYGARCH model is significant at the 1% level and different from 1. By expanding the estimation sample, i.e. from January 6, 1992 to March 24, 2014, we do not observe a dramatic change in the estimation results. The estimates of d are significant at the 1% level in all long memory GARCH models.

Finally, we estimate the whole post-1900 oil price data. The estimates of d in FIGARCH, HYGARCH and FIAPARCH models are now equal to 1 and significant at the 1% confidence level. These results indicate the presence of infinite persistence in the oil price data post-1900.

When we look at the estimation diagnostics, it seems that the three long memory GARCH models perform better in terms of fitting oil price observations over all different periods of time. In sum, the Log(L), AIC and BIC for the long memory models are smaller than those of short memory models. Furthermore, the Ljung-Box tests on the squared residuals and the ARCH tests also speak in favor of the long memory models. For all three long memory models the Ljung-box tests mostly cannot reject the null hypothesis of no serial correlation in the squared standardized residuals at the 5% level and the ARCH tests mostly accept the null hypothesis that the standardized residuals consist of independent identically distributed (*i.i.d*) Gaussian disturbances.

We now turn to the estimation of the Lognormal MSM. The best GMM objective function implies a high number of hierarchical levels, $k = 20$. The estimates of the Lognormal parameter, $\hat{\lambda}$, and the scale factor parameter, $\hat{\sigma}$, are reported in Table 4.9. Higher $\hat{\lambda}$ in the pre-1900 era indicates a higher degree of fractality of the series in the 19th than the 20th and 21st centuries, i.e. more pronounced changes between tranquil and turbulent phases which is in harmony with the visual impression of more "spikyness" in the years 1875-1895.

4.5.2. Forecasting Results

The results of the SPA test for our three forecasting exercises for all our volatility models are reported in Tables 4.10, 4.11, 4.12, 4.13, 4.14, 4.15, 4.16, and 4.17. The first column in each table contains the benchmark models and each model is tested against the remaining eight models. It also contains individual model combination forecasts that are tested against all nine single models. The p-values of the SPA test are computed based on 5000 bootstrap samples in the empirical test under any pre-specified loss function. First, we observe that in each case of the four forecasting exercises none of our volatility models can outperform all other models at short and long horizons across all six different loss functions. The forecasting performance of our volatility models also differs from one sample period to another. Often, a volatility model that provides relatively accurate forecasts for a period of time might perform poorly in terms of forecasting performance when expanding or reducing the sample size. However, all in all it seems that the long memory volatility models are more appropriate to forecast oil price volatility. We also observe that for the more standard loss functions such as MSE or MAE, the MSM model mostly cannot be outperformed. Based on the SPA results, we count for each of our volatility models the cases where it cannot be outperformed by others across all time spans and criteria. The results indicate that in 99 cases the LMSM cannot be outperformed by its competitor models at the 10% confidence level, followed by HYGARCH (94 cases), FIAPARCH (89 cases), GARCH (74 cases), EGARCH (70 cases), IGARCH (49 cases), FIGARCH (45 cases), GJR-GARCH (42 cases), and APARCH (28 cases). Overall, the new multifractal model, therefore, appears to perform better on average than any particular model from the GARCH family. This is particularly remarkable as (i) it has fewer parameters than all GARCH-type models (i.e., only two while the second best, the HYGARCH model, comes with five parameters that have to be estimated), (ii) our estimation and forecasting methods used for the multifractal model are not the most efficient ones, while we have used the most efficient ML estimates and conditional expectations based upon those to compute forecasts for the GARCH family. Across time periods and criteria we find the following tendencies: First, the MSM and FIAPARCH do well and cannot be rejected as non-dominated models for the 19th century data and for the 2010-2014 out-of-sample period. Both do not perform well for the 2007-2009 out-of-sample period. The HYGARCH model gains its prominent rank particularly from its better performance in this period, but also other short-memory GARCH-type models do better in this period than in the others. Presumably, the higher volatility in the crisis period rewards a concentration on the short-run dynamics rather than long trends in volatility. Across criteria, the RLOG statistic is typically an outlier in its patterns of SPA results which is not surprising given the higher weight it attributes to small rather than large events.

The difficulty to discover a uniformly best model across all six different loss functions at short and long horizons motivates us to also try simple average forecast combinations. Granger and Teräsvirta (1999) and Aiolfi and Timmermann (2006) pointed out that it is often preferable to combine forecasts from competitive models in a linear way and thereby generate hopefully superior predictions. Following this idea, we adopt two different combination strategies. The first combination strategy is given by equally weighted linear combinations of short memory GARCH-type models (GARCH, IGARCH, GJR-GARCH, EGARCH and APARCH) and long memory GARCH-type and MSM models (FIGARCH, HYGARCH, FIAPARCH and LMSM). The second one is also obtained by equally weighted linear combinations of long memory GARCH-type models and the LMSM. Both combination strategies shed light on the complementarities of the short- and long-memory GARCH models on the one side and the complementarities of two classes of volatility models, GARCH-type and MSM, on the other side. In fact, both strategies lead to a high number of forecast combinations. To reduce the number of forecast combinations we only considered GARCH-type models that have the highest p-values according to the SPA test results for our single volatility models. Note that this selection criterion does not hold for the LMSM, so that we always combined the best GARCH-type models in terms of their p-values with the LMSM in order to explore their complementarities. This selection criterion for GARCH-type models led to different forecast combinations for different loss functions. The new predictor is tested against the single models and the test results are reported in Tables 4.10, 4.11, 4.12, 4.13, 4.14, 4.15, 4.16, and 4.17. The results are diverse: First, one often observes that forecast combinations of two relatively successful models do not necessarily improve performance against single models. This holds particularly for combinations of short-memory GARCH specifications. Combinations of long-memory GARCH models with the MSM model are more often successful, but we nevertheless find cases where the combination of well performing single models can be outperformed by the forecasts from one or more of those single models. This exercise underlines that forecast combination is a delicate operation: There is apparently no guarantee that two good models are complementary in their virtues, they could also lead to an overall deterioration when applied in combination. This underscores the necessity of finding more elaborate rules for combinations that are data-driven and react on the single models' advantages and disadvantages.

4.6. Conclusion

This chapter has analyzed the forecasting performance of two classes of volatility models, namely the GARCH-type models and the MSM model via six different loss functions and the superior predictive ability test. The analysis is performed by using a large sample of oil prices of the pre- and post-1900 period. Results were largely uniform for the data of the 19th century and the later record of the 20th/21st centuries with the crisis period 2007 - 2009 showing somewhat unusual behavior. Empirical results of the SPA test indicate that none of the volatility models including the MSM model can outperform their competitor models under all loss criteria. As it turned out, however, the new MSM model most often cannot be outperformed when standard loss functions are used. Across all forecasting horizons and subsamples used, it is the model that in the highest number of cases cannot be outperformed by any other models, and, in this respect, it beats all simple models from our broad selection of GARCH-type processes. Forecast combination exercises point to more robustness of combinations of long-memory GARCH and MSM models rather than short-memory GARCH models. However, superior forecast performance of combined models against their single components is in no way guaranteed.

All in all, the MSM model appears a valuable addition to the toolbox of volatility models not only for financial assets, but also for commodities like oil. Given its highest number of non-rejections by the SPA test, it comes out as the more robust model compared to any GARCH specification, and it also is the most parsimonious one among all candidates considered.

Table 4.1.: Descriptive statistics of the data pre-1900

	Log-returns	Absolute returns	Squared returns
6376 observations (from January 02,1875 to Decem 31, 1895)			
Minimum	-16.186	0	0
Maximum	33.647	33.647	1.132E+3
Mean	-0.007	1.439	5.129
Standard deviation	2.265	1.749	21.291
Skewness	0.752	3.715	29.944
Kurtosis	18.240	34.497	1.460E+3
Hurst index	0.540	0.842***	0.868***
Hill tail index at 5% tail	2.547 **[2.485 2.610]**		
Q(10)	79.177	2.467E+3	659.798
Q(20)	100.685	3.169E+3	712.731
JB	5.343E+4		
ADF	- 73.875		

Note: *** indicates 1% significance of Hurst coefficients based on the simulated boundary values of Weron (2002) for Wiener Brownian motion. For the tail index estimates, the brackets contain the 95% percent confidence intervals of the point estimate based upon the limiting distribution of the estimator.

Table 4.2.: Descriptive statistics of the data post-1900 containing oil prices from Jan 06,1992 to December 31, 2009

	Log-returns	Absolute returns	Squared returns
4521 observations (from January 06,1992 to December 31, 2009)			
Minimum	-17.092	0	0
Maximum	16.414	17.092	292.129
Mean	0.031	1.748	6.078
Standard deviation	2.466	1.739	16.326
Skewness	-0.154	2.700	8.361
Kurtosis	8.222	15.256	99.210
Hurst index	0.490	0.856***	0.905***
Hill tail index at 5% tail	2.797 **[2.716 2.879]**		
Q(10)	31.015	1099.7	875.701
Q(20)	42.686	2051.1	1556.1
JB	5155.4		
ADF	- 68.396		

Note: *** indicates 1% significance of Hurst coefficients based on the simulated boundary values of Weron (2002) for Wiener Brownian motion. For the tail index estimates, the brackets contain the 95% percent confidence intervals of the point estimate based upon the limiting distribution of the estimator.

Table 4.3.: Descriptive statistics of the data post-1900 containing oil prices from Jan 06,1992 to March 24, 2014

	Log-returns	Absolute returns	Squared returns
5590 observations (from January 06,1992 to March 24, 2014)			
Minimum	-17.092	0	0
Maximum	16.414	17.092	292.129
Mean	0.030	1.654	5.482
Standard deviation	2.341	1.657	15.030
Skewness	-0.145	2.753	8.889
Kurtosis	8.525	15.993	113.634
Hurst index	0.471	0.868***	0.910***
Hill tail index at 5% tail	2.899 **[2.823 3.975]**		
Q(10)	31.447	1.419E+3	1.127E+3
Q(20)	41.038	2.647E+3	1.999E+3
JB	7.129E+3		
ADF	-75.996		

Note: *** indicates 1% significance of Hurst coefficients based on the simulated boundary values of Weron (2002) for Wiener Brownian motion. For the tail index estimates, the brackets contain the 95% percent confidence intervals of the point estimate based upon the limiting distribution of the estimator.

Table 4.4.: Descriptive statistics of the complete data post-1900

	Log-returns	Absolute returns	Squared returns
9417 observations (from January 03,1977 to March 24, 2014)			
Minimum	-40.204	0	0
Maximum	19.861	40.204	1614.4
Mean	0.021	1.363	4.922
Standard deviation	2.219	1.751	22.929
Skewness	-0.832	3.962	40.176
Kurtosis	22.738	42.191	2624.4
Hurst index	0.517	0.938***	0.944***
Hill tail index at 5% tail	2.668 **[2.614 2.722]**		
Q(10)	52.030	7.804E+3	1.458E+3
Q(20)	75.992	1.458E+4	1.681E+3
JB	1.540E+5		
ADF	-98.326		

Note: *** indicates 1% significance of Hurst coefficients based on the simulated boundary values of Weron (2002) for Wiener Brownian motion. For the tail index estimates, the brackets contain the 95% percent confidence intervals of the point estimate based upon the limiting distribution of the estimator.

Table 4.5.: Estimation results using oil prices from January 2, 1875 to December 31, 1895

	GARCH	IGARCH	EGARCH	GJR-GARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH
ω	0.066	0.065	0.083	0.066	0.074	0.301	0.277	0.224
	(0.033)	**(0.029)**	**(0.016)**	**(0.035)**	**(0.052)**	**(0.040)**	**(0.054)**	**(0.051)**
α	0.127	0.128	0.007	0.127	0.131			
	(0.027)	**(0.033)**	**(0.012)**	**(0.028)**	**(0.029)**			
β	0.871	0.872	0.961	0.872	0.869	0.550	0.550	0.524
	(0.030)	**(0.033)**	**(0.011)**	**(0.031)**	**(0.038)**	**(0.080)**	**(0.078)**	**(0.097)**
γ			0.302	0.001	0.002			-0.239
			(0.037)	**(0.003)**	**(0.001)**			**(0.062)**
δ					1.748			1.315
					(0.776)			**(0.114)**
ϕ						0.254	0.255	0.302
						(0.052)	**(0.051)**	**(0.056)**
d						0.493	0.489	0.396
						(0.049)	**(0.047)**	**(0.062)**
τ							1.012	
							(0.019)	
				Diagnostic				
Log(L)	-12806	-12806	-12757	-12806	-12802	-12749	-12749	-12729
AIC	25618	25618	25522	25620	25614	25506	25508	25470
BIC	25638	2538	25550	25647	25648	25533	25541	25511
$Q(20)$	37.248	37.240	34.517	37.263	37.349	32.905	32.782	32.309
	[0.011]	**[0.011]**	**[0.023]**	**[0.011]**	**[0.011]**	**[0.035]**	**[0.036]**	**[0.040]**
$Q^2(20)$	28.591	28.561	30.376	28.585	31.990	16.956	16.862	22.671
	[0.090]	**[0.097]**	**[0.064]**	**[0.096]**	**[0.043]**	**[0.656]**	**[0.662]**	**[0.305]**
Arch(20)	27.544	27.525	29.137	27.536	30.653	16.490	16.395	22.105
	[0.121]	**[0.121]**	**[0.085]**	**[0.121]**	**[0.060]**	**[0.686]**	**[0.692]**	**[0.335]**

Note: The numbers in parentheses are standard errors of the estimations. Log(L) is the logarithm maximum likelihood function. AIC and BIC are the Akaike and Bayesian information criterion respectively. $Q(20)$ and $Q^2(20)$ are the Ljung-Box Q-statistics of order 20 obtained from the standardized residuals and squared standardized residuals respectively. ARCH(20) denotes the no conditional heteroscedasticity statistic of order 20. The values reported in square brackets are the p-values of the statistics.

Table 4.6.: Estimation results using oil prices from January 06,1992 to December 31, 2009

	GARCH	IGARCH	EGARCH	GJR-GARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH
ω	0.052	0.033	0.027	0.051	0.046	0.569	0.255	0.197
	(0.022)	(0.017)	(0.016)	(0.022)	(0.026)	(0.132)	(0.106)	(0.186)
α	0.064	0.068	-0.005	0.068	0.072			
	(0.017)	(0.019)	(0.008)	(0.029)	(0.019)			
β	0.929	0.932	0.988	0.930	0.928	0.469	0.414	0.382
	(0.017)	(0.019)	(0.028)	(0.019)	(0.002)	(0.083)	(0.026)	(0.090)
γ			0.156	-0.008	0.003			-0.132
			(0.019)	(0.028)	(0.021)			(0.068)
δ					1.63			1.889
					(0.241)			(0.175)
ϕ						0.214	0.204	0.211
						(0.092)	(0.056)	(0.072)
d						0.364	0.290	0.261
						(0.046)	(0.059)	(0.052)
τ							1.112	
							(0.281)	
				Diagnostic				
Log(L)	-10024	-10026	-10022	-10023	-10020	-10014	-10013	-10008
AIC	20054	20056	20052	20055	20054	20037	20035	20029
BIC	20073	20069	20078	20081	20082	20062	20068	20067
$Q(20)$	19.655	19.418	21.498	19.502	20.389	22.225	22.506	22.991
	[0.480]	[0.495]	[0.368]	[0.489]	[0.434]	[0.328]	[0.314]	[0.289]
$Q^2(20)$	43.085	42.259	49.062	43.338	45.733	30.682	29.907	30.403
	[0.002]	[0.003]	[<0.001]	[0.002]	[<0.001]	[0.060]	[0.071]	[0.064]
Arch(20)	37.525	36.912	41.822	37.727	39.417	28.478	27.943	28.193
	[0.010]	[0.012]	[0.003]	[0.010]	[0.006]	[0.099]	[0.111]	[0.105]

Note: The numbers in parentheses are standard errors of the estimations. Log(L) is the logarithm maximum likelihood function. AIC and BIC are the Akaike and Bayesian information criterion respectively. $Q(20)$ and $Q^2(20)$ are the Ljung-Box Q-statistics of order 20 obtained from the standardized residuals and squared standardized residuals respectively. ARCH(20) denotes the no conditional heteroscedasticity statistic of order 20. The values reported in square brackets are the p-values of the statistics.

Table 4.7.: Estimation results oil prices from from January 06,1992 to March 24, 2014

	GARCH	IGARCH	EGARCH	GJR-GARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH
ω	0.041	0.028	0.022	0.041	0.039	0.431	0.149	4.982E-5
	(0.020)	**(0.013)**	**(0.010)**	**(0.019)**	**(0.071)**	**(0.105)**	**(0.153)**	**(3.558E−5)**
α	0.063	0.067	0.144	0.057	0.069			
	(0.017)	**(0.017)**	**(0.038)**	**(0.021)**	**(0.034)**			
β	0.931	0.933	0.989	0.932	931	0.491	0.431	0.135
	(0.018)	**(0.017)**	**(0.005)**	**(0.017)**	**(0.053)**	**(0.138)**	**(0.124)**	**(0.049)**
γ			0.221	0.011	-0.017			-0.999
			(0.026)	**(0.020)**	**(0.009)**			**(0.312)**
δ					1.610			1.574
					(0.817)			**(0.195)**
ϕ						0.251	0.238	0.111
						(0.132)	**(0.127)**	**(0.041)**
d						0.360	0.285	0.070
						(0.044)	**(0.054)**	**(0.020)**
τ							1.113	
							(0.069)	
				Diagnostic				
Log(L)	-12074	-12076	-12069	-12073	-12068	-12061	-12059	-12069
AIC	24154	24156	24147	24154	24146	24130	24128	24149
BIC	24174	24170	24173	24181	24179	24157	24161	24189
$Q(20)$	17.345	17.195	19.311	17.682	18.507	20.055	20.352	22.334
	[0.631]	**[0.640]**	**[0.502]**	**[0.608]**	**[0.554]**	**[0.455]**	**[0.436]**	**[0.323]**
$Q^2(20)$	47.267	46.878	55.311	47.102	50.154	29.610	28.514	36.200
	[<0.001]	**[<0.001]**	**[<0.001]**	**[<0.001]**	**[<0.001]**	**[0.076]**	**[0.098]**	**[0.015]**
Arch(20)	41.457	41.263	47.483	41.304	43.473	27.511	26.755	33.973
	[0.003]	**[0.003]**	**[<0.001]**	**[0.003]**	**[0.001]**	**[0.122]**	**[0.142]**	**[0.026]**

Note: The numbers in parentheses are standard errors of the estimations. Log(L) is the logarithm maximum likelihood function. AIC and BIC are the Akaike and Bayesian information criterion respectively. $Q(20)$ and $Q^2(20)$ are the Ljung-Box Q-statistics of order 20 obtained from the standardized residuals and squared standardized residuals respectively. ARCH(20) denotes the no conditional heteroscedasticity statistic of order 20. The values reported in square brackets are the p-values of the statistics.

Table 4.8.: Estimation results oil prices from from January 03,1977 to March 24, 2014

	GARCH	IGARCH	EGARCH	GJR-GARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH
ω	3.852E-4	3.846E-4	0.031	3.791E-4	2.098E-4	0.004	0.004	0.011
	(1.986E−4)	**(1.946E−4)**	**(0.006)**	**(1.949E−4)**	**(4.187E−4)**	**(0.002)**	**(0.002)**	**(0.013)**
α	0.079	0.079	0.221	0.074	0.088			
	(0.008)	**(0.009)**	**(0.027)**	**(0.015)**	**(0.017)**			
β	0.921	0.921	0.987	0.921	912	0.924	0.914	0.921
	(0.009)	**(0.009)**	**(0.004)**	**(0.009)**	**(0.033)**	**(0.014)**	**(0.018)**	**(0.016)**
γ			-0.016	0.010	-0.029			-0.203
			(0.015)	**(0.019)**	**(0.268)**			**(0.016)**
δ					2.250			1.591
					(0.893)			**(0.097)**
ϕ						4.632E-8	3.123E-8	1.898E-7
						(1.405E−8)	**(1.173E−8)**	**(9.175E−7)**
d						1.000	1.000	1.00
						(0.023)	**(0.038)**	**(0.031)**
τ							1.017	
							(0.006)	
				Diagnostic				
Log(L)	-16182	-16180	-15963	-16180	-16131	-16114	-16047	-16030
AIC	32370	32365	31933	32368	32273	32236	32103	32072
BIC	32392	32379	31962	32397	32308	32264	32139	32115
$Q(20)$	177.286	177.450	132.202	179.483	177.855	187.193	193.544	204.216
	[0.000]	**[0.000]**	**[0.000]**	**[0.000]**	**[0.000]**	**[0.000]**	**[0.000]**	**[0.000]**
$Q^2(20)$	12.904	12.866	7.523	13.313	12.836	11.969	9.567	8.820
	[0.882]	**[0.883]**	**[0.995]**	**[0.864]**	**[0.884]**	**[0.917]**	**[0.974]**	**[0.984]**
Arch(20)	12.708	12.672	7.530	13.110	13.073	11.764	9.807	8.824
	[0.890]	**[0.891]**	**[0.994]**	**[0.873]**	**[0.874]**	**[0.924]**	**[0.972]**	**[0.985]**

Note: The numbers in parentheses are standard errors of the estimations. Log(L) is the logarithm maximum likelihood function. AIC and BIC are the Akaike and Bayesian information criterion respectively. $Q(20)$ and $Q^2(20)$ are the Ljung-Box Q-statistics of order 20 obtained from the standardized residuals and squared standardized residuals respectively. ARCH(20) denotes the no conditional heteroscedasticity statistic of order 20. The values reported in square brackets are the p-values of the statistics.

Table 4.9.: Estimation results of LMSM model

Parameters	Jan 2, 1875 to Dec 31, 1895	Jan 6, 1992 to Dec 31, 2009	Jan 6, 1992 to March 24, 2014	Jan 3, 1977 to March 24, 2014
λ	1.320	1.016	1.034	1.011
σ	2.252	2.465	2.341	2.218

Note that the optimal objective function of the GMM estimation is obtained for $k = 20$.

Table 4.10.: Superior predictive ability (SPA) test results using oil price observations from January 3, 1875 to December 31, 1892 as in-sample and from January 3, 1893 to December 31, 1895 as out-of-sample.

Base model	Forecast horizons											
	1	5	10	20	30	40	50	60	70	80	90	100
							MSE					
GARCH	0.020	0.062	**0.156**	0.090	0.070	0.040	0.074	0.046	**0.134**	**0.432**	0.012	0.004
IGARCH	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.008	0.032	0.016	0.022	0.084	0.058	**0.338**	0.080	**0.226**	**0.226**	**0.604**	**0.632**
EGARCH	0.020	**0.118**	**0.740**	**0.644**	**0.656**	**0.688**	**1.000**	**0.766**	**0.980**	**0.752**	**0.578**	**0.170**
APARCH	0.004	0.014	0.002	0.004	0.018	0.014	**0.542**	0.032	**0.234**	**0.400**	**0.928**	**0.836**
FIGARCH	0.012	0.092	**0.700**	**0.554**	0.040	0.016	0.076	0.066	0.088	0.018	0.004	0.000
HYGARCH	**1.000**	**0.380**	**0.148**	0.032	0.024	0.038	0.038	0.034	0.038	0.026	0.020	0.014
FIAPARCH	0.044	**0.732**	**0.602**	**0.450**	**0.176**	0.052	0.078	0.050	0.050	0.028	0.022	0.018
LMSM	0.042	**0.170**	**0.880**	**0.924**	**0.404**	**0.312**	**0.316**	**0.270**	**0.240**	**0.204**	**0.130**	**0.126**
FCOM1	0.004	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM11	0.026	0.202	**1.000**	**0.931**	**0.992**	**1.000**	**0.846**	**0.305**	**0.348**	**0.305**	**0.185**	**0.139**
FCOM111	0.049	**0.216**	**0.999**	**0.864**	**0.468**	0.086	**0.209**	**0.127**	**0.153**	**0.231**	0.047	0.012
							MAE					
GARCH	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.002	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	**1.000**	**0.512**	**0.360**	0.098	0.060	0.066	0.058	0.046	0.052	0.080	0.062	0.080
FIAPARCH	0.020	**0.488**	**0.984**	**1.000**	**0.574**	**0.292**	**0.344**	**0.276**	**0.192**	**0.170**	**0.142**	**0.144**
LMSM	0.000	0.018	**0.250**	0.056	**0.486**	**0.708**	**0.656**	**0.724**	**0.808**	**1.000**	**1.000**	**1.000**
FCOM2	0.000	0.000	0.005	0.000	0.007	0.048	0.015	0.062	**0.131**	**0.102**	0.073	0.076
FCOM21	0.000	0.045	**0.460**	**0.207**	**0.725**	**0.485**	**0.988**	**0.505**	**0.355**	**0.222**	**0.198**	**0.193**
							HMSE					
GARCH	**0.102**	**0.118**	0.070	0.058	0.042	0.062	0.034	0.054	0.014	0.022	0.038	0.032
IGARCH	0.024	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
GJR-GARCH	0.020	0.032	0.066	0.046	0.008	0.022	0.010	0.036	0.022	0.016	0.032	0.018
EGARCH	0.008	0.056	0.042	0.038	0.018	0.032	0.024	0.028	0.020	0.016	0.028	0.024
APARCH	0.008	0.036	0.076	0.050	0.014	0.030	0.018	0.040	0.008	0.020	0.034	0.030
FIGARCH	0.084	**0.130**	0.072	0.058	0.032	0.062	0.032	0.054	0.020	0.022	0.042	0.030
HYGARCH	**1.000**	0.048	0.082	0.064	0.022	0.088	0.010	0.050	0.006	0.034	0.058	0.030
FIAPARCH	0.040	**0.112**	0.072	0.052	0.014	0.052	0.008	0.026	0.020	0.030	0.056	0.014
LMSM	0.068	0.040	0.032	0.036	0.026	0.050	0.036	0.046	0.022	0.016	0.034	0.030
FCOM3	0.007	0.232	0.079	0.061	0.033	0.053	0.036	0.047	0.032	0.035	0.041	0.021

Note: The table entries are the p-values of the SPA test of Hansen (2005). The null hypothesis is that a base model cannot be outperformed by other competitor models. The values in bold face represent the p-values that are greater than or equal to the 10% confidence level under a pre-specified loss function. We combine: FCOM1=EGARCH+APARCH, FCOM11=EGARCH+LMSM, FCOM111=FIAPARCH+LMSM, FCOM2=HYGARCH+FIAPARCH+LMSM, FCOM21=FIAPARCH+LMSM, and FCOM3=IGARCH+HYGARCH+ LMSM.

Table 4.11.: Superior predictive ability (SPA) test results using oil price observations from January 3, 1875 to December 31, 1892 as in-sample and from January 3, 1893 to December 31, 1895 as out-of-sample.

Base model	Forecast horizons											
	1	5	10	20	30	40	50	60	70	80	90	100
						HMAE						
GARCH	0.012	0.060	0.046	0.046	0.018	0.044	0.012	0.020	0.004	0.012	0.006	0.006
IGARCH	0.006	**1.000**	**1.000**	**1.000**	**1.000**	0.016	0.008	0.012	0.000	0.002	0.012	0.002
GJR-GARCH	0.010	0.020	0.040	0.034	0.012	0.014	0.010	0.020	0.010	0.014	0.022	0.016
EGARCH	0.006	0.006	0.006	0.006	0.005	0.030	0.012	0.006	0.004	0.004	0.006	0.002
APARCH	0.006	0.008	0.020	0.016	0.006	0.026	0.008	0.012	0.004	0.004	0.022	0.008
FIGARCH	0.002	0.064	0.042	0.038	0.006	0.028	0.008	0.014	0.004	0.006	0.014	0.002
HYGARCH	**1.000**	0.010	0.018	0.010	0.000	0.026	0.000	0.004	0.000	0.002	0.004	0.000
FIAPARCH	0.050	0.094	0.042	0.020	0.004	0.018	0.002	0.002	0.000	0.002	0.004	0.000
LMSM	0.002	0.008	0.006	0.006	0.004	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
FCOM4	0.000	**0.695**	0.089	0.058	0.017	0.025	0.011	0.031	0.013	0.018	0.019	0.000
FCOM41	0.000	**0.165**	0.056	0.044	0.015	0.024	0.007	0.017	0.004	0.007	0.012	0.000
						QLIKE						
GARCH	0.000	**0.566**	**0.960**	**0.846**	**0.160**	0.032	**0.260**	**0.154**	**0.362**	**0.500**	0.066	0.024
IGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.006	**0.540**	**0.606**	**0.778**	**0.664**	**0.508**	**0.546**	**0.650**	**0.870**
EGARCH	0.000	0.060	**0.520**	**0.250**	**0.992**	**0.962**	**0.570**	**0.846**	**0.714**	**0.878**	**0.432**	**0.162**
APARCH	0.000	0.000	0.000	0.004	**0.600**	**0.524**	**0.402**	**0.308**	**0.632**	**0.604**	**0.524**	**0.302**
FIGARCH	0.000	**0.518**	**0.220**	**0.172**	0.072	0.052	0.014	0.022	0.006	0.014	0.022	0.000
HYGARCH	**1.000**	**0.786**	**0.224**	**0.128**	0.062	0.062	0.022	0.030	0.010	0.014	0.024	0.004
FIAPARCH	0.000	**0.156**	**0.146**	0.082	0.042	0.042	0.018	0.024	0.012	0.014	0.014	0.004
LMSM	0.000	**0.164**	**0.930**	**0.616**	**0.396**	**0.366**	**0.344**	**0.238**	**0.192**	**0.128**	0.082	0.076
FCOM5	0.000	0.002	**0.893**	**0.322**	**0.997**	**0.676**	**0.742**	**0.666**	**0.758**	**0.600**	**0.324**	0.075
FCOM51	0.000	0.000	**0.939**	**0.275**	**0.775**	**0.997**	**0.671**	**0.698**	**0.554**	**0.462**	**0.347**	**0.239**
FCOM511	0.000	0.001	**0.793**	**0.382**	**1.000**	**1.000**	**1.000**	**0.540**	**0.619**	**0.593**	**0.560**	**0.330**
FCOM5111	0.000	**0.494**	**0.870**	**0.676**	**0.455**	**0.143**	**0.247**	**0.100**	**0.155**	0.088	0.074	0.038
						RLOG						
GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	0.000	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
FIAPARCH	**0.834**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
LMSM	0.023	**0.902**	**0.453**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM61	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM611	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: The table entries are the p-values of the SPA test of Hansen (2005). The null hypothesis is that a base model cannot be outperformed by other competitor models. The values in bold face represent the p-values that are greater than or equal to the 10% confidence level under a pre-specified loss function. We combine: FCOM4=IGARCH+HYGARCH+LMSM, FCOM41=HYGARCH+LMSM, FCOM5=GARCH+GJR-GARCH+EGARCH, FCOM51=GJR-GARCH+EGARCH+HYGARCH, FCOM511=EGARCH+GJR-GARCH+LMSM, FCOM5111=HYGARCH+LMSM, FCOM6=GARCH+HYGARCH, FCOM61=HYGARCH+FIAPARCH, and FCOM611=HYGARCH+LMSM.

Table 4.12.: Superior predictive ability (SPA) test results using oil price observations from January 6, 1992 to December 29, 2006 as in-sample and from January 2, 2007 to December 31, 2009 as out-of-sample.

Base model	Forecast horizons											
	1	5	10	20	30	40	50	60	70	80	90	100
						MSE						
GARCH	0.002	**0.282**	**0.780**	**0.820**	**0.294**	**0.532**	**0.410**	**0.894**	**0.302**	**0.314**	**0.100**	**0.100**
IGARCH	0.002	0.004	0.002	0.004	0.000	0.004	0.016	0.088	**0.960**	**0.938**	**0.550**	**0.828**
GJR-GARCH	0.004	0.008	0.004	0.002	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.002	0.000	0.004	0.006	0.000	0.006	0.010	0.036	**0.128**	**0.260**	**0.626**	**0.172**
APARCH	0.000	0.002	0.002	0.004	0.000	0.002	0.000	0.000	0.008	0.002	0.004	0.040
FIGARCH	0.000	**0.122**	0.020	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	**1.000**	**0.838**	**0.332**	**0.222**	**0.706**	**0.536**	**0.654**	**0.170**	**0.294**	0.010	0.000	0.000
FIAPARCH	0.000	0.006	0.004	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.000
LMSM	0.002	0.080	0.018	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM1	0.003	0.068	0.027	0.042	0.012	0.014	0.089	0.006	0.000	0.000	0.001	0.002
FCOM11	0.002	**0.235**	**0.111**	**0.124**	**0.383**	**0.767**	**0.573**	**0.286**	0.004	0.000	0.000	0.000
FCOM111	0.003	**0.310**	**0.500**	**0.670**	**0.636**	**0.993**	**0.994**	**0.302**	0.012	0.000	0.000	0.000
						MAE						
GARCH	0.000	0.016	0.016	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	**0.826**	**0.680**	**0.268**	**0.810**	**0.238**	**0.180**	0.002	0.002	0.000	0.000	0.000
GJR-GARCH	0.000	**0.596**	**0.420**	**0.114**	**0.930**	**0.970**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
EGARCH	0.000	0.074	**0.254**	**0.798**	**0.786**	**0.444**	0.064	0.012	0.000	0.000	0.000	0.000
APARCH	0.000	**0.202**	**0.760**	**0.184**	**0.310**	0.058	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.006	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	**0.124**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIAPARCH	**0.902**	0.084	0.054	0.028	**0.518**	**0.144**	0.078	0.056	0.038	**0.126**	0.036	**0.116**
LMSM	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM2	0.002	**0.342**	0.002	**0.194**	**0.858**	**0.706**	0.086	0.015	0.000	0.000	0.000	0.000
FCOM21	0.002	**0.588**	**0.166**	0.017	**0.909**	**0.394**	**0.255**	0.016	0.009	0.000	0.000	0.000
FCOM211	0.025	**0.456**	**0.297**	0.043	**0.868**	**0.640**	**0.426**	**0.333**	**0.252**	**0.414**	**0.179**	**0.382**
						HMSE						
GARCH	0.038	**0.152**	**0.796**	**0.104**	**1.000**	**0.854**	**0.148**	**0.806**	**0.390**	**0.282**	0.086	0.000
IGARCH	0.000	0.002	0.002	0.000	0.000	0.022	0.018	**0.104**	**0.308**	**0.874**	**0.408**	**0.734**
GJR-GARCH	0.000	0.002	0.002	0.000	0.000	0.004	0.000	0.008	0.006	0.000	0.002	0.002
EGARCH	0.000	0.002	0.002	0.000	0.000	0.000	0.010	0.030	0.036	**0.512**	**0.592**	**0.266**
APARCH	0.000	0.000	0.006	0.000	0.000	0.002	0.006	0.012	0.008	0.010	0.008	0.004
FIGARCH	0.008	**0.128**	0.096	0.010	0.054	0.028	0.020	0.024	0.026	0.008	0.006	0.000
HYGARCH	**1.000**	**1.000**	**0.204**	**1.000**	**0.108**	**0.146**	**0.852**	**0.240**	**0.874**	**0.144**	0.056	0.020
FIAPARCH	0.000	0.086	0.070	0.018	0.040	0.028	0.002	0.000	0.002	0.010	0.000	0.010
LMSM	0.078	**0.122**	0.020	0.026	0.050	0.034	0.022	0.016	0.034	0.018	0.012	0.006
FCOM3	0.000	0.001	0.015	0.002	0.011	0.033	**0.175**	0.032	0.007	0.011	0.016	0.007
FCOM31	0.000	0.027	0.057	0.031	0.022	0.064	**0.831**	**0.158**	0.013	0.020	0.019	0.012

Note: The table entries are the p-values of the SPA test of Hansen (2005). The null hypothesis is that a base model cannot be outperformed by other competitor models. The values in bold face represent the p-values that are greater than or equal to the 10% confidence level under a pre-specified loss function. We combine: FCOM1=GARCH+IGARCH, FCOM11=GARCH+IGARCH+HYGARCH, FCOM111=HYGARCH+LMSM, FCOM2=GJR-GARCH+EGARCH, FCOM21=GJR-GARCH+IGARCH, FCOM211=GJR-GARCH+FIAPARCH, FCOM3=GARCH+IGARCH and
FCOM31=GARCH+IGARCH+HYGARCH.

Table 4.13.: Superior predictive ability (SPA) test results using oil price observations from January 6, 1992 to December 29, 2006 as in-sample and from January 2, 2007 to December 31, 2009 as out-of-sample.

Base model	Forecast horizons											
	1	5	10	20	30	40	50	60	70	80	90	100
							HMAE					
GARCH	0.000	**0.136**	**0.292**	0.062	**0.594**	**0.490**	0.060	**0.518**	**0.216**	**0.758**	0.068	0.000
IGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.038	**0.666**	**0.798**	**1.000**
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	**0.144**	**0.312**	0.026
APARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.030	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	**1.000**	**1.000**	**0.708**	**1.000**	**0.406**	**0.510**	**1.000**	**0.482**	**0.784**	**0.412**	0.010	0.000
FIAPARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
LMSM	0.000	0.016	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM4	0.000	**0.302**	**0.721**	**0.216**	**0.837**	**1.000**	**0.517**	**0.167**	0.000	0.000	0.000	0.000
FCOM41	0.000	0.018	0.002	0.003	0.005	0.003	0.001	0.012	0.001	0.000	0.000	0.000
							QLIKE					
GARCH	0.000	**0.286**	**0.662**	**0.158**	**0.808**	**0.764**	**0.116**	**0.772**	**0.546**	0.068	0.010	0.002
IGARCH	0.000	0.000	0.000	0.000	0.000	0.002	0.014	0.098	**0.752**	**1.000**	**0.560**	**1.000**
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.034	**0.138**	**0.238**	**0.440**	0.084
APARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	**0.144**	0.022	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	**1.000**	**0.790**	**0.338**	**0.842**	**0.192**	**0.236**	**0.884**	**0.316**	**0.564**	0.076	0.000	0.000
FIAPARCH	0.000	0.004	0.002	0.002	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
LMSM	0.000	0.022	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM5	0.000	0.034	0.011	0.026	0.013	0.015	0.090	0.005	0.000	0.000	0.001	0.001
FCOM51	0.000	**0.379**	**0.995**	**0.360**	**0.532**	**0.641**	**0.925**	0.054	0.000	0.001	0.002	0.000
FCOM511	0.000	**0.170**	**0.129**	**0.131**	**0.167**	**0.141**	0.025	0.043	0.001	0.001	0.002	0.000
							RLOG					
GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.002
IGARCH	0.002	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.002	0.014	0.006	0.002	0.004	0.016	0.012	0.012	0.008
EGARCH	**0.742**	**0.882**	**1.000**	**1.000**	0.082	0.004	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	**0.458**	**0.118**	0.008	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIAPARCH	**0.444**	0.000	0.000	0.080	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
LMSM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM6	**1.000**	0.006	0.031	**0.200**	**0.180**	0.016	0.002	0.002	0.006	0.006	0.002	0.000

Note: The table entries are the p-values of the SPA test of Hansen (2005). The null hypothesis is that a base model cannot be outperformed by other competitor models. The values in bold face represent the p-values that are greater than or equal to the 10% confidence level under a pre-specified loss function. We combine: FCOM4=GARCH+HYGARCH, FCOM41=HYGARCH+LMSM, FCOM5=GARCH+IGARCH, FCOM51=GARCH+HYGARCH, FCOM511=HYGARCH+LMSM and FCOM6=EGARCH+FIAPARCH.

Table 4.14.: Superior predictive ability (SPA) test results using oil price observations from January 6, 1992 to December 31, 2009 as in-sample and from January 4, 2010 to March 24, 2014 as out-of-sample.

Base model	Forecast horizons											
	1	5	10	20	30	40	50	60	70	80	90	100
							MSE					
GARCH	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.004	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.000	0.020	0.040	0.016	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.002	0.024	0.002	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000
HYGARCH	**1.000**	0.032	0.000	0.012	0.008	**0.100**	0.094	0.076	0.086	0.002	0.042	0.076
FIAPARCH	0.018	**0.106**	**0.160**	**0.186**	**0.196**	**0.370**	**0.694**	**0.526**	**0.728**	**0.754**	**0.864**	**1.000**
LMSM	0.004	**1.000**	**1.000**	**0.866**	**0.872**	**0.704**	**0.306**	**0.558**	**0.334**	**0.246**	**0.136**	0.082
FCOM1	**0.104**	**0.495**	**0.400**	**0.387**	**0.484**	**0.824**	**0.852**	**0.850**	**0.997**	**0.967**	**0.678**	**0.716**
FCOM11	0.018	**0.652**	**0.606**	**0.695**	**0.808**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
							MAE					
GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.007	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.008	0.006
HYGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.011
FIAPARCH	**1.000**	0.000	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	0.000	0.000	0.000
LMSM	0.000	**1.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**1.000**	**0.876**	**0.901**
FCOM2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM21	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM211	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							HMSE					
GARCH	0.054	0.068	0.008	**0.178**	**0.194**	**0.560**	**0.450**	**0.314**	**0.104**	0.098	0.072	**0.110**
IGARCH	0.004	0.002	0.004	0.008	0.008	0.004	0.010	0.018	0.022	0.012	0.044	0.020
GJR-GARCH	0.004	0.002	0.004	0.006	0.006	0.008	0.008	0.014	0.004	0.008	0.018	0.008
EGARCH	0.006	0.002	0.006	0.012	0.004	0.028	0.032	**0.720**	**1.000**	**1.000**	**1.000**	**1.000**
APARCH	0.006	0.004	0.004	0.008	0.004	0.014	0.006	0.010	0.000	0.026	0.054	0.066
FIGARCH	0.052	**0.194**	**1.000**	**0.908**	**0.826**	**0.440**	**0.762**	**0.156**	0.036	0.052	0.086	0.060
HYGARCH	**1.000**	**0.888**	0.040	**0.124**	0.076	0.064	0.090	0.032	0.026	0.024	0.046	0.058
FIAPARCH	0.004	0.032	0.022	0.014	0.008	0.010	0.000	0.004	0.000	0.000	0.000	0.002
LMSM	0.012	**0.212**	0.088	0.018	0.044	0.084	0.076	0.054	0.056	0.036	0.054	0.038
FCOM3	0.035	0.005	0.014	0.013	0.005	0.097	**0.134**	**0.995**	0.067	0.066	0.086	0.084
FCOM31	0.039	0.006	0.016	0.019	0.012	**0.130**	0.092	**0.957**	0.051	0.061	0.078	0.083
FCOM311	0.002	**0.482**	0.068	**0.134**	0.076	**0.270**	**0.111**	0.040	0.083	0.065	0.072	0.085

Note: The table entries are the p-values of the SPA test of Hansen (2005). The null hypothesis is that a base model cannot be outperformed by other competitor models. The values in bold face represent the p-values that are greater than or equal to the 10% confidence level under a pre-specified loss function. We combine: FCOM1=HYGARCH+FIAPARCH, FCOM11=FIAPARCH+LMSM, FCOM2=GARCH+FIAPARCH, FCOM21=GARCH+LMSM, FCOM211=FIAPARCH+LMSM, FCOM3=GARCH+EGARCH, FCOM31=EGARCH +FIGARCH and FCOM311=FIGARCH+HYGARCH.

Table 4.15.: Superior predictive ability (SPA) test results using oil price observations from January 6, 1992 to December 31, 2009 as in-sample and from January 4, 2010 to March 24, 2014 as out-of-sample.

Base model	Forecast horizons											
	1	5	10	20	30	40	50	60	70	80	90	100
						HMAE						
GARCH	0.000	0.026	0.010	**0.124**	**0.116**	0.074	**0.200**	**0.488**	**0.170**	**0.200**	**0.314**	**0.242**
IGARCH	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.002	0.000	0.002	0.000	0.008	**0.186**	**0.330**	**0.280**	**0.308**
APARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.004
FIGARCH	0.000	0.034	**1.000**	**1.000**	**0.884**	**1.000**	**0.800**	**0.512**	**0.854**	**0.752**	**0.896**	**0.890**
HYGARCH	**1.000**	**1.000**	**0.132**	0.038	0.008	0.006	0.004	0.000	0.002	0.000	0.002	0.014
FIAPARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
LMSM	0.000	0.038	0.000	0.000	0.002	0.004	0.004	0.002	0.000	0.004	0.002	0.008
FCOM4	0.000	0.058	**0.324**	**0.121**	0.018	0.048	0.010	0.027	0.044	0.003	0.029	0.088
FCOM41	0.000	**0.469**	**0.526**	**0.496**	0.052	0.022	0.005	0.048	0.046	0.009	0.066	0.017
						QLIKE						
GARCH	0.012	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	0.002	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.002	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.000	0.000	0.004	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.022	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	**1.000**	0.000	0.000	**0.126**	0.076	**0.686**	**0.628**	**0.764**	**0.796**	**0.624**	**1.000**	**1.000**
FIAPARCH	0.034	0.006	0.002	0.012	0.008	0.024	0.050	0.014	0.026	0.036	0.052	0.070
LMSM	0.004	**1.000**	**1.000**	**0.874**	**1.000**	**0.314**	**0.372**	**0.236**	**0.204**	**0.376**	0.080	0.018
FCOM5	0.000	0.058	0.020	**0.180**	**0.408**	**0.916**	**0.587**	**0.966**	**0.986**	**0.971**	**0.137**	0.050
FCOM51	**0.202**	**0.602**	**0.628**	**0.904**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
						RLOG						
GARCH	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIAPARCH	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
LMSM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM61	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: The table entries are the p-values of the SPA test of Hansen (2005). The null hypothesis is that a base model cannot be outperformed by other competitor models. The values in bold face represent the p-values that are greater than or equal to the 10% confidence level under a pre-specified loss function. We combine: FCOM4=FIGARCH+HYGARCH, FCOM41=FIGARCH+LMSM, FCOM5=HYGARCH+LMSM, FCOM51=FIAPARCH+LMSM, FCOM6=GARCH+FIAPARCH and FCOM61=FIAPARCH+LMSM.

Table 4.16.: Superior predictive ability (SPA) test results using oil price observations from January 3, 1977 to December 31, 2009 as in-sample and from January 4, 2010 to March 24, 2014 as out-of-sample.

Base model	Forecast horizons											
	1	5	10	20	30	40	50	60	70	80	90	100
							MSE					
GARCH	0.015	0.007	0.053	**0.125**	**0.392**	**0.232**	**0.749**	**0.772**	**0.360**	**0.303**	**0.538**	**0.571**
IGARCH	0.029	0.094	0.046	0.053	0.037	**0.158**	**0.261**	0.010	0.010	0.026	0.005	0.003
GJR-GARCH	0.021	**0.106**	0.028	**0.127**	**0.773**	**0.218**	**0.517**	**0.364**	**0.178**	0.058	0.014	0.007
EGARCH	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.012	0.044	0.008	0.082	**0.368**	0.016	**0.144**	0.084	0.004	0.002	0.000	0.000
FIGARCH	0.014	0.084	0.031	**0.156**	**0.581**	**0.160**	**0.817**	**0.192**	**0.272**	0.026	**0.219**	**0.259**
HYGARCH	**1.000**	0.006	0.001	0.005	0.002	0.009	0.000	0.001	0.002	0.000	0.000	0.000
FIAPARCH	0.021	0.013	0.015	0.040	0.071	0.096	**0.323**	**0.131**	**0.298**	**0.397**	**0.379**	**0.524**
LMSM	0.009	**1.000**	**1.000**	**1.000**	**0.814**	**1.000**	**0.730**	**0.732**	**0.884**	**0.812**	**0.872**	**0.817**
FCOM1	0.030	**0.242**	**0.201**	**0.322**	**0.695**	**0.272**	**0.803**	**0.883**	**0.415**	**0.488**	**0.895**	**0.884**
FCOM11	0.017	**0.117**	**0.163**	**0.253**	**0.645**	**0.350**	**0.846**	**1.000**	**0.731**	**0.775**	**0.850**	**0.771**
							MAE					
GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
APARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIAPARCH	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	0.021	0.005	0.004	0.006	0.000
LMSM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
FCOM2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM21	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
							HMSE					
GARCH	0.019	0.003	0.028	0.004	0.067	0.030	0.034	0.000	0.016	0.021	0.074	0.026
IGARCH	0.015	0.013	0.025	0.018	0.087	0.070	0.049	0.018	0.059	0.054	0.074	0.009
GJR-GARCH	0.016	0.013	0.012	0.005	0.086	0.068	0.047	0.020	0.067	0.053	0.074	0.008
EGARCH	0.032	0.023	**0.847**	**0.863**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
APARCH	0.091	0.031	0.003	0.034	0.086	0.011	0.023	0.023	0.043	0.035	0.030	0.120
FIGARCH	0.035	0.001	0.019	0.026	0.052	0.056	0.053	0.017	0.028	0.064	0.069	0.022
HYGARCH	**1.000**	0.036	0.025	0.010	0.038	0.001	0.056	0.009	0.030	0.055	0.073	0.119
FIAPARCH	0.011	0.024	0.018	0.041	0.073	0.015	0.000	0.014	0.003	0.000	0.000	0.004
LMSM	0.023	**1.000**	**0.153**	**0.137**	0.046	0.018	0.019	0.026	0.019	0.030	0.036	0.050
FCOM3	0.006	0.042	0.046	0.015	0.059	0.026	0.069	0.040	0.047	0.069	**0.119**	**0.140**

Note: The table entries are the p-values of the SPA test of Hansen (2005). The null hypothesis is that a base model cannot be outperformed by other competitor models. The values in bold face represent the p-values that are greater than or equal to the 10% confidence level under a pre-specified loss function. We combine: FCOM1=GARCH+GJR-GARCH+FIGARCH, FCOM11=GARCH+GJR-GARCH+LMSM, FCOM2=GJR-GARCH+FIAPARCH, FCOM21=FIAPARCH+LMSM and FCOM3=EGARCH+HYGARCH+LMSM.

Table 4.17.: Superior predictive ability (SPA) test results using oil prices observation from January 3, 1977 to December 31, 2009 as in-sample and from January 4, 2010 to March 24, 2014 as out-of-sample.

Base model	Forecast horizons											
	1	5	10	20	30	40	50	60	70	80	90	100
							HMAE					
GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.004	0.001
GJR-GARCH	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.006	0.001
EGARCH	0.000	0.002	**0.426**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**0.868**	**0.789**	**0.546**	0.384
APARCH	0.000	0.000	0.000	0.000	0.004	0.001	0.001	0.012	0.011	0.010	0.012	0.012
FIGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	**1.000**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.003	0.001
FIAPARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
LMSM	0.000	**1.000**	**0.574**	0.045	0.006	0.020	0.029	0.083	**0.132**	**0.211**	**0.454**	**0.616**
FCOM4	0.000	0.045	**0.877**	**0.159**	0.043	**0.158**	**0.312**	**0.646**	**0.825**	**0.901**	**0.773**	**0.604**
							QLIKE					
GARCH	0.002	0.003	0.042	0.022	0.052	0.032	0.067	**0.135**	0.045	0.079	**0.118**	0.076
IGARCH	0.002	0.082	**0.208**	**0.343**	**0.140**	**0.419**	**0.538**	**0.980**	**0.315**	**0.835**	**0.256**	**0.817**
GJR-GARCH	0.003	**0.309**	**0.485**	**0.630**	**0.542**	**0.569**	**0.305**	**0.466**	**0.172**	**0.372**	**0.175**	**0.415**
EGARCH	0.000	**0.104**	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.003	**0.131**	**0.348**	**0.315**	**0.489**	**0.188**	**0.416**	**0.393**	0.068	0.036	0.038	0.001
FIGARCH	0.004	**0.351**	**0.135**	**0.462**	**0.639**	**0.465**	**0.780**	**0.871**	**0.697**	**0.692**	**0.986**	**0.869**
HYGARCH	**1.000**	**0.373**	**0.470**	**0.529**	**0.735**	**0.777**	**0.766**	**0.396**	**0.664**	**0.647**	**0.779**	**0.345**
FIAPARCH	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
LMSM	0.000	**0.916**	**0.854**	**0.866**	**0.833**	**0.714**	**0.793**	**0.534**	**0.783**	**0.710**	**0.497**	**0.524**
FCOM5	0.003	0.097	**0.186**	**0.388**	**0.166**	**0.618**	**0.457**	**0.722**	**0.193**	**0.607**	**0.275**	**0.659**
FCOM51	0.000	**0.584**	**0.652**	**0.775**	**0.844**	**1.000**	**1.000**	**0.880**	**0.902**	**0.873**	**0.868**	**0.890**
							RLOG					
GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
IGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
GJR-GARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
EGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
APARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
HYGARCH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FIAPARCH	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
LMSM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FCOM61	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: The table entries are the p-values of the SPA test of Hansen (2005). The null hypothesis is that a base model cannot be outperformed by other competitor models. The values in bold face represent the p-values that are greater than or equal to the 10% confidence level under a pre-specified loss function. We combine: FCOM4=EGARCH+LMSM, FCOM5=IGARCH+GJR-GARCH, FCOM51=FIGARCH+HYGARCH+LMSM, FCOM6=HYGARCH+FIAPARCH and FCOM61=FIAPARCH+LMSM.

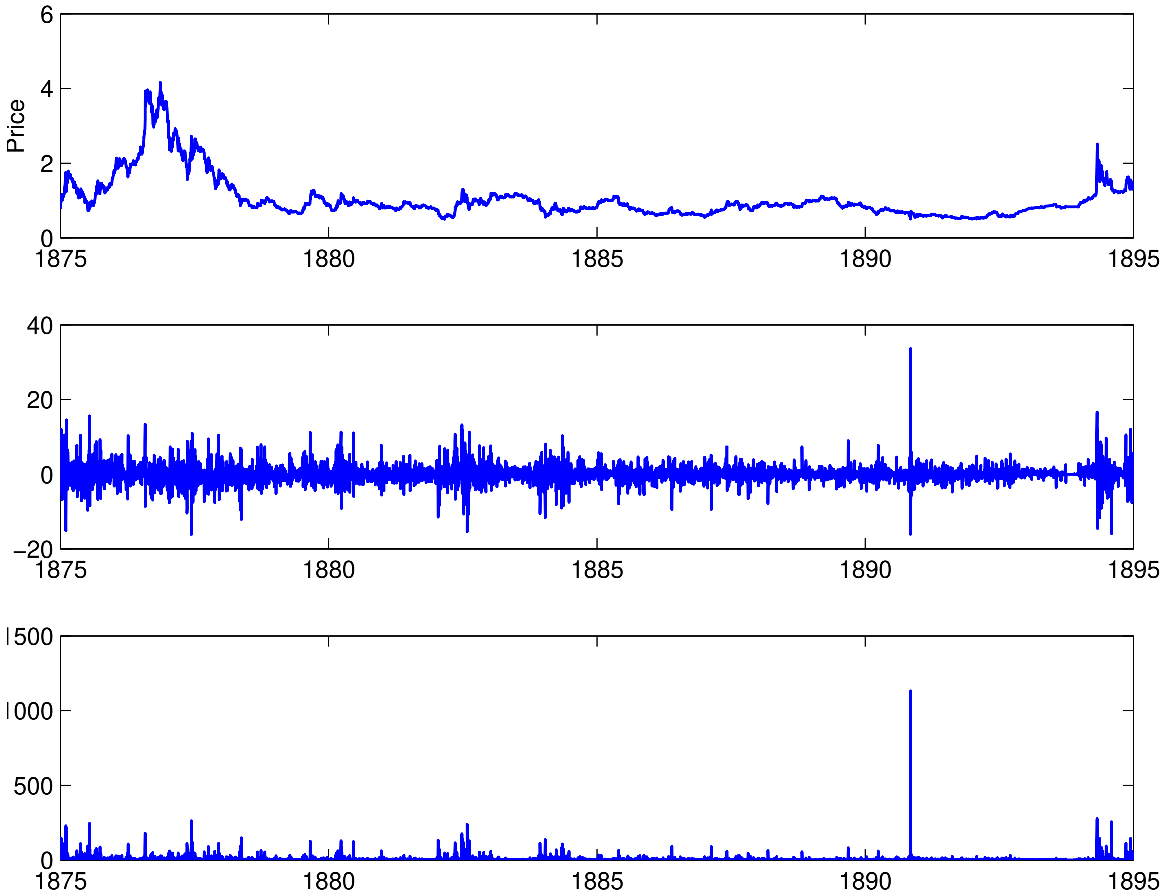

Figure 4.1.: Plot of oil prices, log-returns and squared returns (from January 2, 1875 to December 31, 1895)

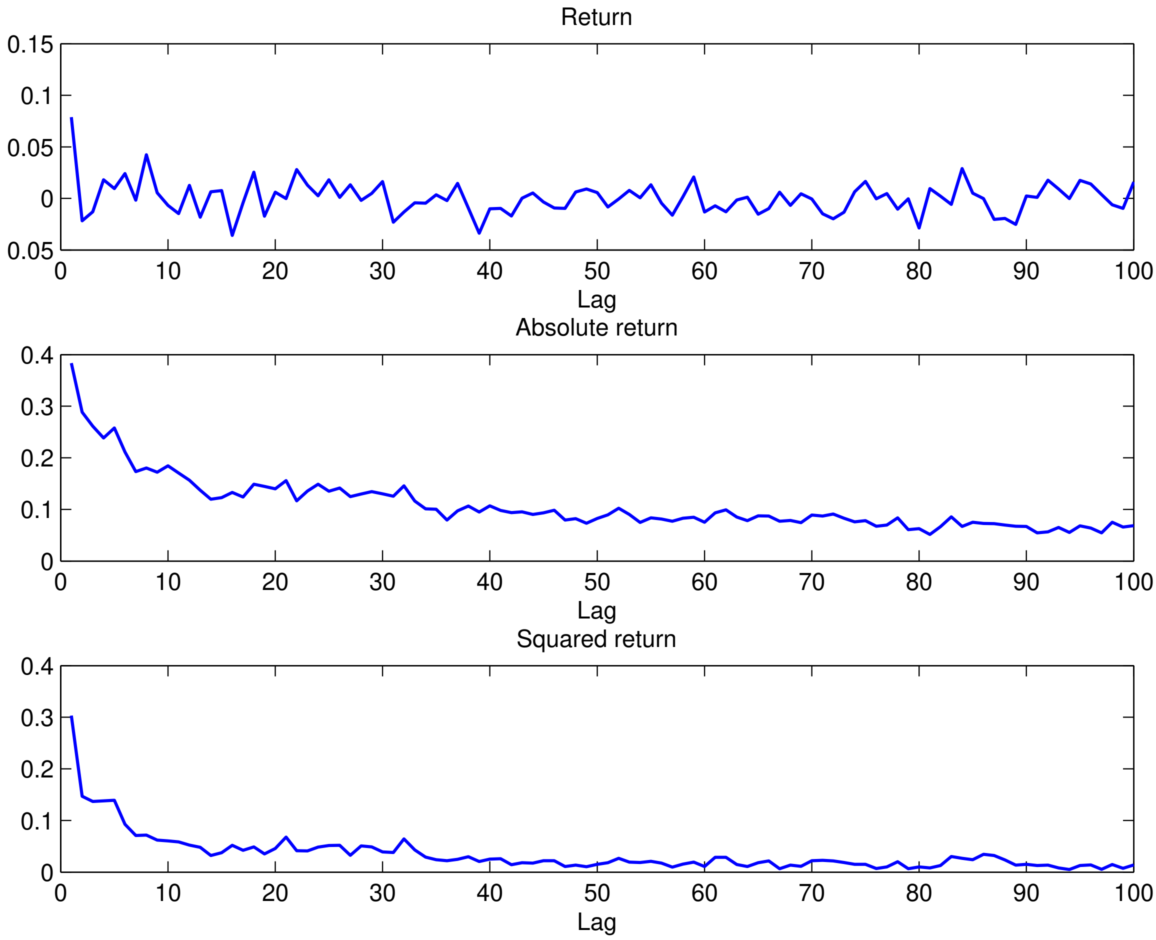

Figure 4.2.: Plot of autocorrelation functions of log-returns, absolute and squared log-returns (from January 2, 1875 to December 31, 1895)

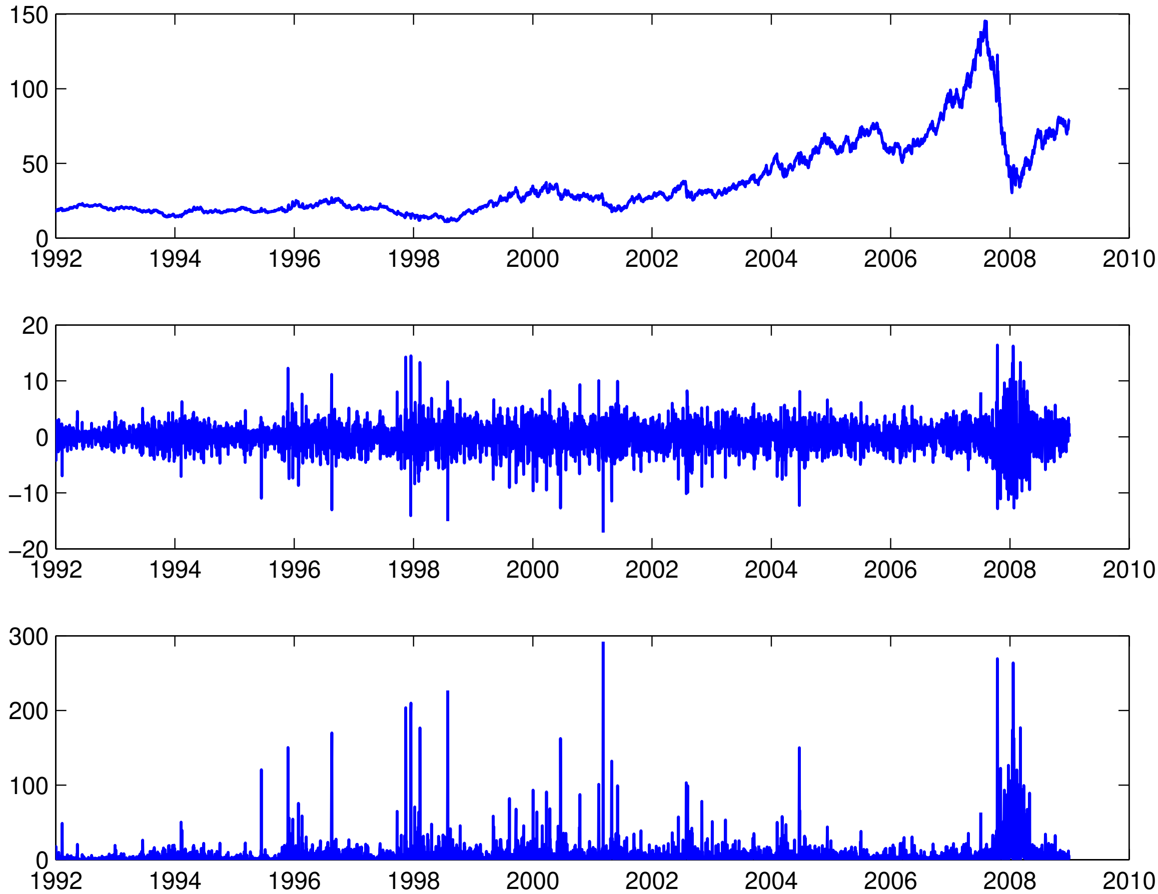

Figure 4.3.: Plot of oil prices, log-returns and squared returns (from January 6, 1992 to December 31, 2009)

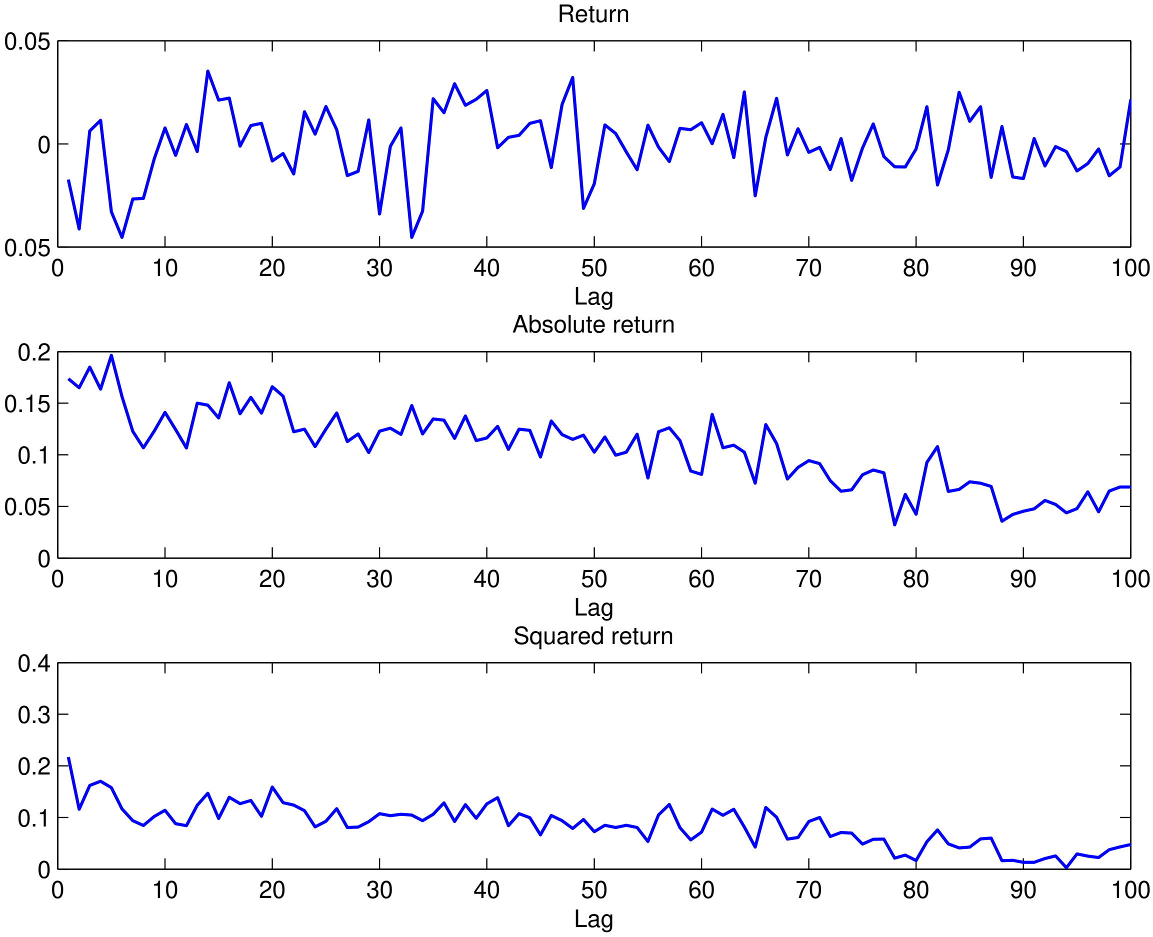

Figure 4.4.: Plot of autocorrelation functions of log-returns, absolute and squared log-returns (from January 6, 1992 to December 31, 2009)

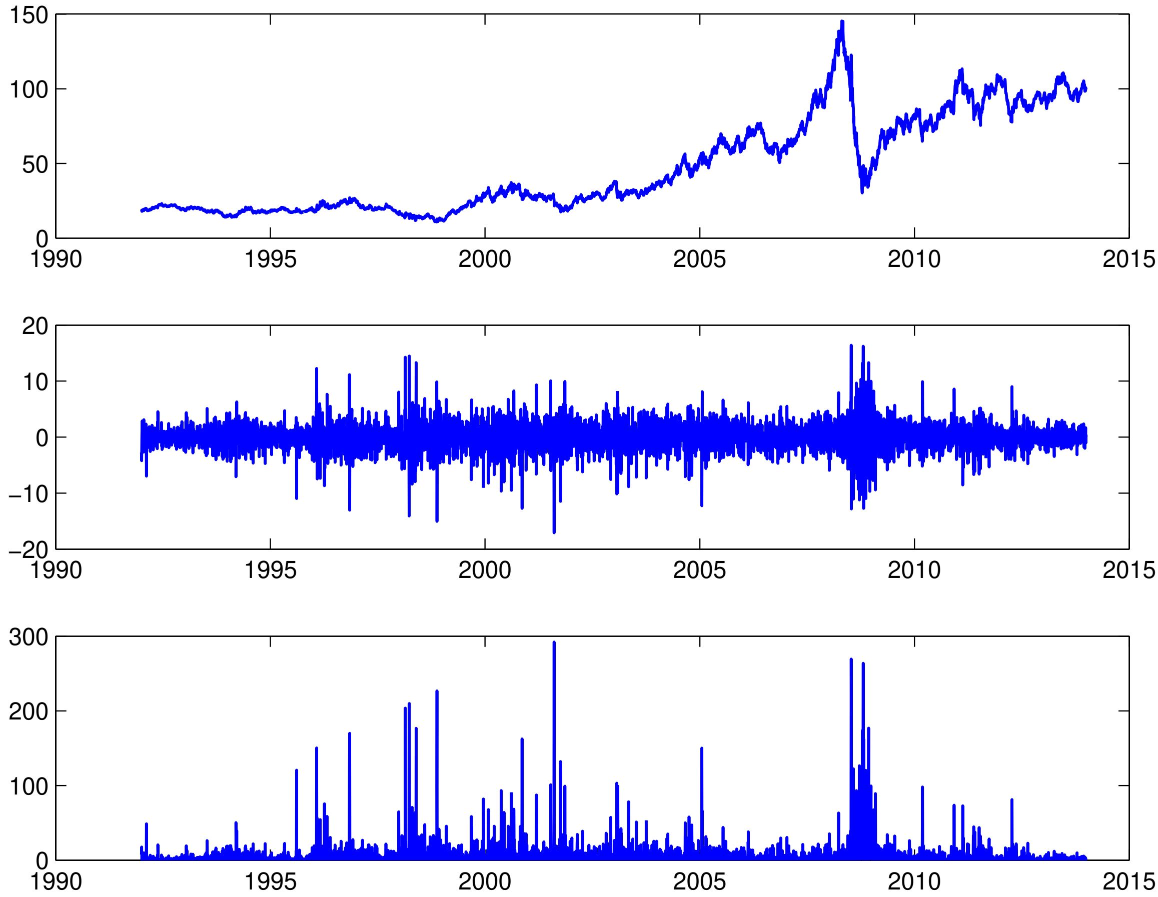

Figure 4.5.: Plot of oil prices, log-returns and squared returns (from January 6, 1992 to March 24, 2014)

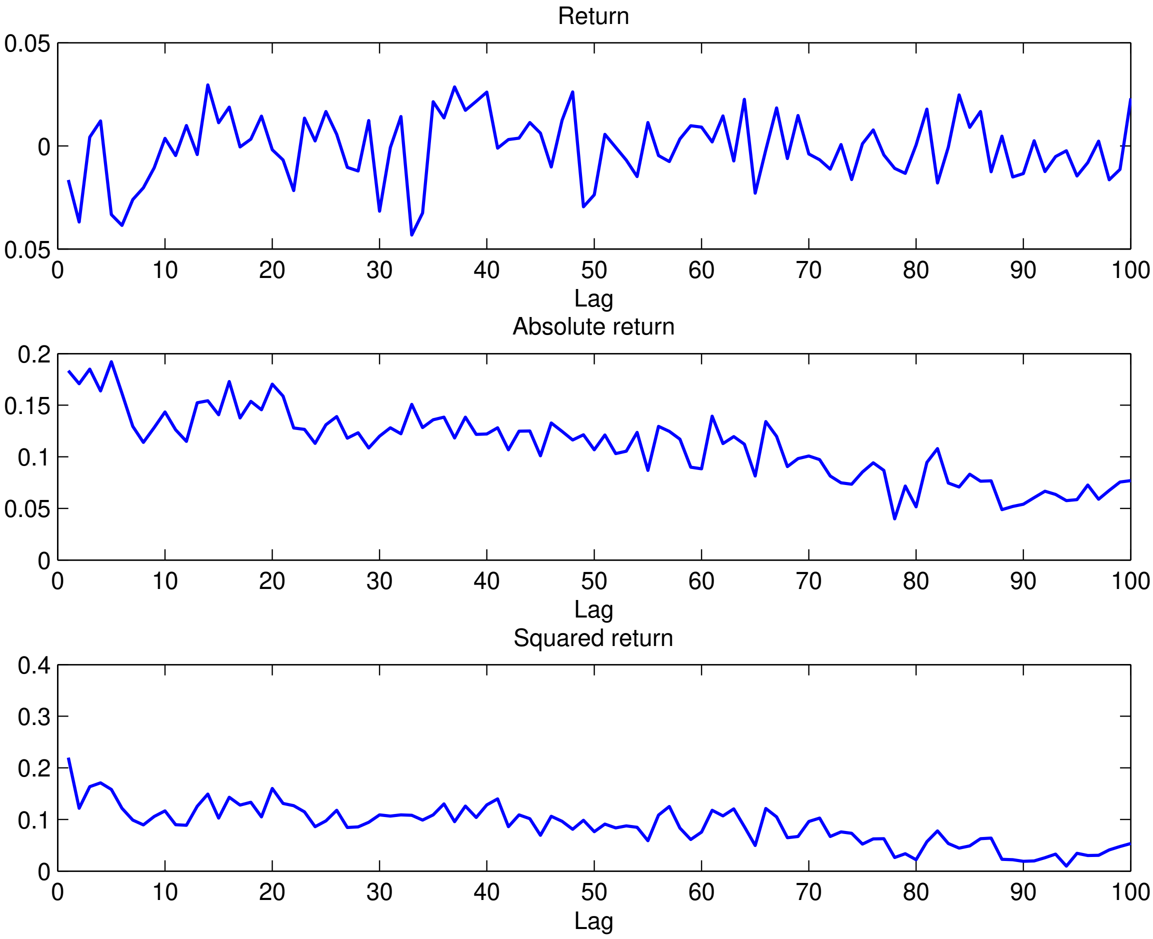

Figure 4.6.: Plot of autocorrelation functions of log-returns, absolute and squared log-returns (from January 6, 1992 to March 24, 2014)

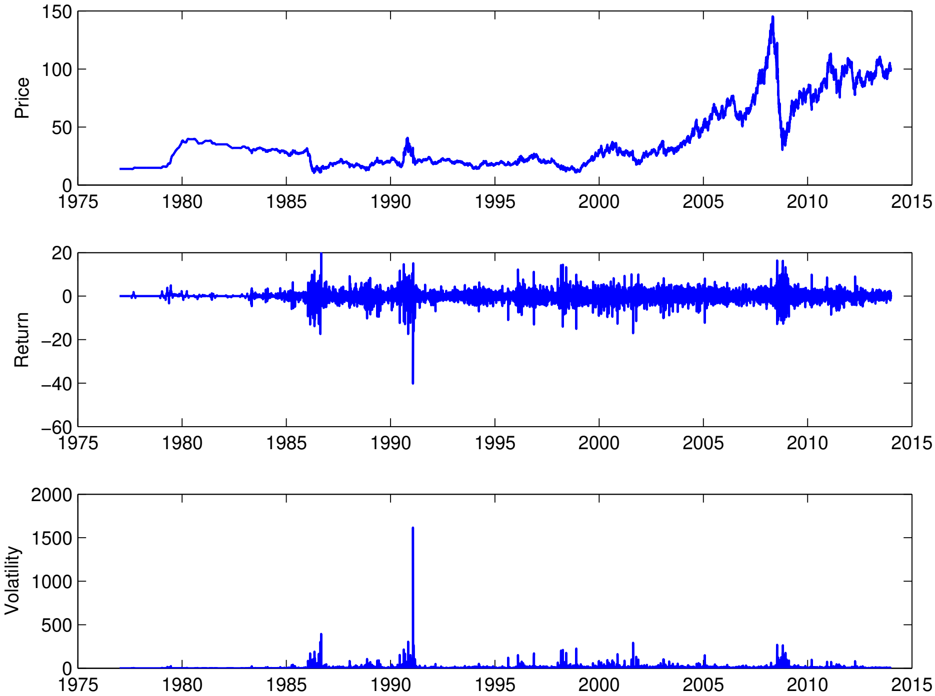

Figure 4.7.: Plot of oil prices, log-returns and squared returns (from January 6, 1977 to March 24, 2014)

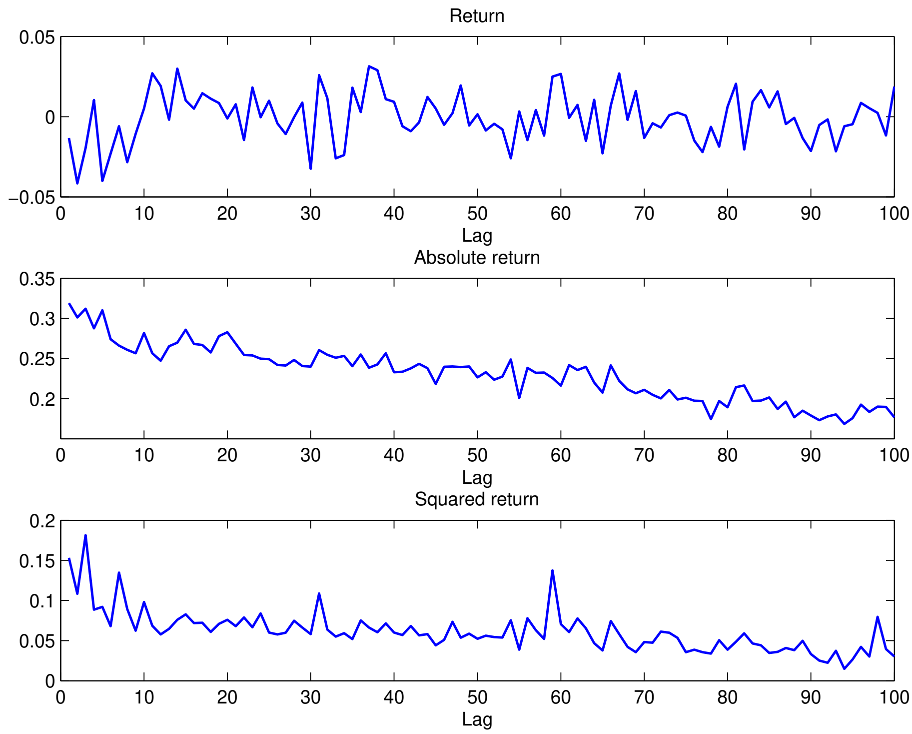

Figure 4.8.: Plot of autocorrelation functions of log-returns, absolute and squared log-returns (from January 6, 1977 to March 24, 2014)

Part III.

Application Of Multifractal Processes To Modeling Financial Markets Microstructure

5. Assessing Forecast Performance of Financial Duration Models via Density Forecasts and Likelihood Ratio Test

5.1. Introduction

The evaluation of a model's forecasting performance plays an important role in financial econometrics. This is due to the fact that it provides information on the adequacy of the models used and on the reliability of their forecasts. In the past point forecasts have been applied in most research works. However, with the development of interval forecasts (cf. Chatfield, 1993) and density forecasts (cf. Diebold et al., 1998) the attention has shifted from point forecasts to interval and density forecasts.

The purpose of this chapter is to evaluate in detail the predictive ability of the Markov switching multifractal duration (MSMD) model recently proposed by Chen et al. (2013). Although the MSMD model has already been evaluated by Chen et al. (2013) via point forecasts, we broaden the scope of the evaluation of the model in four important aspects: (i) we use a broad set of benchmark processes for comparison, (ii) we use a larger number of distributions for the innovations, (iii) we apply additional criteria for comparison of the performance of alternatives, i.e. density forecasts, and (iv) we look at further data beyond those investigated by Chen et al. (2013).

Given the existing strand of research on the evaluation of different specifications of ACD models (cf. 3.3) and the promising new competitor in the form of the Binomial MSMD model, we find that it is of paramount importance to assess the predictive ability of the Binomial MSMD model and compare its forecasting performance to those of the ACD, Log-ACD and FIACD models. We use Weibull, Lognormal, Burr, and generalized gamma distributions for the innovations in the ACD and Log-ACD models and exponential one for the innovation in the FIACD model. The forecasting performance comparison between the Binomial MSMD model and the alternative ACD models is done via density forecasts and likelihood ratio test for a sample of eight stocks traded on the New York Stock Exchange.

The rest of the chapter is organized as follows. In Section 5.2, we give a brief overview of the different financial duration models, and we present in Section 5.3 the different methodologies for estimation of these models. We describe the methodologies for the evaluation of density forecasts in Section 5.4. Section 5.5 reports the empirical results, and we conclude in Section 5.6.

5.2. Model Review

In this section, we briefly describe the standard ACD(1,1), the logarithmic ACD(1,1), the FI-ACD(1,d,1), and the Binomial Markov switching multifractal duration (Binomial MSMD) models.

5.2.1. The ACD Model

In the standard ACD(1,1) model a financial duration, x_t, is modeled as

$$\begin{aligned} x_t &= \Psi_t \xi_t \\ \Psi_t &= \omega + \beta_1 x_{t-1} + \delta_1 \Psi_{t-1}, \end{aligned} \tag{5.1}$$

where $\omega > 0, \beta_1 > 0, \delta_1 > 0$, and $\beta_1 + \delta_1 < 1$. The innovation term ξ_t is independent identically Exponentially distributed with unit-mean. The constraints on the coefficients are to ensure positive durations and the existence of the unconditional mean of the durations.

Empirical observations of the distribution of financial durations indicate that the hazard function of x_t, which is defined as the density function divided by the survivor function, may be increasing for small durations and decreasing for long durations. In the following, we propose four different distributions for the innovations which can replicate these variations in the hazard function. The corresponding density functions of durations are obtained using a transform of variables technique (cf. Appendix A.1).

1. *The Weibull distribution*:

   The Weibull distribution allows for either increasing or decreasing hazard functions. This flexibility makes it more attractive than the exponential distribution, which leads to constant hazard function. In financial duration models it requires a distribution with unit expectation. Assuming that ξ_t in *eq.* (5.1) are independent and identically distributed and follow a Weibull(1,α) distribution, the corresponding pdf is given by

$$f(\xi_t; \alpha) = \begin{cases} \alpha \xi_t^{\alpha-1} \exp\left[-\xi_t^{\alpha}\right] & \xi_t \geq 0 \\ 0 & \xi_t < 0. \end{cases} \tag{5.2}$$

   It is clear that the expectation of ξ_t is: $\mathbb{E}(\xi_t) = \Gamma\left(1 + \frac{1}{\alpha}\right)$ which does not fulfil the unit-mean requirement. Therefore, we normalize the distribution so that $\mathbb{E}[\xi_t] = 1$. The normalization leads to the following pdf

$$g(\xi_t;\alpha) = \begin{cases} \alpha\xi_t^{\alpha-1}c^{\alpha}\exp\left[-(\xi_t c)^{\alpha}\right] & \xi_t \geq 0 \\ 0 & \xi_t < 0, \end{cases} \tag{5.3}$$

where $c = \Gamma\left(1 + \frac{1}{\alpha}\right)$.

Given $g(\xi_t;\alpha)$, and the fact that x_t is a monotonic transformation of ξ_t, cf. (5.1), we can easily derive the pdf of x_t by making the change of variable as follows. We have

$$\begin{cases} x_t = \Psi_t\xi_t \\ f(x_t;\alpha) = g(\xi_t;\alpha)\left|\frac{d\xi_t}{dx_t}\right|. \end{cases} \tag{5.4}$$

eq. (5.4) leads to the pdf of x_t

$$f(x_t;\alpha) = \begin{cases} \frac{\alpha}{x_t}\left[\frac{x_t}{\phi_t}\right]^{\alpha}\exp\left[-\left(\frac{x_t}{\phi_t}\right)^{\alpha}\right] & x_t \geq 0 \\ 0 & x_t < 0, \end{cases} \tag{5.5}$$

where $\phi_t = \Psi_t\left[\Gamma(1 + 1/\alpha)\right]^{-1}$.

Given the information set $\mathfrak{I}$ available at time $t-1$, the conditional density of x_t is then

$$W(x_t|\mathfrak{I}_{t-1};\alpha) = \frac{\alpha}{x_t}\left[\frac{x_t}{\phi_t}\right]^{\alpha}\exp\left[-\left(\frac{x_t}{\phi_t}\right)^{\alpha}\right], \quad x \geq 0. \tag{5.6}$$

2. *The Burr distribution*:

The Burr distribution goes back to Burr (1942). Lancaster (1990) points out that the Burr distribution can be derived as a gamma mixture of Weibull distributions. This distribution offers more flexibility because it has the exponential, Weibull and log-logistic distributions as limiting cases. We assume that ξ_t are independent and identically distributed and follow a Burr(κ, σ^2) distribution. We normalize the distribution so that $\mathbb{E}[\xi_t] = 1$ and the corresponding pdf is given by

$$f(\xi_t;\kappa,\sigma^2) = \frac{\kappa a^{\kappa}\xi_t^{\kappa-1}}{(1+\sigma^2 a^{\kappa}\xi_t^{\kappa})^{\left(\frac{1}{\sigma^2}+1\right)}}, \quad \xi_t \geq 0, \tag{5.7}$$

where

$$a = \left[\frac{\Gamma\left(1+\frac{1}{\kappa}\right)\Gamma\left(\frac{1}{\sigma^2}-\frac{1}{\kappa}\right)}{(\sigma^2)^{\left(1+\frac{1}{\kappa}\right)}\Gamma\left(\frac{1}{\sigma^2}+1\right)}\right] \quad \text{and} \quad \kappa > 0, \quad \sigma^2 > 0. \tag{5.8}$$

We proceed as previously done for the case of the Weibull distribution (cf. *eq.* (5.4)) and obtain the conditional density of x_t as

$$B(x_t|\Im_{t-1};\kappa,\sigma^2) = \frac{\kappa\pi_t^{\kappa}x_t^{\kappa-1}}{(1+\sigma^2\pi_t^{\kappa}x_t^{\kappa})^{\left(\frac{1}{\sigma^2}+1\right)}}, \quad x \geq 0, \tag{5.9}$$

where

$$\pi_t = a\Psi_t^{-1}. \tag{5.10}$$

3. *The generalized gamma (GG) distribution*:

The generalized gamma density also provides more flexibility to fit the data. The generalized gamma distribution introduced by Stacy (1962) has exponential, gamma, and Weibull as subfamilies. The Lognormal distribution is obtained as a limiting distribution when ν approaches ∞. Assuming that ξ_t are independent and identically distributed and follow a GG(τ, ν) distribution and normalizing the distribution so that $\mathbb{E}[\xi_t] = 1$ the corresponding pdf is given by

$$f(\xi_t;\nu,\tau) = \begin{cases} \dfrac{\nu\xi_t^{\nu\tau-1}}{\beta^{\nu\tau}\Gamma(\tau)}\exp\left[-\left(\dfrac{\xi_t}{\beta}\right)^{\nu}\right], & \text{if } \xi_t > 0 \\ 0, & \text{otherwise,} \end{cases} \tag{5.11}$$

where $\beta = \dfrac{\Gamma(\tau)}{\Gamma(\tau+\frac{1}{\nu})}$ and $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}dt. \tag{5.12}$$

Following the same procedures as previously done in the case of the Weibull distribution (cf. *eq.* (5.4)) we obtain the conditional density of x_t as

$$GG(x_t|\Im_{t-1}) = \frac{\nu(x_t)^{\nu\tau-1}}{(\theta_t)^{\nu\tau}\Gamma(\tau)}\exp\left[-\left(\frac{x_t}{\theta_t}\right)^{\nu}\right], \quad x \geq 0, \tag{5.13}$$

where

$$\theta_t = \Psi_t\frac{\Gamma(\tau)}{\Gamma(\tau+\frac{1}{\nu})}, \quad \tau > 0 \quad \nu > 0. \tag{5.14}$$

4. *The Lognormal distribution*:

The log-normal distribution has only recently been introduced in the context of durations modeling by some authors (cf. Allen et al., 2008, 2009; Sun et al., 2008) who find that this distribution also constitutes a good candidate for modeling of financial durations. Assuming that ξ_t are independent and identically distributed and follow a $\ln N(\mu, \sigma^2)$ distribution and normalizing the distribution so that $\mathbb{E}[\xi_t] = 1$, the associated pdf is given by

$$f(\xi_t;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}\xi_t} \exp\left[-\frac{\left(\ln(\xi_t) + \frac{1}{2}\sigma^2\right)^2}{2\sigma^2}\right], \quad \xi_t \geq 0. \tag{5.15}$$

Normalization via $\mathbb{E}[\xi_t] = 1$ leads to $\exp\left(-\mu + 0.5\sigma^2\right) = 1$. From this restriction one can establish the following relationship: $\mu = 0.5\sigma^2$. As result, the distribution is completely determined by the scale parameter σ^2.

By applying the change of variable technique we obtain the conditional density function of x_t, given the past information set $\Im_{t-1}$, as

$$f(x_t;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}x_t} \exp\left[-\frac{\left(\ln(x_t/\Psi_t) + \frac{1}{2}\sigma^2\right)^2}{2\sigma^2}\right], \quad x_t \geq 0. \tag{5.16}$$

5.2.2. The Log-ACD Model

Bauwens and Giot (2000) investigated the logarithmic ACD (1,1) model. In the Log-ACD model a financial duration, x_t, is formalized as

$$\begin{aligned} x_t &= \exp(\psi_t)\xi_t, \qquad \xi_t \sim i.i.d., \text{ with } E(\xi_t) = 1 \\ \psi_t &= \omega + \beta_1\xi_{t-1} + \delta_1\psi_{t-1}, \end{aligned} \tag{5.17}$$

where ψ_t is the logarithm of the conditional duration $\Psi_t = \exp(\psi_t)$.

5.2.3. The Fractionally Integrated ACD Model

As defined in sec. 3.3.4 a financial duration x_t in the FIACD(p,d,q) can be formalized as

$$\begin{aligned} [1-\delta(L)]\Psi_t &= \omega^* + [1-\delta(L) - [1-\phi(L)](1-L)^d]x_t \\ &= \omega^* + A(L)x_t \end{aligned} \tag{5.18}$$

where $A(L) = a_1L + a_2L^2 + a_3L^3 + \ldots$ is a polynomial of infinite order with $a_k \geq 0$, $k = 1, 2, \ldots$, $\omega^* > 0$, and $0 \leq d \leq 1$.

Here we study a special specification of the FIACD(p,d,q) where $p = q = 1$ and assumed that the innovation, ξ_t, is standard exponential distributed. In the FIACD(1,d,1) *eq.* (5.18) becomes

$$\begin{aligned}\Psi_t &= \omega^* (1-\delta_1 L)^{-1} + \left[1 - (1-\delta_1 L)^{-1} (1-\phi_1 L)(1-L)^d\right] x_t \\ &= \omega + B(L) x_t,\end{aligned} \tag{5.19}$$

where

$$B(L) = b_1 L + b_2 L^2 + \cdots = 1 - (1-\delta_1 L)^{-1} (1-\phi_1 L)(1-L)^d \tag{5.20}$$

is a polynomial of infinite order with $b_k \geq 0$, $k = 1, 2, \ldots$, and $\omega > 0$.

The parameters, b_k, of the $B(L)$ can be expressed as

$$\begin{aligned} b_1 &= \phi_1 - \delta_1 + d \\ b_2 &= (d-\delta_1)(\delta_1 - \phi_1) + \frac{d(1-d)}{2} \\ &\vdots \\ b_k &= \delta_1 b_{k-1} + \left(\frac{k-1-d}{k} - \phi_1\right) \pi_{d,k-1} \quad k = 2, 3, \ldots; \end{aligned} \tag{5.21}$$

where $\pi_{d,k} = \pi_{d,k-1}(k-1-d)k^{-1}$. Note that $\pi_{d,k}$ represents the terms of the expansion of $(1-L)^d$ that can be expressed as

$$\pi_d(L) = \sum_{k=0}^{\infty} \pi_{d,k} L^k. \tag{5.22}$$

To guarantee positivity of durations in the FIACD(1,d,1) the parameters ϕ, δ and d have to fulfill the following conditions:

$$\delta_1 - d \leq \phi_1 \leq \frac{2-d}{3}, \quad d\left(\phi - \frac{1-d}{2}\right) \leq \delta_1 (d - \delta_1 + \phi_1).$$

As stressed in Jasiak (1998) the FIACD(1,d,1) can be easily estimated via maximum likelihood method by choosing a suitable truncation point that we set to 1000 in our empirical study (cf. Jasiak, 1998; Baillie et al., 1996).

5.2.4. The Binomial MSMD Model

In the MSMD model proposed by Chen et al. (2013) a financial duration, x_t, can be expressed as

$$x_t = \frac{\zeta_t}{\lambda_t}, \tag{5.23}$$

where ζ_t is *i.i.d.* standard exponential distributed, and λ_t is the mean intensity. The dynamic process governing the mean intensity is described in detail in sec. 3.4.1

5.3. Estimation Methods

5.3.1. ML Estimation for ACD Models

For the standard ACD, Log-ACD and FIACD models the estimation of parameters can be obtained via a maximum likelihood method. Let $L(x_1,\dots,x_T;\Xi)$ be the likelihood function with parameter vector $\Xi = (\omega,\beta_1,\delta_1,d;\eta)$ with ω, β_1, δ_1 and d the structural parameters of *eq.* (5.1) or (5.17) or (5.19) and η the vector of distributional parameters for the pertinent choice of the distribution for innovations. The maximum likelihood estimator (MLE) is given by

$$\hat{\Xi} = \arg\max_{\Xi} \ln L(x_1,\dots,x_T;\Xi), \tag{5.24}$$

where $L(x_1,\dots,x_T;\Xi)$ is the product of the T appropriate density functions. It is clear that the functional form of the likelihood function depends on the distributional assumption on ξ_t.

5.3.2. ML Estimation for the Binomial MSMD Model

With standard filtering methods a closed form solution can also be obtained for the likelihood function of the Binomial MSMD model and a maximum likelihood method can be used to estimate the parameter vector $\varphi = (m_0,\bar{\lambda},b,\gamma_1)$ of the model. Let

$$f(x_1,\dots,x_T;\varphi) = \prod_{t=1}^{T} f(x_t|x_{t-1},\dots,x_1;\varphi), \tag{5.25}$$

the joint probability density function of durations $(x_1\dots,x_T)$. Given the information set $\Im_{t-1}$ available at the time $t-1$, the one step ahead density can be written as

$$f(x_t|x_{t-1},\dots,x_1;\varphi) = \sum_{k=1}^{2^k} f(x_t|M_{t-1}=m^i;\varphi)Pr(M_{t-1}=m^i|x_{t-1},\dots,x_1), \tag{5.26}$$

and the joint probability density function becomes

$$f(x_1,\dots,x_T;\varphi) = \prod_{t=1}^{T} \omega(x_t;\varphi)(\pi_{t-1}A). \tag{5.27}$$

$\omega(x_t;\varphi)$ is a vector of dimension 2^k of conditional densities ($f(x_t|M_{t-1}=m^i;\varphi) = \lambda_t(m^i)\exp[-\lambda_t(m^i)x_t]$) of any observation x_t for intensity regime m^i and the transition matrix A has components $a_{i,j} = Pr(M_{t+1}=m^j|M_t=m^i)$. M_t is a latent variable, but one can recursively compute the conditional probabilities $\pi_t^i = Pr(M_t=m^i|x_t,\dots,x_1)$ through Bayesian updating

$$\pi_t = \frac{\omega(x_t;\varphi)*(\pi_{t-1}A)}{\sum \omega(x_t;\varphi)*(\pi_{t-1}A)}. \tag{5.28}$$

where $*$ represents the element by element product. The estimates of the parameters are obtained as

$$\hat{\varphi} = \arg\max_{\varphi} \sum_{t=1}^{T} \ln\left[\omega(x_t)\,(\pi_{t-1}A)\right]. \tag{5.29}$$

5.4. Density Forecasts

Since we also wish to discriminate between different candidate distributions for the innovations, we evaluate the predictive ability of different duration models not via point forecasts, but via their density forecasts. Point forecasts just tell us how well a model captures the dynamics of the durations around the mean, but do not shed much light on the accuracy of the shape of the residual distribution. In order to compare ACD models, we need information about the suitability of each model and the appropriateness of the residual distribution. To this end, we use tests of density forecasts as developed by Diebold et al. (1998). With their tools it is possible to evaluate nested and non-nested models. The methodology for evaluating density forecasts is based on the integral transform which goes back to Rosenblatt (1952). Let us denote by $\{p_t(x_t|\Im_t)\}_{t=1}^{\infty}$ a sequence of densities identifying the data generating process governing the durations x_t and $\{f_t(x_t|\Im_t)\}_{t=1}^{\infty}$, the sequence of one-step-ahead density forecasts produced by any duration model. Diebold et al. (1998) prove that the correct density is weakly superior to all other forecasts. This suggests to test whether $\{f_t(x_t|\Im_t)\}_{t=1}^{\infty} = \{p_t(x_t|\Im_t)\}_{t=1}^{\infty}$ which seems not to be feasible at first sight. Rosenblatt (1952) derived that under the null hypothesis the probability integral transform, $z_t = \int_{-\infty}^{x_t} f_t(y)dy$, is uniformly distributed. Diebold et al. (1998) extended Rosenblatt's research and showed that under the null hypothesis the probability integral transform, z_t, is *i.i.d.* uniformly distributed. This implies that the evaluation of forecasts consists in assessing whether the probability integral transform series, $\{z_t\}_{t=1}^{T}$, are *i.i.d.* $U(0,1)$. Diebold et al. (1998) recommended simple tests of *i.i.d.* $U(0,1)$ behavior such as those of Kolmogorov-Smirnov and Cramer-von Mises and graphical tools as complements to these tests. For visual inspection, one can plot a histogram based on an empirical z sequence as well as the autocorrelation function of $(z_t - \bar{z})$, $(z_t - \bar{z})^2$. A visual inspection of the histogram can help to detect departures from uniformity and Pearson's goodness-of-fit test can be computed by exploiting statistical properties of the histogram under the null hypothesis of uniformity. The autocorrelation functions reveal potential deficiencies of a model to account for the dynamics of a duration process and Ljung-Box Q-statistic for $(z_t - \bar{z})$ and $(z_t - \bar{z})^2$ may be used to test independencies.

Recently, Berkowitz (2001) proposes a more powerful tool for evaluating density forecasts. His suggestion is to use the inverse Normal transform of the z sequence. This transformation allows us to simply use the likelihood ratio test.

The density forecasts of ACD or Binomial MSMD models are the conditional densities of x_t given the past information $\Im_{t-1}$. The z sequences of ACD and of Binomial MSMD models are obtained by integral transforms of the conditional densities.

5.4.1. Testing Density Forecasts

To compare the predictive ability of ACD models to Binomial MSMD model, we used two goodness-of-fit measures, namely the Kolmogorov-Smirnov (referred to as KS henceforth) statistic and the Anderson-Darling (referred to as AD henceforth) statistic proposed by Rachev and Mittnik (2000) as well as the likelihood ratio test developed by Berkowitz (2001).

5.4.1.1. Kolmogorov-Smirnov Distance

The Kolmogorov-Smirnov test is used in order to know whether a sample comes from a hypothesized continuous distribution. As stressed in Conover (1999) the KS-statistic or distance counts among the supremum class of empirical distribution function (EDF) test statistics[1] and is designed on the largest vertical difference between the hypothesized and empirical distribution.

Here the Kolmogorov-Smirnov statistic helps us to test *i.i.d.* $U(0,1)$ behavior of $\{z_t\}_{t=1}^{T}$. We test the hypothesis

$$H_0\colon F(x) = F_0(x) \qquad \text{against} \qquad H_a\colon F(x) \neq F_0(x) \quad \forall x, \tag{5.30}$$

where F_0 denotes a known cumulative distribution function (cdf).

The test statistic (KS) is defined as

$$\text{KS} = \sup_{x\in\mathbb{R}} |\hat{F}(x) - F_0(x)| \tag{5.31}$$

where

$$\hat{F}(z) := \frac{1}{T}\sum_{t=1}^{T} 1\{x_t \leq z\} \tag{5.32}$$

is the empirical distribution function (EDF), and $F_0(x)$ is a uniform cumulative distribution function. The major drawback of the KS statistic is that it tends to be more sensitive near the center of the distribution, i.e., around the median value, $F_0(x) = 0.5$, than at the tails where $F_0(x)$ is close to 0 or 1.

5.4.1.2. Anderson-Darling Distance

As an alternative to the KS statistic we also used the AD statistic proposed by Rachev and Mittnik (2000) to test *i.i.d.* $U(0,1)$ behavior of $\{z_t\}_{t=1}^{T}$. This test statistic is designed in such a way that discrepancies in the tails of the distribution are conveniently weighted. Rachev and Mittnik (2000) defined the AD statistic as follows

$$\text{AD} = \sup_{x\in\mathbb{R}} \frac{|\hat{F}(x) - F_0(x)|}{\sqrt{F_0(x)(1 - F_0(x))}}, \tag{5.33}$$

[1] An EDF test statistic is defined as a statistic that measures the difference between the empirical distribution function (EDF) and the hypothesized one.

where $\hat{F}(x)$ and $F_0(x)$ are defined as in *eq.* (5.31).

With both statistics; the KS distance appropriate for the deviations around the median of the distribution and the AD distance convenient for the tails, we can obtain reliable results of testing the empirical distribution. Research results by Arshed et al. (2003) pointed out that the AD test is the most powerful EDF test. By comparing the AD test to the KS test Razali and Wah (2011) found that the AD test is more powerful than the KS test. However, they pointed out that the power of both tests remains still low for small sample size.

5.4.1.3. Likelihood Ratio Test

The likelihood ratio test is a more powerful tool for evaluating density forecasts. Using a simple transformation to normality, Berkowitz (2001) obtained the following proposition:

- If the sequence $z_t = \int_{\infty}^{x_t} f(u)du$ is distributed as an *i.i.d.* $U(0,1)$, then

$$v_t = \Phi^{-1}\left[\int_{\infty}^{x_t} f(u)du\right] \quad \text{is an } \textit{i.i.d.} \quad N(0,1). \tag{5.34}$$

With the new sequence v, one can test the joint null hypothesis (H_0) of independence and normality against a first-order autoregressive AR(1) with mean and variance different from 0 and 1, respectively.

Let us consider the following AR(1) process $v_t - \mu = \rho(v_{t-1} - \mu) + \varepsilon_t$, where μ is the mean of v_t, ρ is the AR(1) parameter and ε_t is a white noise process. The exact log-likelihood (L) function associated with the AR(1) process is given by

$$\begin{aligned} L = &-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln[\sigma^2/(1-\rho^2)] - \frac{[v_1 - \mu/(1-\rho)]^2}{2\sigma^2/(1-\rho^2)} \\ &- \frac{T-1}{2}\ln(2\pi) - \frac{T-1}{2}\ln(\sigma^2) - \sum_{t=2}^{T}\left[\frac{(v_t - \mu - \rho v_{t-1})^2}{2\sigma^2}\right]. \end{aligned} \tag{5.35}$$

The likelihood ratio test statistic is given by

$$LR = -2\left[L(0,1,0) - L(\hat{\mu}, \hat{\sigma}^2, \hat{\rho})\right], \tag{5.36}$$

where $L(0,1,0)$ is the value of the log-likelihood function under H_0 and $L(\hat{\mu}, \hat{\sigma}^2, \hat{\rho})$ is the estimated log-likelihood function associated with the AR(1) process. Under the null hypothesis, the test statistic is chi-squared distributed with three degrees of freedom ($\chi^2(3)$).

The LR test by Berkowitz has some shortcomings. It can happen that the null hypothesis is accepted because the conditions $\mu = 0$, $\rho = 0$ and $\mathbb{V}ar(\varepsilon_t) = 1$ are true, but the sequence v is not normal, as it should be under the null hypothesis. In other words, Berkowitz's test can fail to detect model failure arising from non-normality of the sequence v. To solve this problem, Dowd (2004) suggests the Jarque-Bera test as a complement to the Berkowitz's test.

5.5. Empirical Application

5.5.1. Data

We use data of eight stocks traded on the New York Stock Exchange (NYSE): Citigroup (C), International Business Machines (IBM), Bank of America (BAC), Coca Cola (KO), Walt Disney (DIS), Boeing (BA), General Motors (GM), and Ford Motor (F). For each stock trade durations are defined as time elapsed between two consecutive trades. The sampling period corresponds to July 2004 which has 21 trading days. We also consider price durations for only two stocks, namely IBM, and BAC. The price duration ($x_t(\iota_p)$) is the minimal time interval needed to observe a change in the mid-price[2] (p) not less than a threshold (ι_p) that is set to \$0.0625. Mathematically, we define the price duration as:

$$x_t = \inf\{x \in \mathbb{R}_+, \text{ such that } |p_{T_t+x} - p_{T_t}| \geq \iota_p\}. \tag{5.37}$$

For price durations, the sampling period covers July and August 2004. The data were extracted from the Trade and Quote (TAQ) database[3] available at the NYSE. For each stock we only take into account the transactions in the period from 10 : 00 to 16 : 00 in order to avoid the effects of opening auction (cf. (Engle and Russell, 1998); and (Ghysels et al., 2004)). Note that we do not include the over night durations and zero-values. Table 5.1 reports the statistical characteristics of the raw data for trade durations.

5.5.1.1. Seasonal Adjustment

The Raw data shows a strong seasonal patterns which can be categorized in two groups: a *day-of-the-week* effect and a *time-of-the-day* effect. The first one comes from the fact that the trading activity at the beginning of the week is low and becomes very high at the end of the week. This can be observed in *Fig.* 5.1 where durations remain constantly high between Monday and Wednesday, then decrease continuously afterward, and finally show their smallest values on Friday. The second one is due to systematic variations of trade arrivals over the trading day. Typically, the average duration is short at morning opening time and afternoon closing time, and long around noon or at lunch time. At the start of the day, trading activities are very high due to new events (macroeconomic news or news released by firms after the previous market close) that occurred during the night. At the end of the day, traders want to close their positions before the end of the trading session (cf. *Fig.* 5.2). Following Engle and Russell (1998), raw data can be adjusted as follows: We first remove the *day-of-the-week effect* by averaging the duration for each weekday and divide the raw data for each day of the week by the average duration. To eliminate the remaining effect known as the *time-of-the-day effect* from the data ($\tilde{X}_t$) obtained after the first adjustment, we use a *cubic spline* with 13 knots chosen over the trading day to smooth the time-of-day function. To

[2] Mid-price is used to avoid biases caused by a bid-ask bounce (cf. (Roll, 1984)).

[3] This database consists of two parts: The first reports all trades, while the second lists the best bid and the ask prices posted by market makers.

obtain the values at each knot, the first one being at 10 : 00 and the last one being at 16 : 00, we consider intervals of 30 minutes around each knot and calculate the average duration over these 30 minutes. The adjusted duration is obtained as:

$$x_t = \tilde{X}_t / \varrho(T_t) \tag{5.38}$$

where $\tilde{X}_t$ is the duration after the day-of-week effect has been removed, $\varrho(T_t)$ is the time-of-day effect and x_t denotes the adjusted duration (cf. *Figs.* 5.5 and 5.6 for examples of plots of raw and adjusted durations). Applying the *cubic spline* approach for price durations the structure of data changes drastically and this affects the estimates of parameters. For this reason we also make use of other approaches for seasonal adjustment and find that only the *linear spline* does not modify the structure of raw data (price durations). The adjusted price durations are obtained by using a *linear spline* with 7 knots at round hours, 10 : 00, 11 : 00, 12 : 00, 13 : 00, 14 : 00, 15 : 00, and 16 : 00.

5.5.2. Results of Performance Comparison

For each stock and for each type of data (trade and price durations), we estimate the ACD, the Log-ACD, the FIACD and the Binomial MSMD models. The results of the estimation and the standard errors are reported in Tables 5.7, 5.8, 5.9, and 5.10.

5.5.2.1. Estimation and Comparison of Estimated Models

The design of the simulation consists of two different experiments. We first generate data from the Binomial MSMD model, using different number of intensity components (k) and different values of m_0. Table 5.5 shows the empirical moments obtained from the simulation exercise. The Binomial MSMD model exhibits overdispersion, clustering, asymmetry, and heavy tails. By varying the intensity components k and the value for m_0, we observe that the Binomial MSMD model offers a lot of flexibility to fit data with different degrees of heterogeneity and different dynamic structures. This means that one can generate data from the Binomial MSMD model with different values for k and m_0 and compare the dynamic properties of the simulated data to those of the raw data at hand.

Secondly, we estimate the standard ACD, the Log-ACD, FIACD and the Binomial MSMD models, using trade duration data for IBM. Note that in the estimation procedures of the Binomial MSMD model we use $k = 7$. The choice of intensity components, k, is motivated from the results we obtained in the first simulation exercise. The estimated parameters are used to generate data from each model specification (ACD, Log-ACD, FIACD and Binomial MSMD models) and to compare their empirical moments and autocorrelation functions to those of IBM trade durations. The results are reported in Table 5.6. Except for the ACD or Log-ACD model with Weibull distribution for error terms, all models exhibit overdispersion which is consistent with IBM duration data. Skewness and kurtosis obtained from simulated data are sometimes above or below that obtained from IBM trade duration data. The inability of the Binomial MSMD model to proprely reproduce the asymmetric effects in the data is due to the fact that the MSM process used is symmetric. In *Fig.* 5.7 and 5.8 the autocorrelation functions of WACD, LACD, BACD, and GGACD models decay quickly and after some lags run parallel to that of IBM trade durations. In contrast, the autocorrelation function of the FIACD and Binomial MSMD model decay hyperbolically and nicely approximate that of IBM trade durations. This suggests that both models can better capture the long memory (cf. Deo et al., 2010) which is observed in the real data. *Fig.* 5.9 depicts the autocorrelation functions of simulated data from the FIACD model, the Binomial MSMD model and IBM trade durations.

5.5.2.2. In- and Out-Of-Sample Results

To evaluate and compare the four models we conduct "in-sample" and "out-of-sample" exercises. To do this, we split the data set for each stock into four subsets of equal size. For the "in-sample" we estimate each model with the last fourth of the data, and then forecast densities and probability integral transforms (z) are calculated on the same sample. In the case of "out-of-sample" we use the first three-fourths of the data to estimate each model specification, and then the forecast densities and z sequences are computed on the last fourth of the data, using the parameters obtained from the estimation using the first three-fourths of the data.

With z sequences obtained from "in-sample" and "out-of-sample" exercises using duration data, we first calculate KS and AD statistics for each model specification and each stock. The results are reported in Table 5.11. Except for KO trade durations the best model under "the AD distance" to fit trade durations in the tails of the distribution by conducting "in-sample" exercises is the Binomial MSMD model while for C, IBM, and KO trade durations the best model under "the KS distance" to fit the data near the median is the Burr ACD model. Expect for C the KS and AD statistics for all other stocks obtained from the Binomial MSMD model by doing "out-of-sample" exercises are smaller than those of the ACD, Log-ACD and FIACD models. This gives evidence that the Binomial MSMD model is the best model to fit the data near the median as well as in the tails of the distribution under "the KS and AD distances". We do not report the results of the Log-ACD model because both the Log-ACD and ACD models exhibit similar results. Secondly, the null hypothesis of the LR test, and the Jarque-Bera (JB) test for trade durations are strongly rejected because the p-values are very small. One can argue, however, that this strong rejection of the models is due to the large sample size of the data used in this empirical application. To have more information about how the models perform, we consider two stocks, namely Ford (F) and General Motors (GM) and visualize the histograms of their probability integral transforms (z) and the autocorrelation functions of z for ACD and Binomial MSMD models. z-histograms of F and GM obtained from WACD, LACD, BACD, and GGACD models have peaks near 0 and 1 (cf. *Fig.* 5.10 and 5.11). This suggests that these models might not be able to capture the heavy tails in trade duration data. This suggestion is in harmony with the AD statistics, which are very high for WACD, LACD, BACD, and GGACD models. By the EFIACD model we also observe a peak near 1. However, this is not so pronounced. Compared to the z-histograms for WACD, LACD, BACD, GGACD, and EFIACD models the z-histograms for Binomial MSMD model are relatively "well-behaved" and seem to approximate the histograms of the uniform distribution.

If a model is successful in capturing the dynamic structure of the durations, then the z sequences should be independent. To get information on how far the models capture the time dependence in duration data, we plot the z-correlograms for F and GM and perform the Ljung-Box test. For WACD, LACD, BACD, and GGACD models the null hypothesis (no autocorrelation) is strongly rejected because the p-values are very small. In the EFIACD model while the null hypothesis is rejected for F, it is accepted at the 1% confidence level for GM. For Binomial MSMD model the results of the Ljung-Box test are fine. The null hypothesis of no autocorrelation is accepted at least at the 1% confidence level for both stocks F and GM (cf. Table 5.13). In sum, there is also evidence that the Binomial MSMD model is more successful in capturing the dynamic structure of the trade durations (cf. Table 5.14).

For IBM price durations, the WACD and EFIACD models followed by the Binomial MSMD model exhibit high KS and AD statistics, while small values for KS and AD statistics are obtained by the generalized gamma distribution. This implies that the best model to fit IBM price durations in the median and in the tail of the distribution is the GGACD model. This result is also in harmony with the LR test, and the Jarque-Bera (JB) test that are significant at the 5% level for the GGACD model, i.e. the null hypothesis of the LR test and the JB test cannot be rejected at the 5% confidence level (cf. Table 5.12). The WACD and Binomial MSMD models are strongly rejected because the p-values are very small. Lognormal and Burr distributions also exhibit small KS and AD statistics compared to those of WACD and Binomial MSMD models.

One has to be careful when interpreting the results of the Berkowitz LR tests we obtained for IBM price durations in LACD and BACD models. In fact, while the null hypothesis of the LR tests are accepted at the 10% level, the null hypothesis of the JB tests (Normality) are strongly rejected at all confidence levels. These results indicate that the LR test fails to detect the non-normality of the v sequences ($v = \Phi^{-1}(z)$) and that it makes sense to supplement the LR test with the JB test in order to avoid wrong inferences.

By observing the z-histograms of IBM price durations, it is clear that the z-histograms for LACD, BACD and GGACD models nicely approximate that of the uniform distribution (cf. *Fig.* 5.12 and 5.13). The null hypothesis of no autocorrelation in the z sequence are almost accepted at the 5% confidence level for all models and the visual inspection of z-correlograms suggests that WADC, LACD, BACD, GGACD, and Binomial MSMD models exhibit similar performance in capturing the dynamic structure of the IBM price durations (cf. *Fig.* 5.14 and 5.15). Surprisingly, the EFIACD model shows some deficiencies by capturing the dynamic structure. We believe that this is due to the uncertainty associated with the parameters estimation.

For BAC price durations the KS and AD statistics for Binomial MSMD model are smaller than those of WACD, LACD, BACD, GGACD, and EFIACD models. The null hypothesis of the LR and the JB tests are accepted at the 5% level for the Binomial MSMD model in both in-sample and out-of-sample. In the case of in-sample exercises the null hypothesis of the LR and the JB tests are accepted at the 5% level for WACD, BACD, and GGACD models and rejected for the EFIACD model. Here the Ljung-Box tests for no autocorrelation are strongly significant for all models at the 5% level. These results are also supported by the visual inspection of z-correlograms for BAC price durations. The z-correlograms reveal that all models are similarly successful in modeling the dynamic structure of BAC price durations (cf. *Fig.* 5.16 and 5.17).

5.6. Conclusion

We analyze four different models, namely the ACD, the Log-ACD, the FIACD models both with different distributional assumptions (exponential, Weibull, Lognormal, Burr, and generalized gamma) for error terms, and the Binomial MSMD model. Our Monte Carlo studies show that the BMSMD model is able to mimic the most important *stylized facts* such as clustering effects, overdispersion, non-linearities, long memory and heavy tails observed in financial duration data. Results from Monte Carlo analysis give evidence that the Binomial MSMD model can better reproduce the long memory property than the ACD or Log-ACD model and is similar to the FIACD in term of reproducing the long range dependence. We applied density forecast evaluation methodologies and likelihood ratio test to compare the predictive ability of the Markov switching multifractal duration model to those of standard ACD, Log-ACD and FIACD models with different distributional assumptions for error terms. The choice of the standard ACD and Log-ACD is due to the fact that there is evidence in the literature (cf. (Bauwens et al., 2004)) that forecast performances of complex models such as TACD, SCD or SVD are not superior to those of ACD and Log-ACD models. The results from empirical application show that the Binomial MSMD model outperforms the ACD, Log-ACD and FIACD models when modeling trade duration data. It is also worthwhile to note that the ACD model and Log-ACD model exhibit similar forecast performances. Compared to ACD and Log-ACD models which only allow for ARMA-type dynamics, the Binomial MSMD model allows for a multiplicative mixture of components determining the duration between trades. The use of the Markov switching multifractal (MSM) process seems to provide for higher flexibility due to its large number of intensity states so that it dominates the performance of the ACD or Log-ACD model.

When applying the Binomial MSMD model to price durations, we observe that the Binomial MSMD model slightly dominates or exhibits similar forecast performance as the ACD or Log-ACD model. Here it is important to note that the distributional assumptions for error terms play an important role in the modeling of price durations. So, the ACD model with generalized gamma distribution for the error term provides in most cases good results and is sometimes preferred to the Binomial MSMD model. We find that the Burr distribution performs well and can also be used when modeling price durations. Surprisingly, the forecast performance of the FIACD is not so good. The FIACD model is dominated by the standard ACD models.

A promising future research avenue seems to try appropriate extensions of the Binomial MSMD modeling approach. The MSM process used in the Binomial MSMD modeling approach is a discrete version of a Poisson multifractal process. Duration data are statistically viewed as point processes. It is well known that the Poisson process is a simple point process itself. For this reason it might be more convenient to use the continuous-time Poisson multifractal process in the modeling approach. Another possibility to improve the model is to introduce asymmetries. It is well documented that financial duration data exhibit asymmetry, but unfortunately, however the MSM process is symmetric. We think that one should be able to enhance predictive ability of the Binomial MSMD model by introducing appropriate asymmetries.

In the next chapter 6 we propose new alternative MSMD models, compare their forecast performance to that of the standard MSMD model developed by Chen et al. (2013), and infer from the results obtained here and those of the next chapter 6.

Table 5.1.: Information on the raw data

	Trade durations							
	C	IBM	BAC	KO	DIS	BA	GM	F
Number of obs.	94520	89162	72899	68052	68002	63067	51797	48295
Minimum value	1	1	1	1	1	1	1	1
Maximum value	141	91	129	281	129	199	218	228
Mean value	4.796	5.086	6.220	6.660	6.668	7.189	8.750	9.384
Standard dev.	5.587	5.829	8.004	9.070	8.338	9.791	11.928	11.613
Overdispersion	1.165	1.146	1.287	1.362	1.250	1.362	1.363	1.238
Skewness	3.861	3.433	3.522	4.357	3.221	3.722	3.694	3.378
Kurtosis	32.726	22.214	22.528	45.249	19.764	26.682	26.420	23.937

Table 5.2.: Information on the adjusted data

	Trade durations							
	C	IBM	BAC	KO	DIS	BA	GM	F
Number of obs.	94520	89162	72899	68052	68002	63067	51797	48295
Minimum value	0.109	0.091	0.068	0.054	0.076	0.065	0.046	0.036
Maximum value	24.592	24.457	25.316	42.461	16.018	21.993	16.961	16.960
Mean value	0.962	0.952	0.940	0.944	0.933	0.941	0.948	0.935
Standard dev.	1.075	1.027	1.138	1.198	1.103	1.204	1.221	1.075
Overdispersion	1.118	1.079	1.211	1.269	1.182	1.280	1.288	1.150
Skewness	3.621	3.095	3.297	4.011	2.885	3.301	3.314	2.880
Kurtosis	27.634	19.871	21.837	44.782	15.804	21.538	20.780	17.427
$Q(10)$	2224.9	1631.7	1306.1	1144.3	2104.3	909.394	798.909	860.316
$Q(100)$	3708.1	4020.5	2774.5	2422.8	5188.5	1797.5	1835.2	2034.5

Note: $Q(10)$ and $Q(100)$ denote the Ljung-Box Q-statistic of order 10 and 100 on the durations.

Table 5.3.: Information on the adjusted data

	Price durations	
	IBM	BAC
Number of obs.	4734	3057
Minimum value	0.009	0.001
Maximum value	11.644	17.160
Mean value	1.002	1.027
Standard dev.	1.062	1.321
Overdispersion	1.060	1.286
Skewness	2.802	4.2538
Kurtosis	15.169	35.738
$Q(10)$	99.400	106.969
$Q(100)$	177.215	195.626

Note: $Q(10)$ and $Q(100)$ denote the Ljung-Box Q-statistic of order 10 and 100 on the durations.

Table 5.4.: Dynamic Properties of the IBM Trading Durations

	Raw Data		Adjusted Data	
Lags	Autocorrelation	Partial Autocorrelation	Autocorrelation	Partial Autocorrelation
1	0.095	0.095	0.070	0.070
2	0.090	0.082	0.063	0.059
3	0.072	0.058	0.047	0.039
4	0.067	0.049	0.040	0.031
5	0.061	0.042	0.032	0.022
6	0.057	0.037	0.029	0.021
7	0.055	0.034	0.028	0.020
8	0.053	0.031	0.026	0.017
9	0.055	0.033	0.028	0.019
10	0.064	0.041	0.040	0.031
11	0.056	0.030	0.028	0.017
12	0.054	0.027	0.026	0.015
13	0.056	0.029	0.026	0.016
14	0.054	0.026	0.028	0.017
15	0.045	0.016	0.019	0.008
16	0.050	0.023	0.024	0.014
17	0.043	0.015	0.017	0.007
18	0.053	0.026	0.025	0.015
19	0.050	0.022	0.025	0.015
20	0.044	0.015	0.016	0.005

Table 5.5.: Empirical Moments of the BMSMD model for different k and m_0 values

		Overdispersion	Skewness	Kurtosis
$m_0 = 1.1$	k=6	1.054	2.303	11.541
	k=7	1.062	2.356	12.119
	k=8	1.073	2.413	12.545
$m_0 = 1.2$	k=6	1.205	3.041	18.817
	k=7	1.244	3.239	21.168
	k=8	1.277	3.451	24.498
$m_0 = 1.3$	k=6	1.451	4.072	31.500
	k=7	1.542	4.580	40.491
	k=8	1.631	5.058	49.868
$m_0 = 1.4$	k=6	1.796	5.359	52.414
	k=7	1.934	5.968	63.996
	k=8	2.122	7.084	93.771
$m_0 = 1.5$	k=6	2.219	6.564	73.831
	k=7	2.502	7.918	108.685
	k=8	2.792	9.286	149.422

Note: The values in the Table are average results based on Monte-Carlo simulations (400 samples of size 5000). We set $b = 2$ and $\gamma_1 = 0.5$.

Table 5.6.: A Comparison of empirical moments of ACD, Log-ACD and BMSMD Models

Distributions	Models	Overdispersion	Skewness	Kurtosis
Weibull	ACD	0.912	1.796	7.844
	Log-ACD	0.909	1.772	7.655
Lognormal	ACD	1.094	4.396	45.752
	Log-ACD	1.98	4.464	48.209
Burr	ACD	4.078	25.944	1113.089
	Log-ACD	6.053	16.884	567.583
gen. gamma	ACD	1.135	4.649	51.515
	Log-ACD	1.144	4.548	46.390
Exponential	FIACD	1.024	2.170	10.392
Model	**Intensity Components**	**Overdispersion**	**Skewness**	**Kurtosis**
BMSMD	k=6	1.070	2.389	12.264
	k=7	1.071	2.416	12.766
	k=8	1.069	2.397	12.416
Trade Duration		**Overdispersion**	**Skewness**	**Kurtosis**
IBM		**1.079**	**3.095**	**19.871**

Note: The values in the Table are average results (except in the last row) based on Monte-Carlo simulation (200 samples of size equals to that of IBM trade data). The parameters used for each model specification are set equal to the values obtained from the estimation of each model using IBM data. The last column reports the descriptive statistics for the IBM data.

Table 5.7.: Estimates of ACD, Log-ACD, and BMSMD models for trade durations

		C					*IBM*					*BAC*				
Weibull		ω	β	δ	α		ω	β	δ	α		ω	β	δ	α	
	ACD	0.076	0.063	0.859	1.110		0.016	0.028	0.956	1.120		0.025	0.035	0.939	1.011	
		(0.006)	**(0.002)**	**(0.008)**	**(0.003)**		**(0.001)**	**(0.001)**	**(0.002)**	**(0.003)**		**(0.002)**	**(0.002)**	**(0.003)**	**(0.003)**	
	Log-ACD	-0.058	0.055	0.925	1.110		-0.027	0.027	0.983	1.120		-0.034	0.032	0.973	1.011	
		(0.002)	**(0.002)**	**(0.005)**	**(0.003)**		**(0.001)**	**(0.001)**	**(0.002)**	**(0.003)**		**(0.002)**	**(0.002)**	**(0.003)**	**(0.003)**	
Lognormal		ω	β	δ	σ^2		ω	β	δ	σ^2		ω	β	δ	σ^2	
	ACD	0.117	0.060	0.814	0.768		0.028	0.030	0.941	0.776		0.029	0.029	0.939	0.955	
		(0.021)	**(0.006)**	**(0.029)**	**(0.004)**		**(0.007)**	**(0.004)**	**(0.012)**	**(0.004)**		**(0.002)**	**(0.002)**	**(0.004)**	**(0.005)**	
	Log-ACD	-0.060	0.051	0.886	0.768		-0.029	0.026	0.973	0.776		-0.030	0.027	0.970	0.955	
		(0.003)	**(0.002)**	**(0.009)**	**(0.004)**		**(0.002)**	**(0.002)**	**(0.003)**	**(0.004)**		**(0.002)**	**(0.002)**	**(0.004)**	**(0.005)**	
Burr		ω	β	δ	σ^2	κ	ω	β	δ	σ^2	κ	ω	β	δ	σ^2	κ
	ACD	0.201	0.090	0.759	1.538	2.323	0.046	0.044	0.921	1.313	2.153	0.055	0.046	0.929	1.570	2.101
		(0.018)	**(0.005)**	**(0.018)**	**(0.025)**	**(0.018)**	**(0.005)**	**(0.003)**	**(0.006)**	**(0.020)**	**(0.015)**	**(0.006)**	**(0.003)**	**(0.006)**	**(0.027)**	**(0.018)**
	Log-ACD	-0.034	0.075	0.847	1.538	2.322	-0.028	0.037	0.965	1.313	2.152	-0.017	0.040	0.965	1.574	2.103
		(0.003)	**(0.003)**	**(0.012)**	**(0.024)**	**(0.017)**	**(0.001)**	**(0.002)**	**(0.004)**	**(0.021)**	**(0.015)**	**(0.002)**	**(0.003)**	**(0.004)**	**(0.028)**	**(0.018)**
gen. gamma		ω	β	δ	τ	ν	ω	β	δ	τ	ν	ω	β	δ	τ	ν
	ACD	0.081	0.055	0.858	3.998	0.560	0.035	0.042	0.918	4.505	0.524	0.047	0.048	0.900	3.881	0.503
		(0.025)	**(0.010)**	**(0.037)**	**(0.069)**	**(0.005)**	**(0.003)**	**(0.002)**	**(0.006)**	**(0.068)**	**(0.004)**	**(0.004)**	**(0.003)**	**(0.007)**	**(0.058)**	**(0.004)**
	Log-ACD	-0.133	0.091	0.686	6.710	0.433	-0.086	0.071	0.882	4.475	0.532	-0.087	0.085	0.907	2.209	0.639
		(0.010)	**(0.003)**	**(0.039)**	**(0.119)**	**(0.004)**	**(0.008)**	**(0.005)**	**(0.019)**	**(0.086)**	**(0.005)**	**(0.005)**	**(0.006)**	**(0.010)**	**(0.022)**	**(0.004)**
Exponential		ω	ϕ	d	δ		ω	ϕ	d	δ		ω	ϕ	d	δ	
	FIACD	0.315	0.297	0.161	0.369		0.293	0.405	0.180	0.510		0.291	0.407	0.178	0.507	
		(0.013)	**(0.037)**	**(0.007)**	**(0.039)**		**(0.032)**	**(0.185)**	**(0.027)**	**(0.073)**		**(0.021)**	**(0.122)**	**(0.017)**	**(0.129)**	
Binomial		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1	
	MSMD	1.114	1.142	1.784	0.335		1.104	1.173	2.318	0.351		1.265	0.609	30.148	0.980	
		(0.002)	**(0.009)**	**(0.070)**	**(0.039)**		**(0.002)**	**(0.016)**	**(0.116)**	**(0.044)**		**(0.004)**	**(0.005)**	(3.339)	**(0.014)**	

Note: the numbers in bold in parentheses are standard errors of the estimation. For the Binomial MSMD model we use $k = 7$.

Table 5.8.: Estimates of ACD, Log-ACD, and BMSMD models for trade durations

		KO					*DIS*					*BA*				
		ω	β	δ	α		ω	β	δ	α		ω	β	δ	α	
Weibull	ACD	0.032	0.038	0.928	0.978		0.022	0.043	0.934	1.011		0.028	0.034	0.936	0.952	
		(0.005)	**(0.003)**	**(0.008)**	**(0.003)**		**(0.003)**	**(0.003)**	**(0.005)**	**(0.003)**		**(0.003)**	**(0.002)**	**(0.004)**	**(0.003)**	
	Log-ACD	-0.035	0.032	0.970	0.978		-0.041	0.039	0.977	1.012		-0.034	0.032	0.970	0.952	
		(0.003)	**(0.003)**	**(0.005)**	**(0.003)**		**(0.002)**	**(0.002)**	**(0.002)**	**(0.003)**		**(0.002)**	**(0.002)**	**(0.003)**	**(0.003)**	
		ω	β	δ	σ^2		ω	β	δ	σ^2		ω	β	δ	σ^2	
Lognormal	ACD	0.043	0.032	0.920	1.029		0.040	0.045	0.912	1.000		0.035	0.028	0.933	1.093	
		(0.012)	**(0.005)**	**(0.017)**	**(0.007)**		**(0.006)**	**(0.004)**	**(0.010)**	**(0.005)**		**(0.003)**	**(0.002)**	**(0.005)**	**(0.006)**	
	Log-ACD	-0.029	0.026	0.964	1.029		-0.040	0.037	0.965	1.000		-0.029	0.026	0.964	1.093	
		(0.002)	**(0.002)**	**(0.004)**	**(0.006)**		**(0.003)**	**(0.002)**	**(0.004)**	**(0.005)**		**(0.002)**	**(0.002)**	**(0.004)**	**(0.006)**	
		ω	β	δ	σ^2	κ	ω	β	δ	σ^2	κ	ω	β	δ	σ^2	κ
Burr	ACD	0.088	0.056	0.907	1.572	2.019	0.081	0.078	0.869	1.194	1.810	0.080	0.053	0.928	1.666	2.012
		(0.013)	**(0.005)**	**(0.011)**	**(0.031)**	**(0.020)**	**(0.020)**	**(0.010)**	**(0.024)**	**(0.023)**	**(0.015)**	**(0.008)**	**(0.004)**	**(0.006)**	**(0.013)**	**(0.020)**
	Log-ACD	-0.010	0.043	0.958	1.577	2.022	-0.038	0.057	0.954	1.198	1.812	-0.002	0.047	0.962	1.678	2.019
		(0.003)	**(0.004)**	**(0.007)**	**(0.044)**	**(0.026)**	**(0.004)**	**(0.006)**	**(0.008)**	**(0.024)**	**(0.016)**	**(0.002)**	**(0.004)**	**(0.005)**	**(0.033)**	**(0.021)**
		ω	β	δ	τ	ν	ω	β	δ	τ	ν	ω	β	δ	τ	ν
gen. gamma	ACD	0.054	0.056	0.881	3.030	0.555	0.052	0.065	0.878	3.915	0.492	0.084	0.056	0.845	6.675	0.357
		(0.007)	**(0.005)**	**(0.013)**	**(0.047)**	**(0.005)**	**(0.006)**	**(0.005)**	**(0.011)**	**(0.059)**	**(0.004)**	**(0.016)**	**(0.005)**	**(0.025)**	**(0.101)**	**(0.003)**
	Log-ACD	-0.105	0.091	0.809	2.054	0.632	-0.067	0.058	0.933	3.606	0.518	-0.064	0.051	0.876	10.851	0.282
		(0.006)	**(0.006)**	**(0.019)**	**(0.020)**	**(0.003)**	**(0.006)**	**(0.005)**	**(0.011)**	**(0.060)**	**(0.005)**	**(0.007)**	**(0.004)**	**(0.030)**	**(0.220)**	**(0.003)**
		ω	ϕ	d	δ		ω	ϕ	d	δ		ω	ϕ	d	δ	
Exponential	FIACD	0.289	0.343	0.180	0.457		0.218	0.391	0.218	0.507		0.286	0.366	0.188	0.500	
		(0.067)	**(0.176)**	**(0.056)**	**(0.206)**		**(0.020)**	**(0.093)**	**(0.020)**	**(0.106)**		**(0.017)**	**(0.033)**	**(0.011)**	**(0.037)**	
		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1	
Binomial	MSMD	1.280	1.051	16.478	0.906		1.267	1.075	14.255	0.496		1.324	1.188	33.850	0.992	
		(0.005)	**(0.021)**	**(1.605)**	**(0.033)**		**(0.003)**	**(0.015)**	**(1.208)**	**(0.038)**		**(0.004)**	**(0.015)**	**(3.050)**	**(0.005)**	

Note: the numbers in bold in parentheses are standard errors of the estimation. For the Binomial MSMD model we use $k = 7$.

Table 5.9.: Estimates of ACD, Log-ACD, and BMSMD models for trade durations

		GM					*F*				
Weibull		ω	β	δ	α		ω	β	δ	α	
	ACD	0.023 **(0.003)**	0.034 **(0.002)**	0.942 **(0.005)**	0.935 **(0.003)**		0.022 **(0.003)**	0.034 **(0.002)**	0.942 **(0.005)**	0.995 **(0.003)**	
	Log-ACD	-0.033 **(0.002)**	0.031 **(0.002)**	0.975 **(0.003)**	0.935 **(0.003)**		-0.034 **(0.002)**	0.032 **(0.002)**	0.976 **(0.003)**	0.995 **(0.003)**	
Lognormal		ω	β	δ	σ^2		ω	β	δ	σ^2	
	ACD	0.027 **(0.009)**	0.028 **(0.005)**	0.944 **(0.014)**	1.178 **(0.007)**		0.029 **(0.003)**	0.034 **(0.002)**	0.937 **(0.006)**	1.148 **(0.007)**	
	Log-ACD	-0.028 **(0.002)**	0.026 **(0.002)**	0.973 **(0.003)**	1.178 **(0.007)**		-0.031 **(0.002)**	0.031 **(0.002)**	0.971 **(0.003)**	1.148 **(0.007)**	
Burr		ω	β	δ	σ^2	κ	ω	β	δ	σ^2	κ
	ACD	0.040 **(0.005)**	0.041 **(0.003)**	0.939 **(0.005)**	1.148 **(0.027)**	1.639 **(0.016)**	0.026 **(0.003)**	0.038 **(0.002)**	0.938 **(0.005)**	0.530 **(0.017)**	1.329 **(0.010)**
	Log-ACD	-0.020 **(0.002)**	0.036 **(0.003)**	0.971 **(0.003)**	1.153 **(0.027)**	1.641 **(0.016)**	-0.034 **(0.002)**	0.034 **(0.002)**	0.973 **(0.003)**	0.528 **(0.017)**	1.328 **(0.010)**
gen. gamma		ω	β	δ	τ	ν	ω	β	δ	τ	ν
	ACD	0.062 **(0.009)**	0.048 **(0.003)**	0.883 **(0.012)**	4.543 **(0.112)**	0.429 **(0.006)**	0.049 **(0.006)**	0.051 **(0.003)**	0.896 **(0.009)**	3.265 **(0.070)**	0.524 **(0.006)**
	Log-ACD	-0.039 **(0.003)**	0.036 **(0.002)**	0.965 **(0.004)**	4.891 **(0.098)**	0.408 **(0.004)**	-0.023 **(0.001)**	0.023 **(0.001)**	0.992 **(0.001)**	3.853 **(0.104)**	0.481 **(0.007)**
Exponential		ω	ϕ	d	δ		ω	ϕ	d	δ	
	FIACD	0.275 **(0.026)**	0.331 **(0.096)**	0.193 **(0.020)**	0.469 **(0.107)**		0.268 **(0.018)**	0.316 **(0.058)**	0.194 **(0.013)**	0.453 **(0.064)**	
Binomial		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1	
	MSMD	1.335 **(0.005)**	1.208 **(0.023)**	35.237 **(4.794)**	0.999 **(0.001)**		1.150 **(0.003)**	1.317 **(0.025)**	2.985 **(0.266)**	0.993 **(0.014)**	

Note: the numbers in bold in parentheses are standard errors of the estimation. For the Binomial MSMD model we use $k = 7$.

Table 5.10.: Estimates of ACD, Log-ACD, and BMSMD models for Price durations

		IBM					*BAC*				
Weibull		ω	β	δ	α		ω	β	δ	α	
	ACD	0.204	0.098	0.700	1.085		0.275	0.128	0.605	0.920	
		(0.043)	**(0.014)**	**(0.050)**	**(0.012)**		**(0.057)**	**(0.020)**	**(0.065)**	**(0.012)**	
	Log-ACD	-0.058	0.055	0.925	1.110		-0.027	0.027	0.983	1.120	
		(0.002)	**(0.002)**	**(0.005)**	**(0.003)**		**(0.001)**	**(0.001)**	**(0.002)**	**(0.003)**	
Lognormal		ω	β	δ	σ^2		ω	β	δ	σ^2	
	ACD	0.154	0.110	0.750	0.993		0.261	0.205	0.618	1.589	
		(0.029)	**(0.014)**	**(0.034)**	**(0.020)**		**(0.049)**	**(0.028)**	**(0.049)**	**(0.041)**	
	Log-ACD	-0.060	0.051	0.886	0.768		-0.029	0.026	0.973	0.776	
		(0.003)	**(0.002)**	**(0.009)**	**(0.004)**		**(0.002)**	**(0.002)**	**(0.003)**	**(0.004)**	
Burr		ω	β	δ	σ^2	κ	ω	β	δ	σ^2	κ
	ACD	0.165	0.098	0.744	0.539	1.481	0.262	0.148	0.605	0.334	1.129
		(0.034)	**(0.014)**	**(0.040)**	**(0.046)**	**(0.034)**	**(0.051)**	**(0.022)**	**(0.058)**	**(0.045)**	**(0.029)**
	Log-ACD	-0.034	0.075	0.847	1.538	2.322	-0.028	0.037	0.965	1.313	2.152
		(0.003)	**(0.003)**	**(0.012)**	**(0.024)**	**(0.017)**	**(0.001)**	**(0.002)**	**(0.004)**	**(0.021)**	**(0.015)**
gen. gamma		ω	β	δ	τ	ν	ω	β	δ	τ	ν
	ACD	0.177	0.100	0.723	4.144	0.481	0.259	0.145	0.606	2.761	0.527
		(0.039)	**(0.014)**	**(0.047)**	**(0.178)**	**(0.011)**	**(0.051)**	**(0.022)**	**(0.059)**	**(0.370)**	**(0.039)**
	Log-ACD	-0.098	0.095	0.856	2.850	0.615	-0.113	0.114	0.759	2.745	0.529
		(0.011)	**(0.011)**	**(0.026)**	**(0.186)**	**(0.022)**	**(0.015)**	**(0.014)**	**(0.043)**	**(0.367)**	**(0.039)**
Exponential		ω	ϕ	d	δ		ω	ϕ	d	δ	
	FIACD	0.385	1.817e-5	0.130	0.029		0.804	0.500	1.258e-8	0.362	
		(0.130)	**(1.000e−4)**	**(0.053)**	**(0.062)**		**(0.080)**	**(0.181)**	**(9.483e−10)**	**(0.171)**	
Binomial		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1	
	MSMD	1.172	1.260	3.853	0.285		1.276	0.892	3.661	0.642	
		(0.028)	**(0.155)**	**(2.353)**	**(0.142)**		**(0.020)**	**(0.104)**	**(1.037)**	**(0.222)**	

Note: the numbers in bold in parentheses are standard errors of the estimation. For the Binomial MSMD model we use $k = 7$.

Table 5.11.: Mean of KS and AD statistics for WACD, LACD, BACD, GGACD and BMSMD models (Trade durations)

	In − Sample															
	C		*IBM*		*BAC*		*KO*		*DIS*		*BA*		*GM*		*F*	
Models	*KS*	*AD*	*KS*	*AD*	*KS*	*AD*	*KS*	*AD*	*KS*	*AD*	*KS*	*AD*	*KS*	*AD*	*KS*	*AD*
WACD	0.092	0.215	0.088	0.206	0.103	0.243	0.114	0.270	0.109	0.264	0.105	0.253	0.110	0.270	0.058	0.140
LACD	0.046	0.120	0.036	0.093	0.035	0.088	0.043	0.107	0.041	0.110	0.038	0.098	0.043	0.114	0.020	0.058
BACD	**0.035**	**0.106**	**0.035**	0.104	**0.033**	0.098	**0.026**	**0.068**	0.037	0.106	**0.028**	0.080	0.036	0.103	0.020	0.052
GGACD	0.057	0.148	0.038	0.093	0.040	0.099	0.050	0.121	0.046	0.122	0.043	0.109	0.050	0.131	0.019	0.053
EFIACD	0.047	0.111	0.038	0.093	0.043	0.104	0.046	0.112	0.037	0.091	0.048	0.117	0.040	0.099	0.029	0.073
BMSMD	0.047	0.112	0.037	**0.090**	**0.033**	**0.081**	0.034	0.084	**0.027**	**0.068**	**0.028**	**0.070**	**0.026**	**0.066**	**0.011**	**0.033**
	Out − of − Sample															
WACD	0.122	0.284	0.110	0.264	0.104	0.244	0.114	0.267	0.109	0.265	0.106	0.252	0.113	0.271	0.060	0.148
LACD	0.061	0.155	0.041	0.107	0.036	0.089	0.042	0.105	0.042	0.113	0.038	0.097	0.043	0.109	0.026	0.075
BACD	0.062	0.169	0.044	0.124	0.038	0.103	0.040	0.104	0.043	0.120	0.029	0.082	0.044	0.116	0.035	0.095
GGACD	0.069	0.166	0.054	0.141	0.053	0.129	0.058	0.139	0.059	0.156	0.064	0.164	0.063	0.158	0.034	0.099
EFIACD	**0.043**	**0.103**	**0.038**	0.093	0.044	0.108	0.045	0.117	0.038	0.098	0.048	0.121	0.041	0.103	0.029	0.077
BMSMD	0.050	0.117	**0.038**	**0.091**	**0.035**	**0.084**	**0.035**	**0.084**	**0.027**	**0.067**	**0.027**	**0.068**	**0.025**	**0.064**	**0.015**	**0.042**

Note: the values in bold refer to the smallest mean of KS and AD statistics obtained using our portfolio of financial duration models.

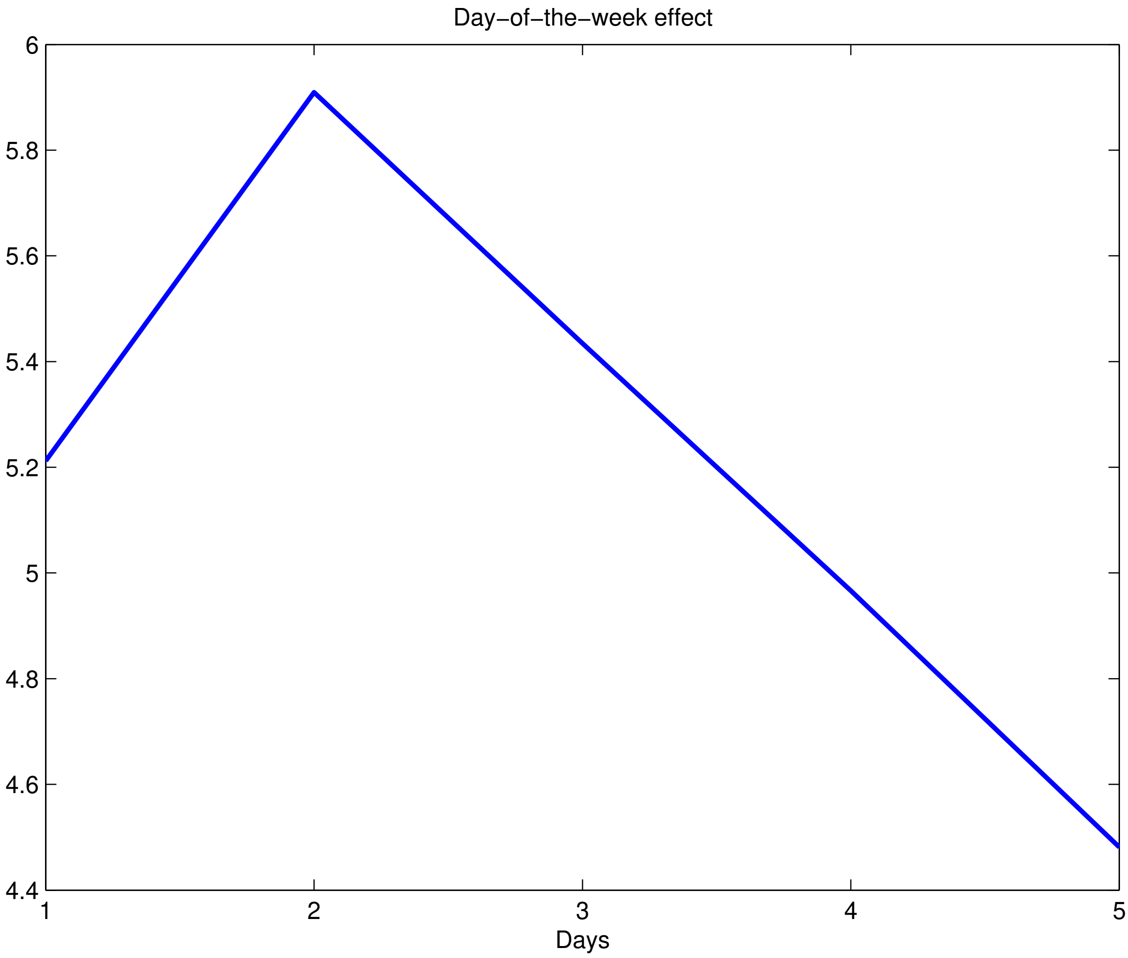

Figure 5.1.: Day-of-the-week effect.

Table 5.12.: Mean of KS and AD statistics, the likelihood ratio (LR) test, and Jarque Bera (JB) test for price durations.

	In − Sample							
	IBM				*BAC*			
Distributions	*KS*	*AD*	*LR*	*JB*	*KS*	*AD*	*LR*	*JB*
Weibull	0.034	0.077	25.207 **(0.000)**	3.540 **(0.170)**	0.014	0.034	2.384 **(0.497)**	4.165 **(0.125)**
Lognormal	**0.012**	0.031	0.463 **(0.927)**	23.741 **(0.000)**	0.033	0.090	0.186 **(0.980)**	147.254 **(0.000)**
Burr	0.013	0.032	2.993 **(0.393)**	8.022 **(0.018)**	0.011	0.029	1.576 **(0.665)**	5.351 **(0.069)**
gen. gamma	0.006	**0.016**	1.038 **(0.792)**	2.766 **(0.251)**	0.010	0.029	1.354 **(0.716)**	5.504 **(0.064)**
EFIACD	0.024	0.061	34.373 **(0.000)**	361.614 **(0.000)**	0.040	0.107	34.117 **(0.000)**	39.456 **(0.000)**
BMSMD	0.026	0.065	57.697 **(0.000)**	35.064 **(0.000)**	**0.007**	**0.022**	2.464 **(0.482)**	1.433 **(0.488)**
	Out − Sample							
Weibull	0.025	0.060	19.576 **(0.000)**	3.444 **(0.179)**	0.026	0.068	12.846 **(0.005)**	0.089 **(0.957)**
Lognormal	0.018	0.045	1.845 **(0.605)**	23.783 **(0.000)**	0.021	0.139	96.563 **(0.000)**	157.829 **(0.000)**
Burr	0.009	0.025	0.904 **(0.824)**	11.139 **(0.004)**	0.029	0.094	44.952 **(0.000)**	10.204 **(0.006)**
gen. gamma	**0.008**	**0.020**	1.089 **(0.780)**	4.170 **(0.124)**	0.024	0.088	54.446 **(0.000)**	20.746 **(0.000)**
EFIACD	0.031	0.082	57.558 **(0.000)**	239.209 **(0.000)**	0.033	0.098	53.587 **(0.000)**	49.327 **(0.000)**
BMSMD	0.028	0.069	60.262 **(0.000)**	36.700 **(0.000)**	**0.011**	**0.030**	4.194 **(0.241)**	0.570 **(0.752)**

Note: LR and JB denote the Berkowitz's likelihood ratio test and the Jarque Bera test, respectively. The numbers in parentheses are the p-values.

Table 5.13.: Ljung-Box tests for z (the probability integral transforms) and z^2

Trade Durations					Price Durations			
	F				IBM			
Model	In-sample		Out-of-sample		In-sample		Out-of-sample	
	pv	pv^2	pv	pv^2	pv	pv^2	pv	pv^2
WACD	0.000	0.000	0.000	0.001	0.512	**0.692**	0.570	**0.685**
LACD	0.000	0.000	0.000	0.002	**0.550**	0.604	0.569	0.611
BACD	0.000	0.000	0.000	0.001	0.467	0.604	0.541	0.603
GGACD	0.000	0.000	0.000	0.001	0.523	0.623	**0.584**	0.620
EFIACD	0.001	0.126	0.000	0.079	0.372	0.622	0.000	0.006
BMSMD	**0.068**	**0.102**	**0.021**	**0.092**	0.294	0.533	0.352	0.509
	GM				BAC			
Model	In-sample		Out-of-sample		In-sample		Out-of-sample	
	pv	pv^2	pv	pv^2	pv	pv^2	pv	pv^2
WACD	0.000	0.000	0.000	0.000	0.833	0.518	**0.817**	0.395
LACD	0.000	0.000	0.000	0.000	**0.868**	**0.629**	0.806	0.399
BACD	0.000	0.000	0.000	0.000	0.836	0.531	0.809	**0.423**
GGACD	0.000	0.000	0.000	0.000	0.838	0.533	0.810	0.404
EFIACD	0.011	0.105	0.000	0.001	0.734	0.402	0.121	0.089
BMSMD	**0.023**	**0.075**	**0.061**	**0.106**	0.723	0.280	0.765	0.342

Note: pv and pv^2 represent the p-values of the Ljung-Box Q statistic based on the first 50 autocorrelations of the z (probability integral transforms) and z^2, respectively. The values in bold refer to the highest p-values.

Table 5.14.: Number of significant autocorrelations out of 50 for z at 5% level

Trade Durations				
	F		GM	
Model	In-sample AC(z)	Out-of-sample AC(z)	In-sample AC(z)	Out-of-sample AC(z)
WACD	9	7	14	9
LACD	10	9	15	11
BACD	9	6	14	11
GGACD	10	10	16	10
EFIACD	7	9	5	7
BMSMD	**5**	**4**	**6**	**5**
Price Durations				
	IBM		BAC	
Model	In-sample AC(z)	Out-of-sample AC(z)	In-sample AC(z)	Out-of-sample AC(z)
WACD	3	3	1	1
LACD	3	3	2	2
BACD	3	3	1	1
GGACD	3	3	1	1
EFIACD	3	7	1	3
BMSMD	**3**	**3**	**1**	**1**

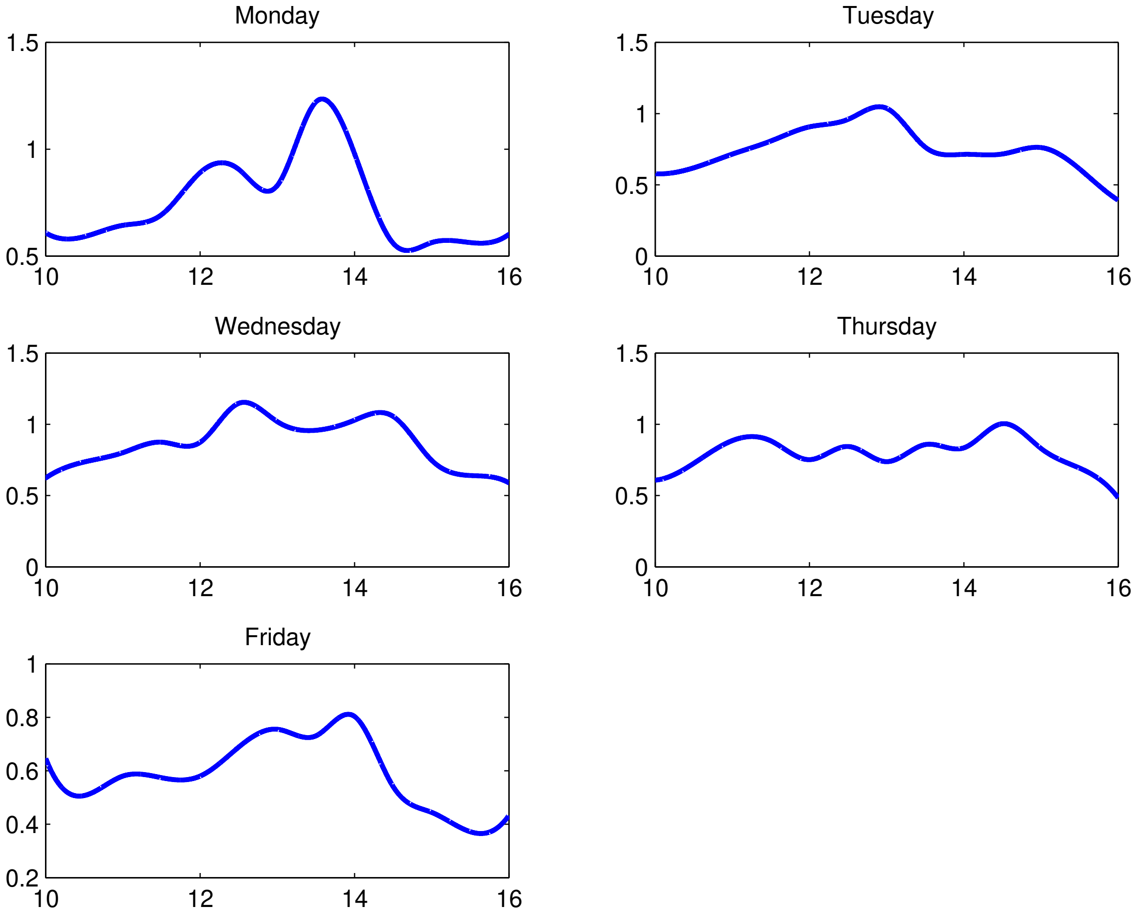

Figure 5.2.: Time-of-the-day function for IBM trade durations

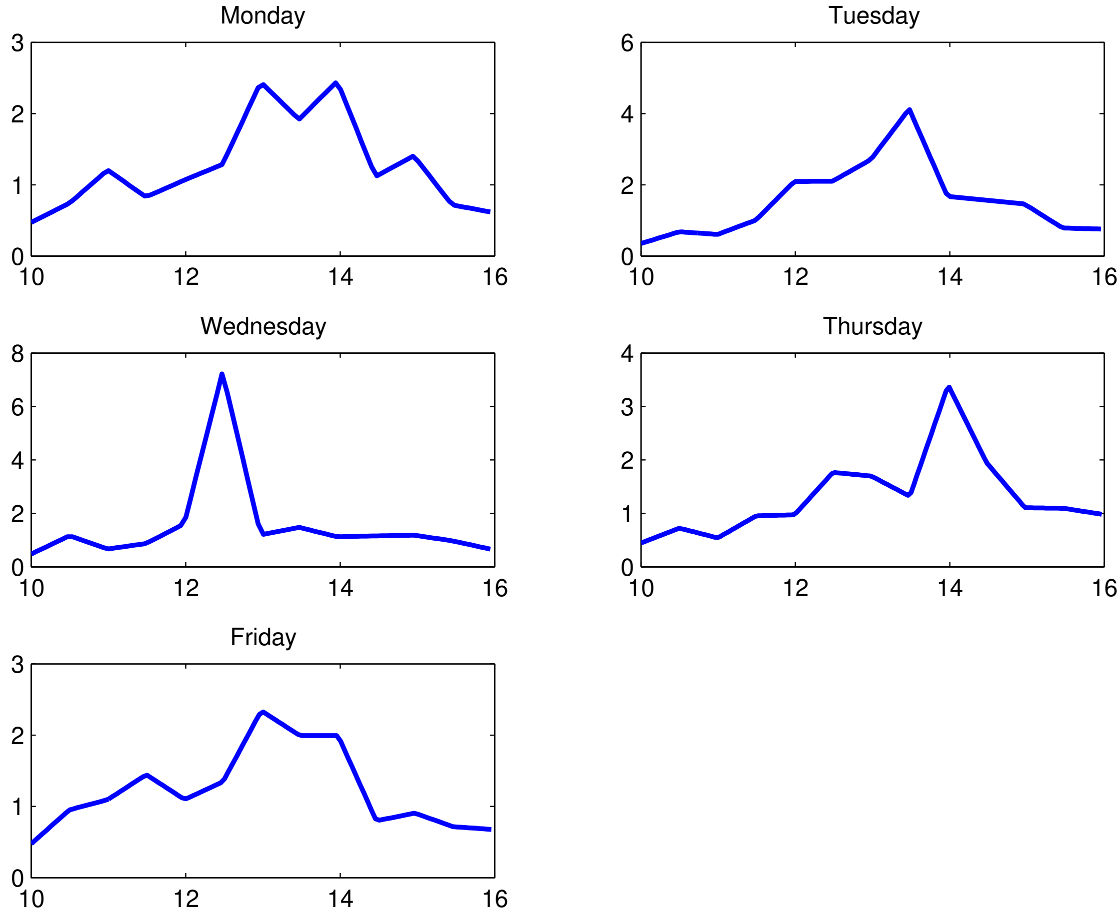

Figure 5.3.: Time-of-the-day function for IBM price durations

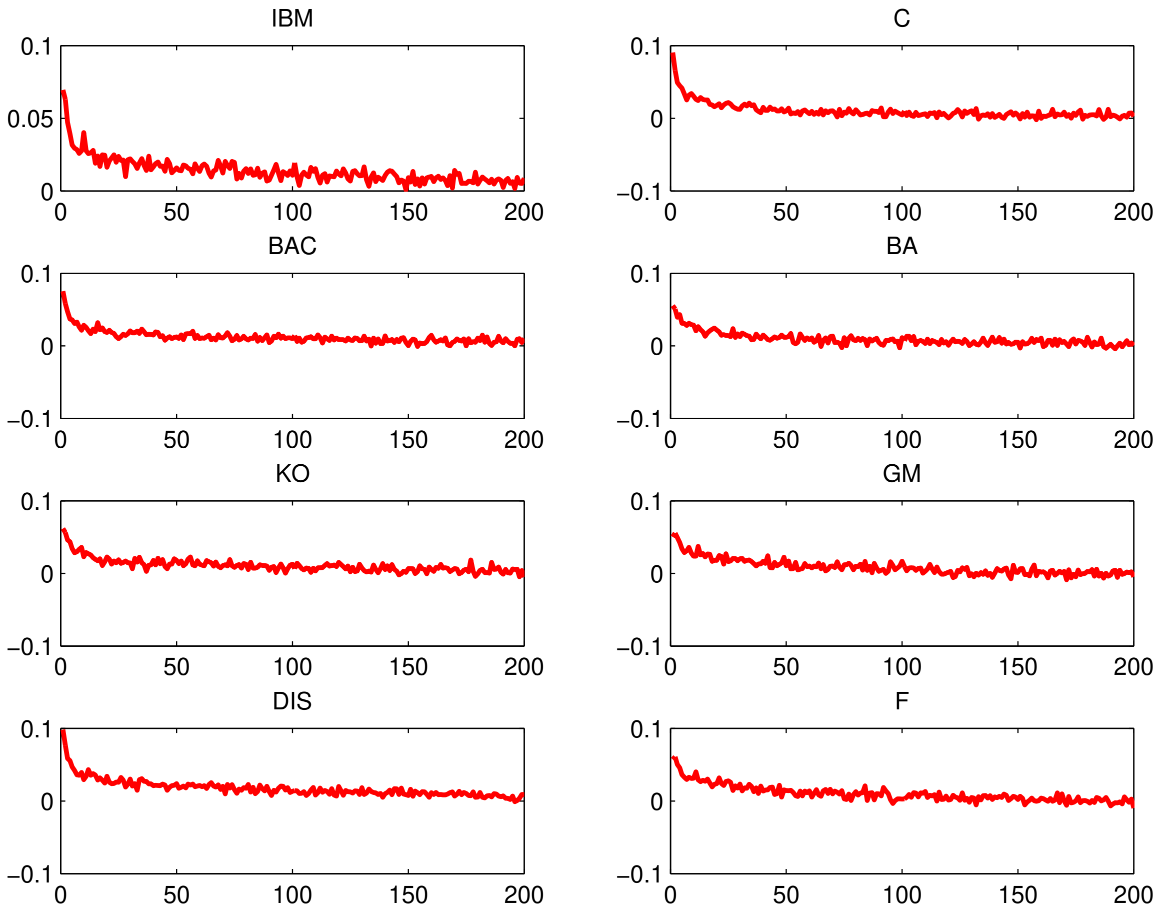

Figure 5.4.: Autocorrelation functions of adjusted trade duration data for the eight stocks

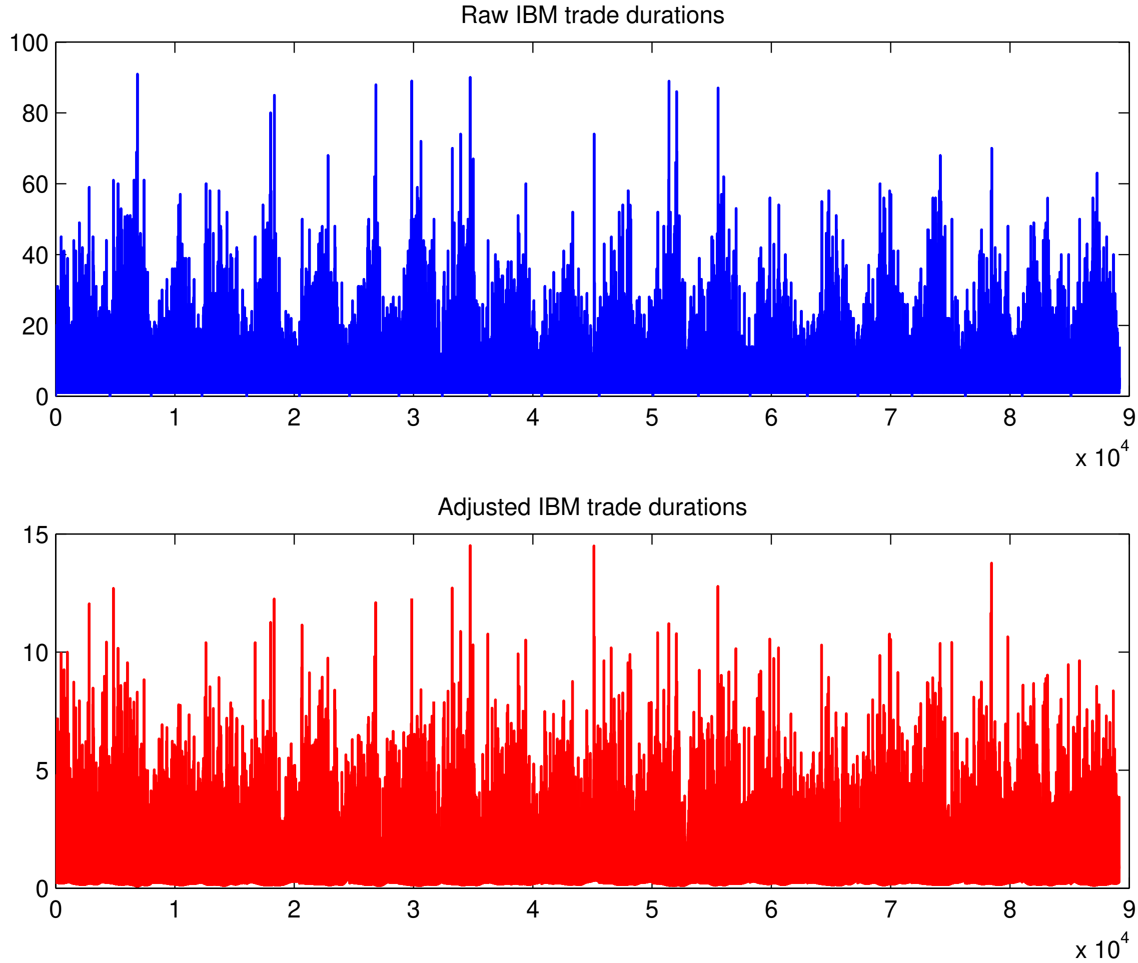

Figure 5.5.: Plot of raw and adjusted IBM trade durations

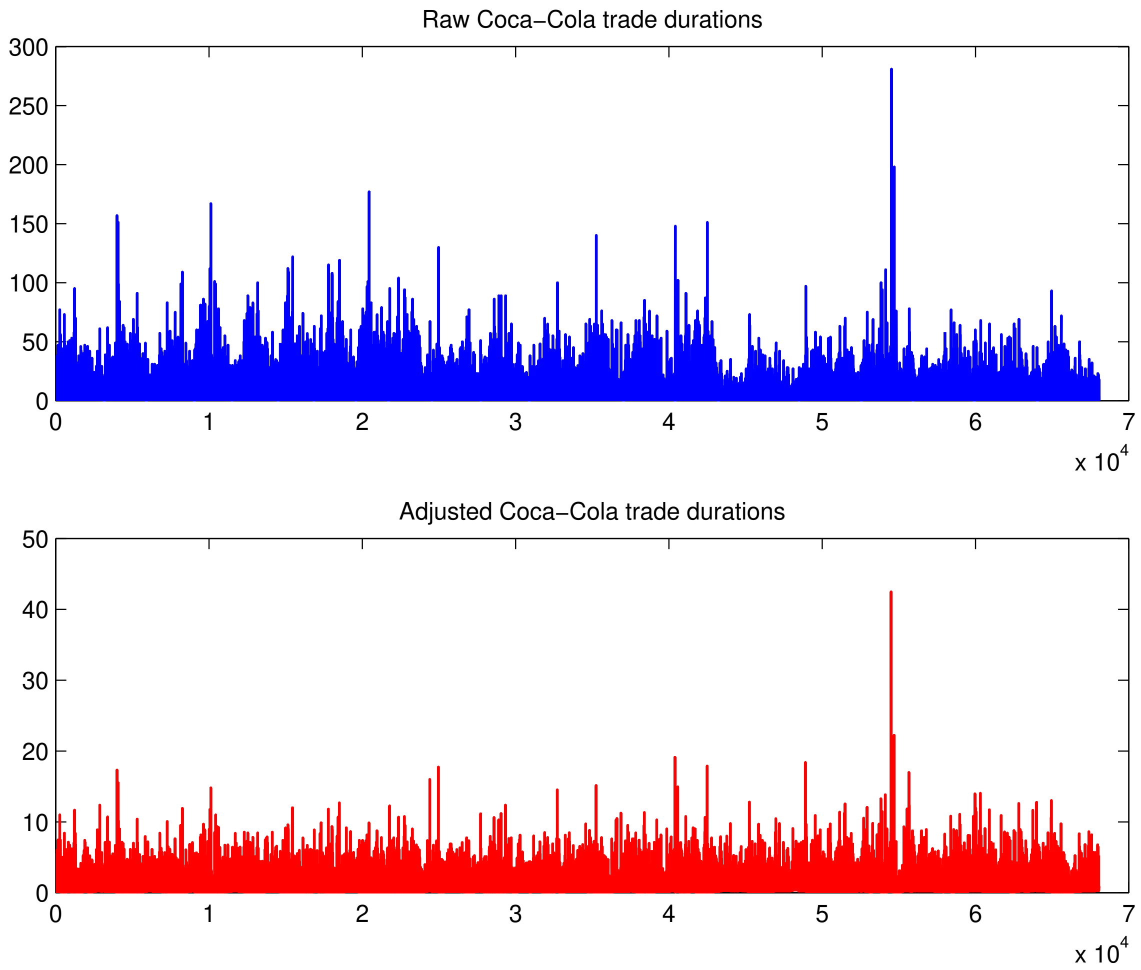

Figure 5.6.: Plot of raw and adjusted Coca-Cola trade durations

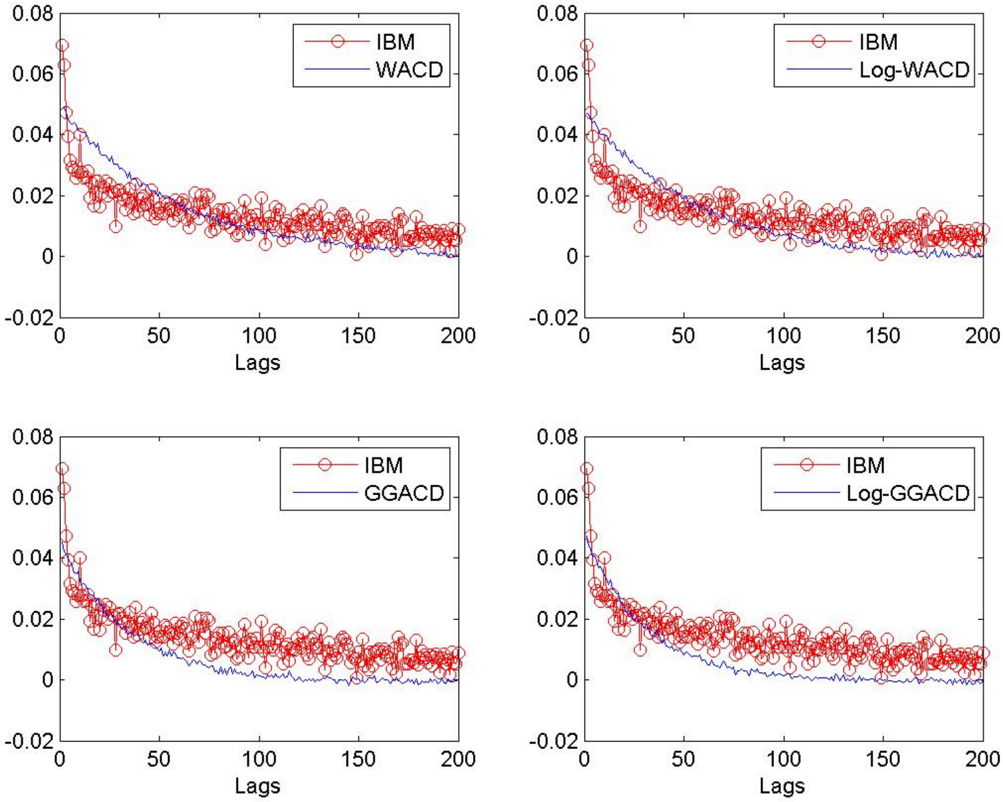

Figure 5.7.: Autocorrelation functions of IBM trade durations and simulated data sets corresponding to WACD, Log-WACD, GGACD, and Log-GGACD specifications. The parameters used for the simulation are set equal to their estimated value for the IBM data.

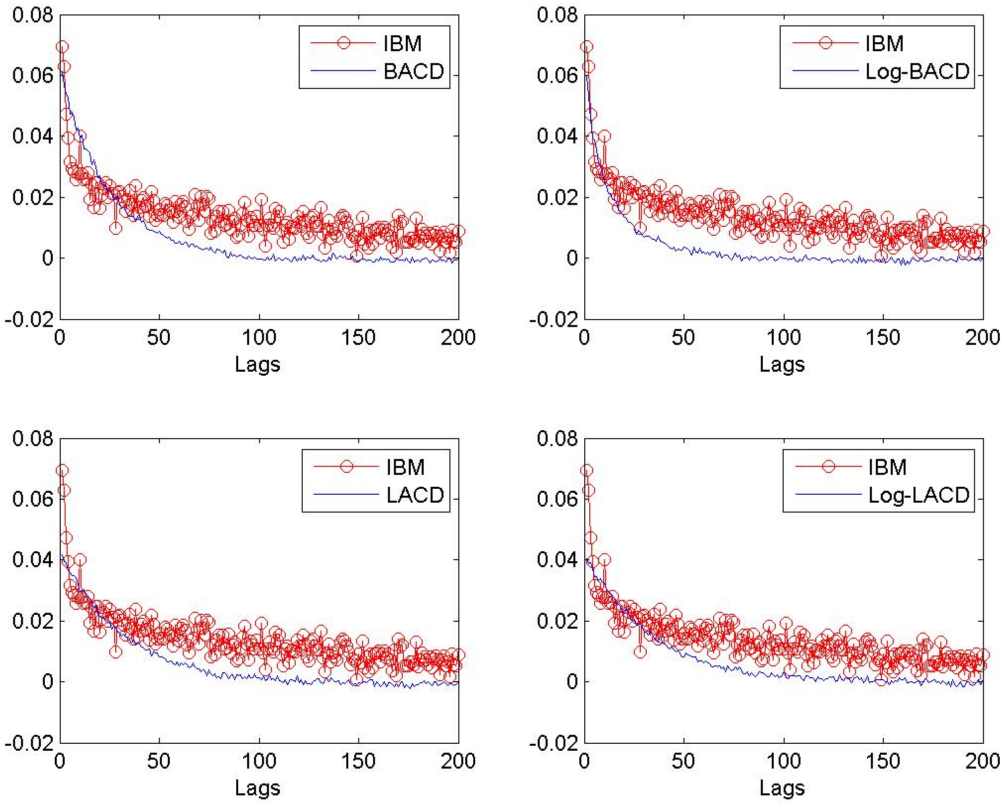

Figure 5.8.: Autocorrelation functions of IBM trade durations and simulated data sets corresponding to BACD, Log-BACD, LACD, and Log-LACD specifications. The parameters used for the simulation are set equal to their estimated value for the IBM data.

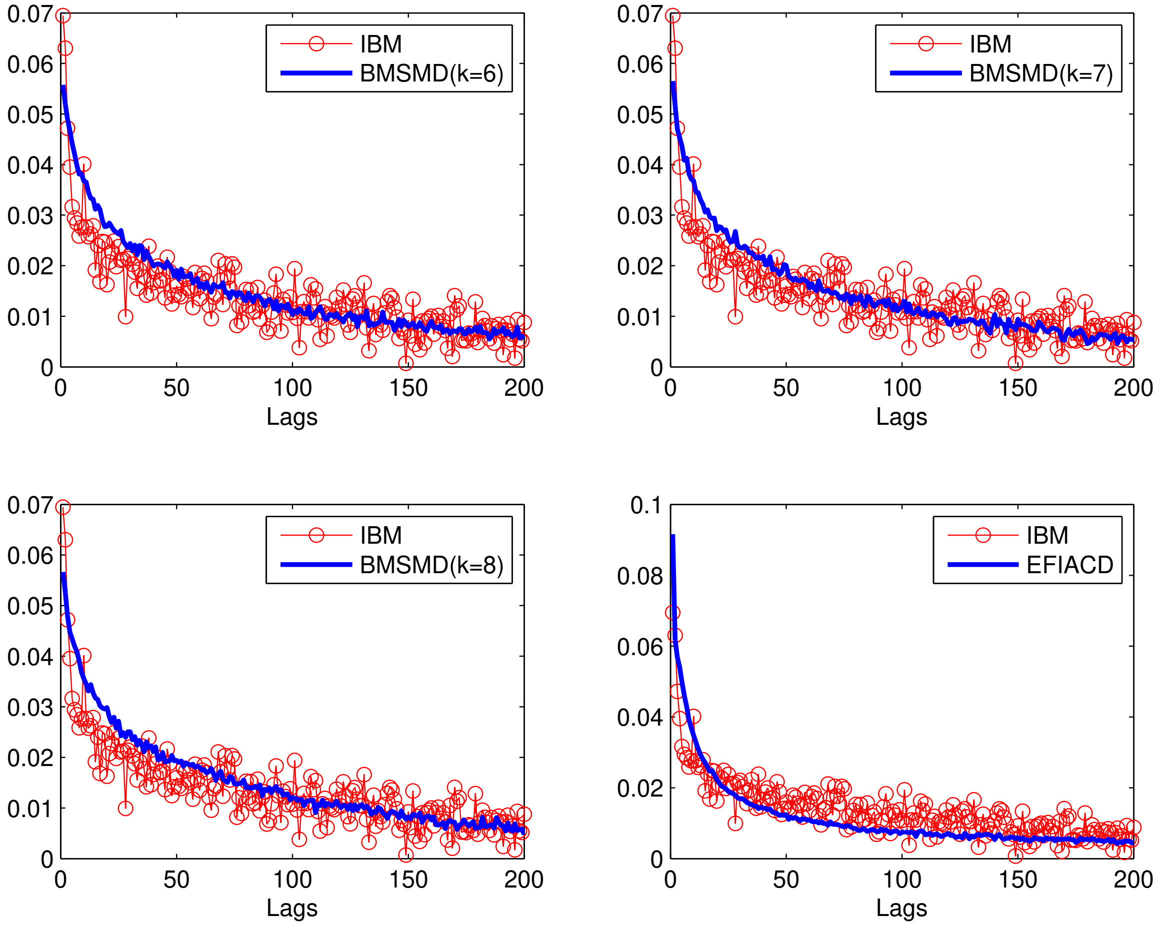

Figure 5.9.: Autocorrelation functions of IBM trade durations and simulated data sets corresponding to EFIACD specification and BMSMD specification with different intensity components (k). The parameters used for the simulation are set equal to their estimated value for the IBM data. Note that BMSMD stands for Binomial MSMD.

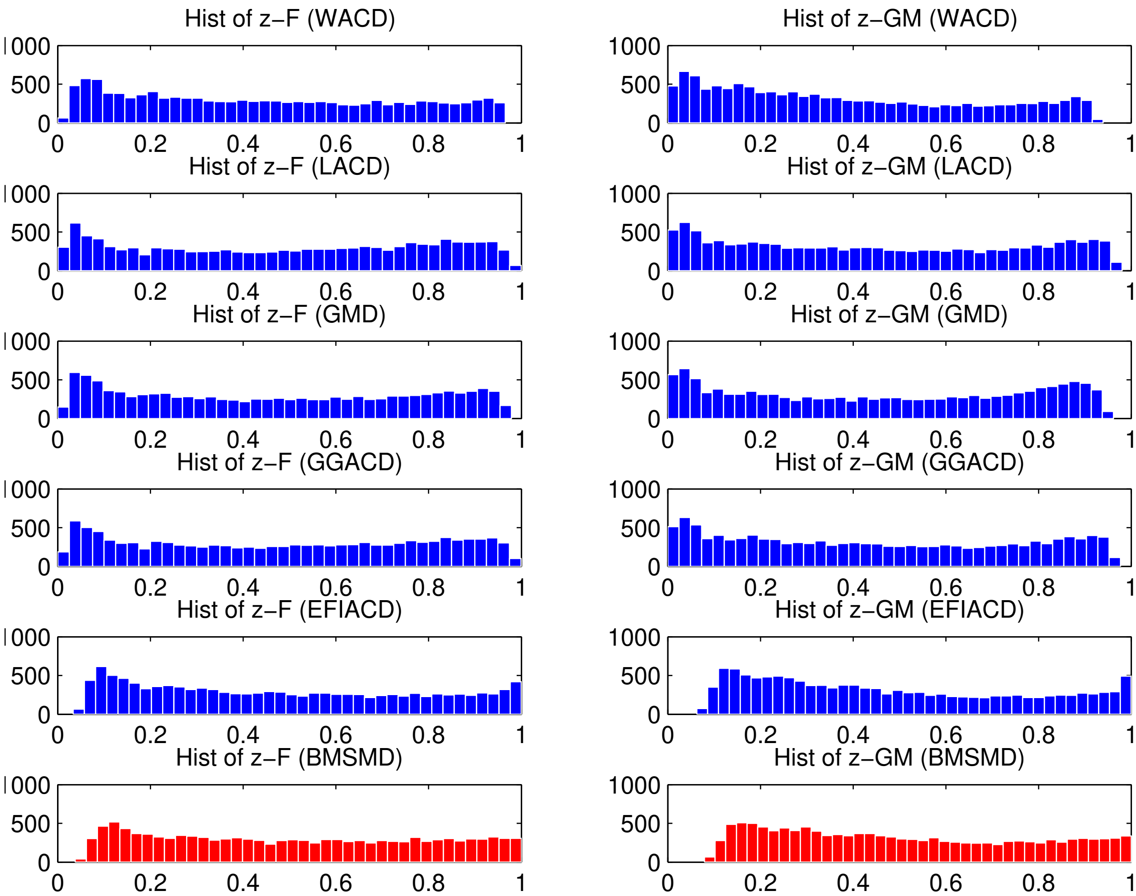

Figure 5.10.: Histograms of the Probability Integral Transforms for Ford and General Motors Trade Durations (In-sample). Note that BMSMD stands for Binomial MSMD.

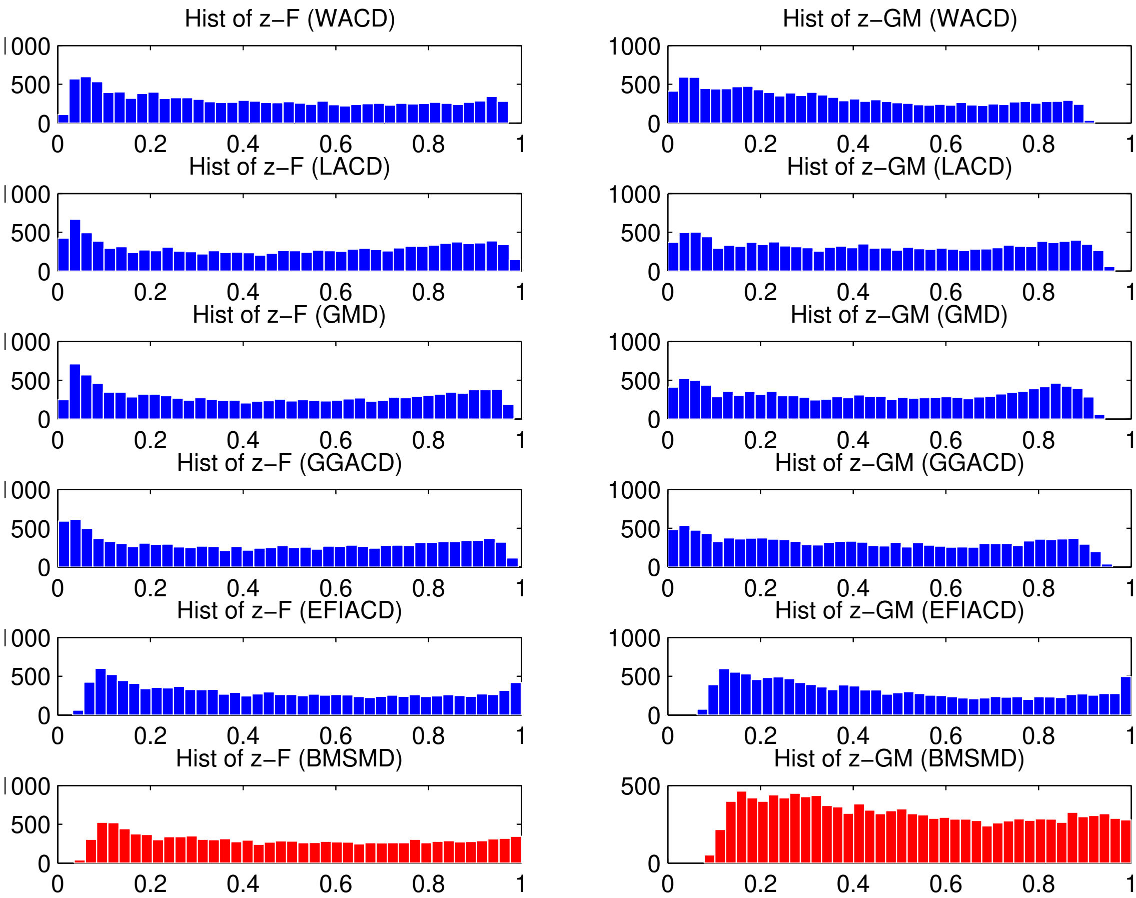

Figure 5.11.: Histograms of the Probability Integral Transforms for Ford and General Motors Trade Durations (Out-of-sample). Note that BMSMD stands for Binomial MSMD.

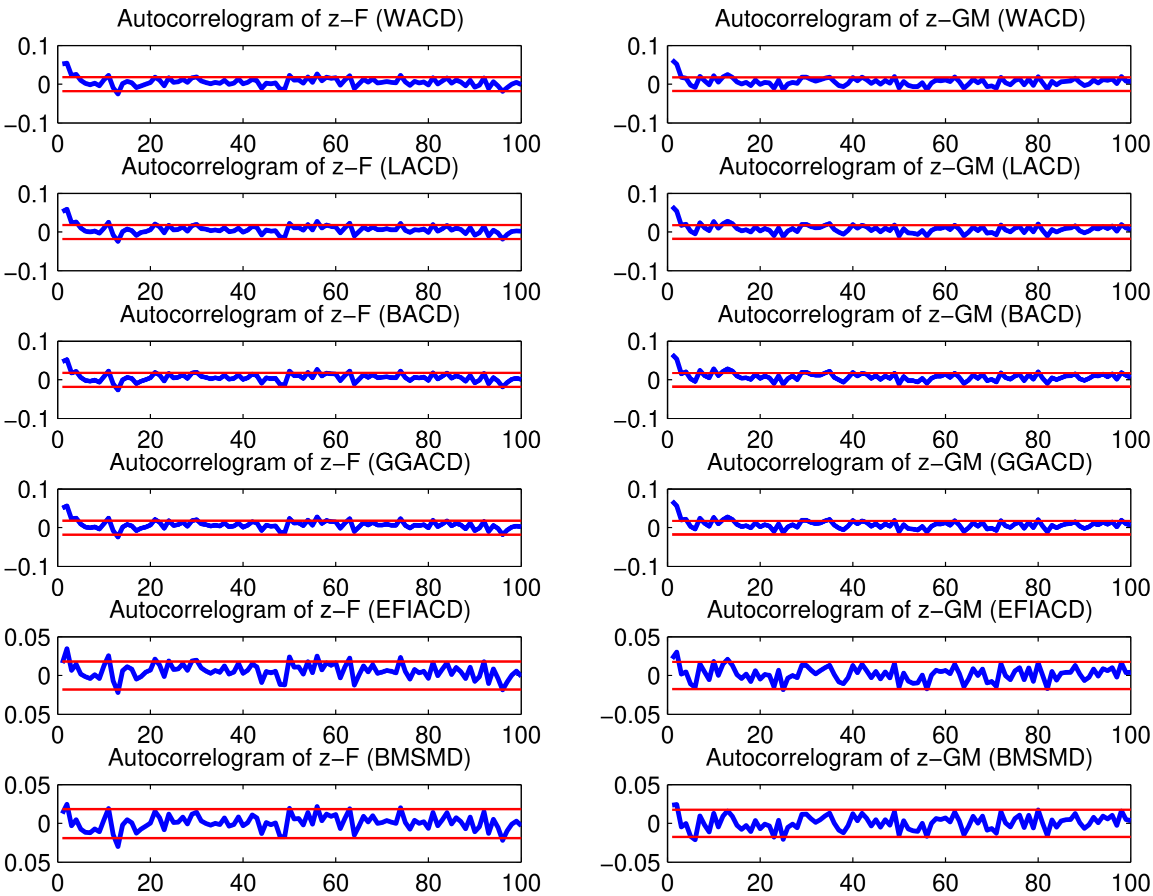

Figure 5.12.: z-Correlograms for Ford and General Motors Trade Durations (In-sample). Note that BMSMD stands for Binomial MSMD.

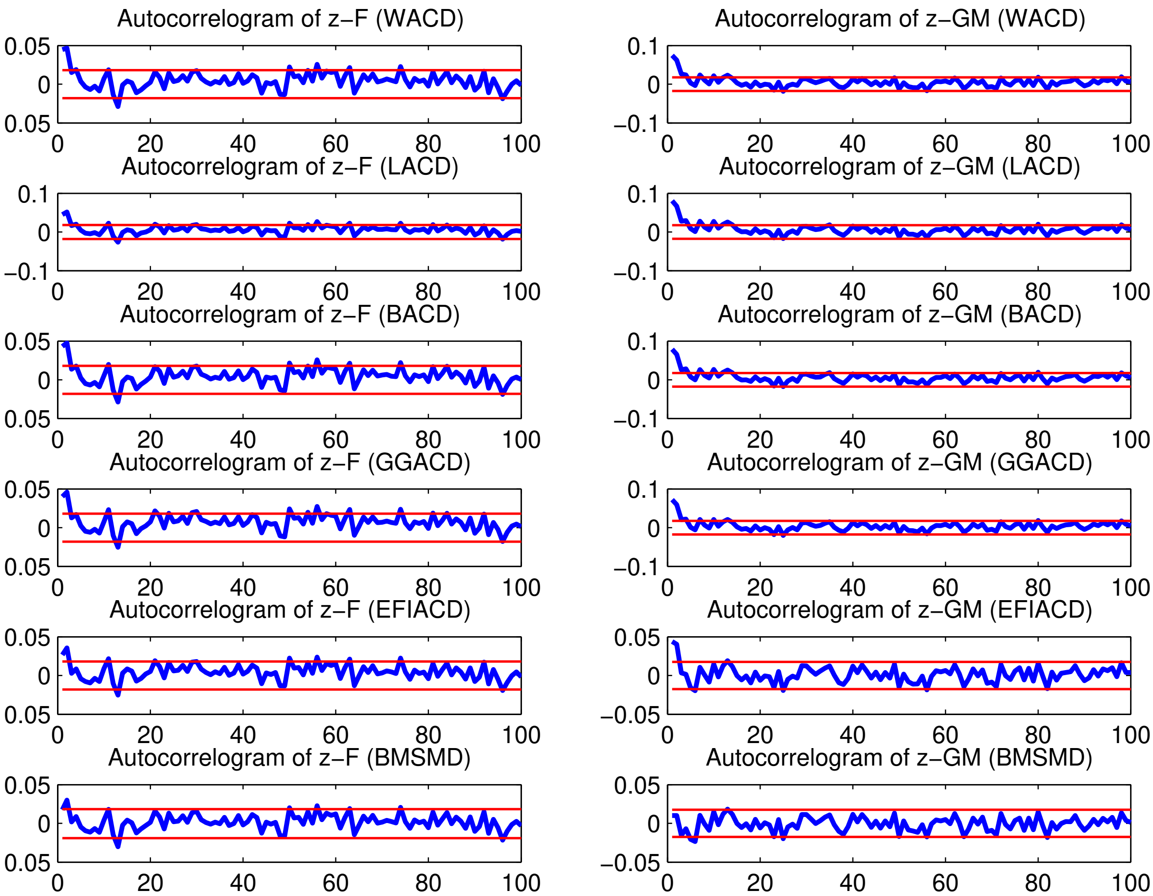

Figure 5.13.: z-Correlograms for Ford and General Motors Trade Durations (Out-of-sample). Note that BMSMD stands for Binomial MSMD.

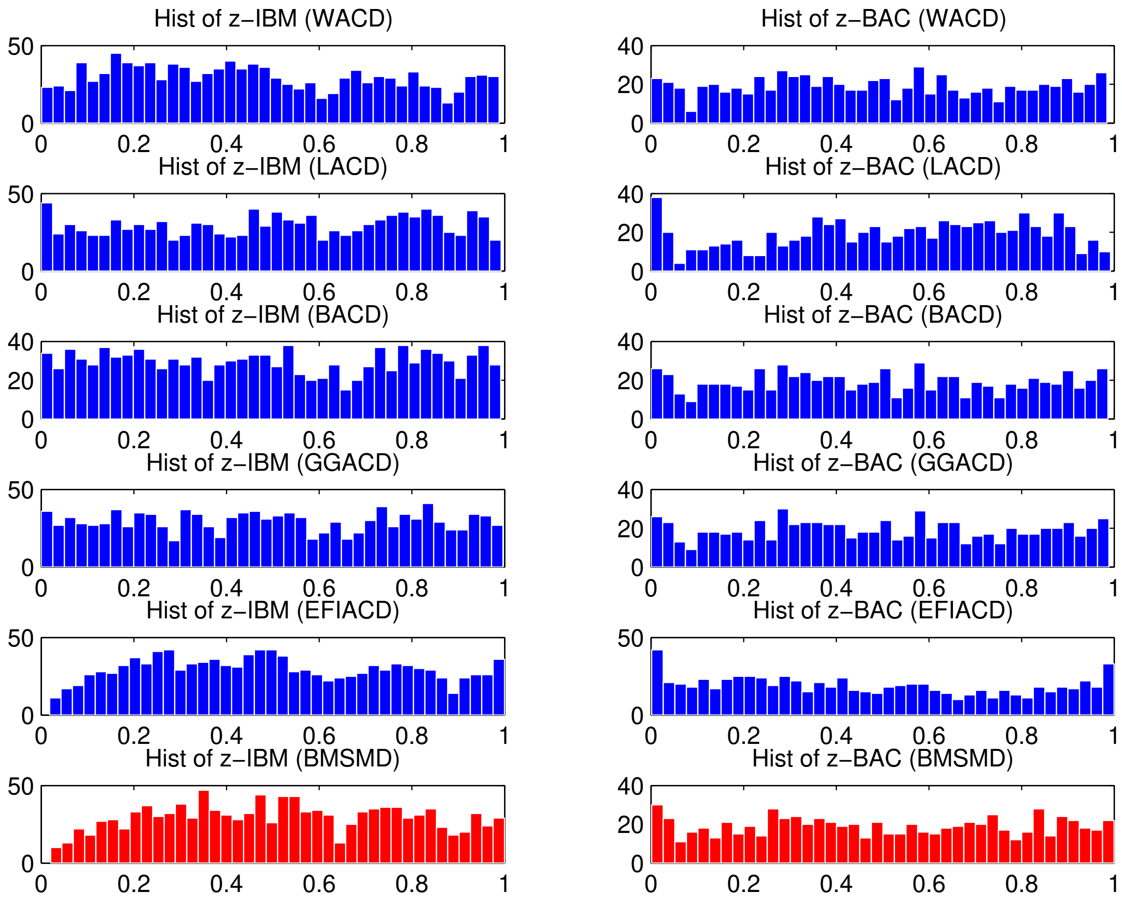

Figure 5.14.: Histograms of the Probability Integral Transforms for IBM and BAC Price Durations (In-sample). Note that BMSMD stands for Binomial MSMD.

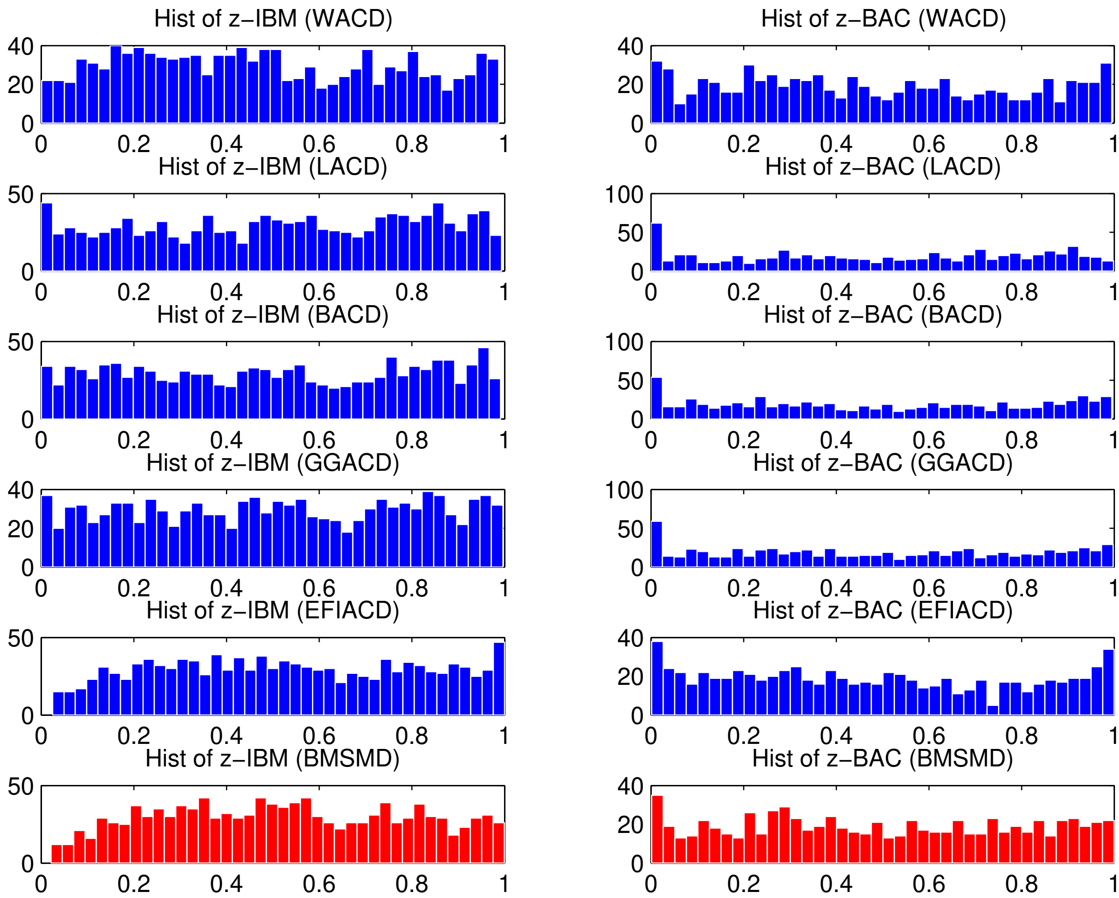

Figure 5.15.: Histograms of the Probability Integral Transforms for IBM and BAC Price Durations (Out-of-sample). Note that BMSMD stands for Binomial MSMD.

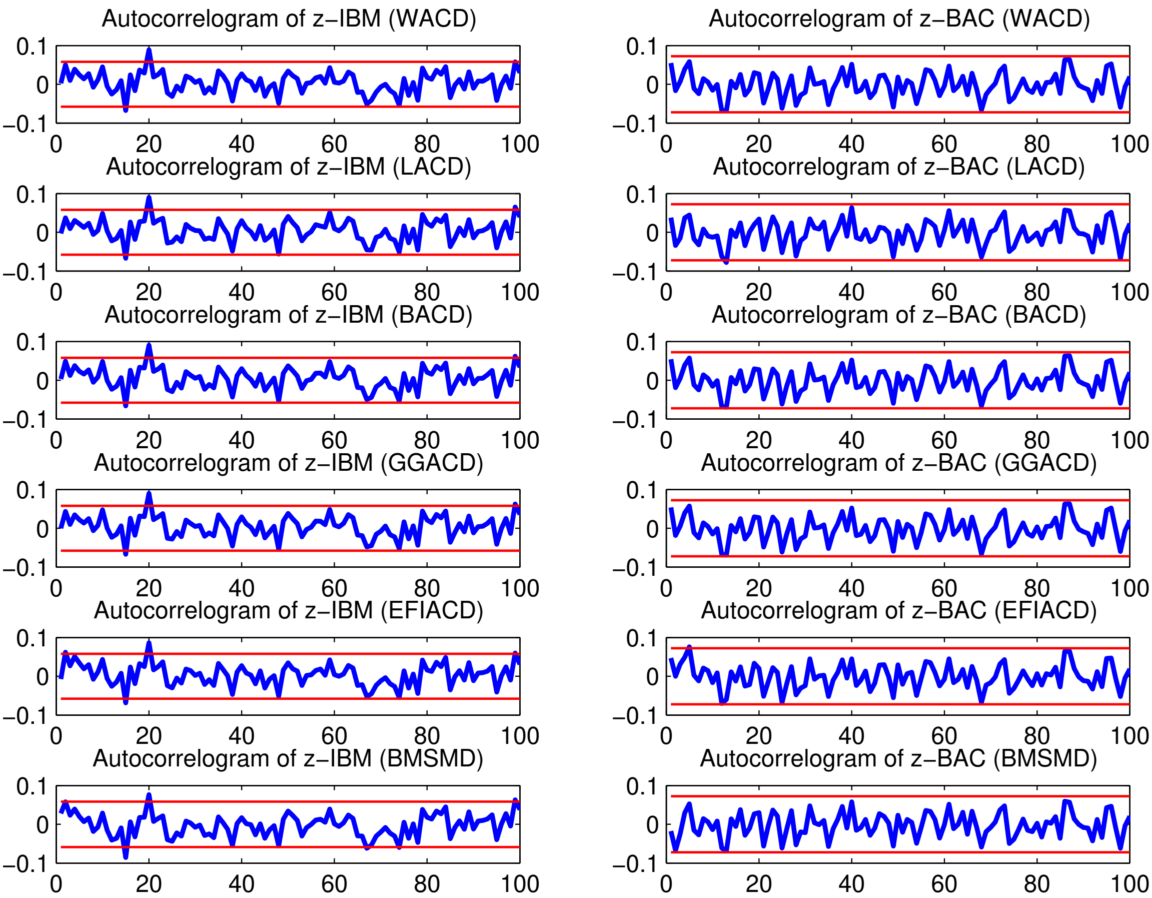

Figure 5.16.: z-Correlograms for IBM and BAC Price Durations (In-sample). Note that BMSMD stands for Binomial MSMD.

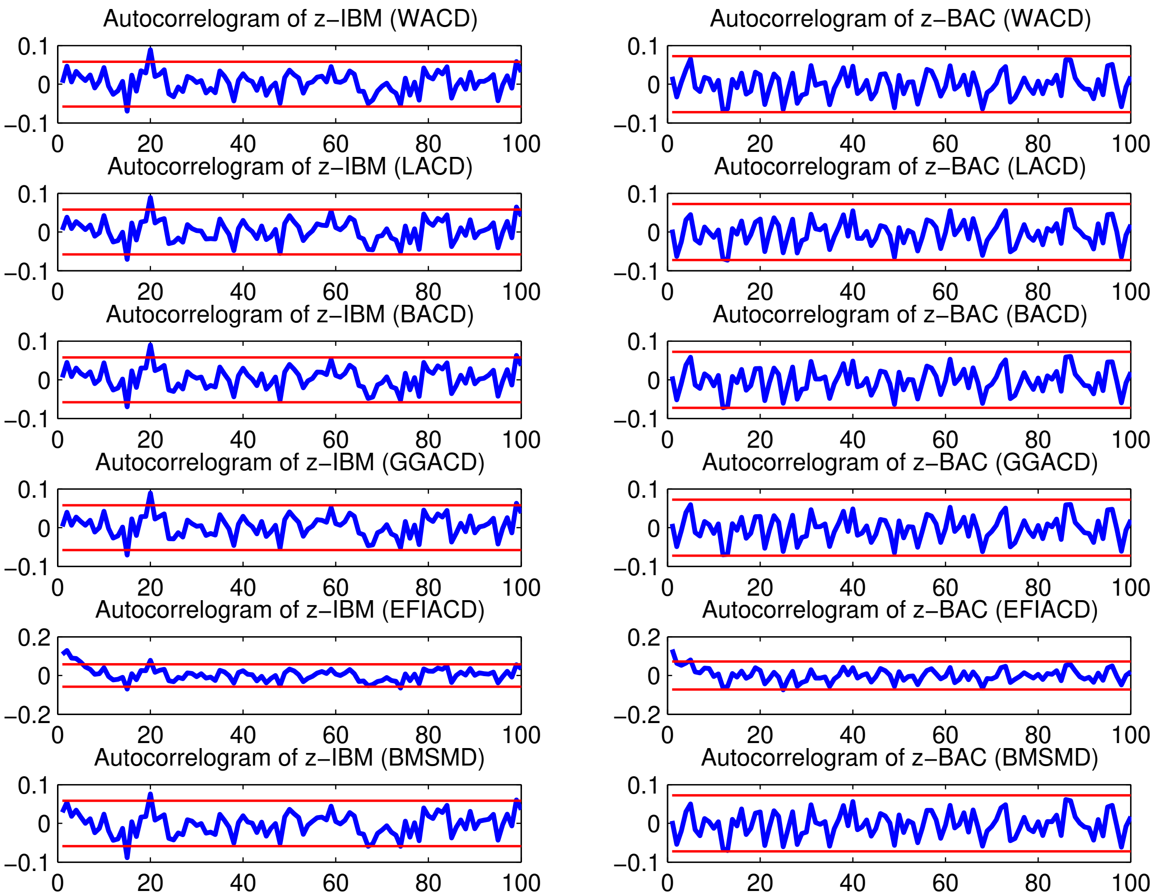

Figure 5.17.: z-Correlograms for IBM and BAC Price Durations (Out-of-sample). Note that BMSMD stands for Binomial MSMD.

6. Modeling Financial Duration Data Using Alternative Markov Switching Multifractal Duration Models

6.1. Introduction

The mixture of distributions is not new in finance and goes back to Clark (1973). This idea of mixing distribution has also been used by Tauchen and Pitts (1983) to explain the positive association between daily price variability and the trading volume. The mixture of distribution models have been applied for modeling financial duration data and achieved successful results in the literature. For instance, the SVD model of Ghysels et al. (2004) is obtained from a combination of a gamma distribution and an exponential distribution. The SCD model of Bauwens and Veredas (2004) combines a Lognormal distribution and a Weibull (or gamma) distribution. De Luca and Gallo (2004) also used a mixture of two exponential distributions for modeling intra-daily durations. In this chapter we propose a mixture of generalized gamma and Burr representations for financial durations. In fact, both representations are generalized versions of the MSMD model.

Empirical studies on financial duration data give evidence that flexible distributional assumptions as such generalized gamma, Burr or Lognormal for the innovations in the ACD models provide a satisfactory fit for the data and contribute a lot to the forecast performance of the models. In line with this, we find that it would be interesting to explore generalized gamma and Burr distributional assumptions for innovations in the MSMD model in order to verify whether a better fit can be achieved. We compare the ability of the new models to fit financial duration data to that of the standard MSMD model of Chen et al. (2013) via the likelihood ratio test, the Akaike and the Bayesian information criterion.

We find that the mixture of generalized gamma, Burr and their particular cases (Weibull, gamma and exponential) are able to capture long memory, heavy tails, and clustering effects simultaneously, and thus, may provide relatively accurate forecasts. The main finding here is that the functional form of the MSMD models is more determinant for obtaining accurate forecasts than the distributional assumptions for innovations.

The rest of the chapter is organized as follows. In Section 6.2 we introduce the mixture of generalized gamma and Burr representations for financial durations. Section 6.3 presents statistical properties of the models. The estimation procedures and model selection criteria are described in Section 6.4 and 6.5, respectively. Section 6.6 illustrates the empirical application and we conclude in Section 6.7.

6.2. Alternative Markov Switching Multifractal Duration Models

The MSMD model in its generalized form is defined as

$$\begin{cases} x_t &= \dfrac{\zeta_t}{\lambda_t}, \\ \lambda(M_t) &= \bar{\lambda} \displaystyle\prod_{i=1}^{k} M_t^{(i)}, \end{cases} \tag{6.1}$$

where x_t denotes the duration between two consecutive financial events occurring at times T_{t-1} and T_t and ζ_t is an *i.i.d.* unit-mean innovation. In principle, any distribution with positive support can be used for the innovation. The dynamic processes governing $\lambda(M_t)$ are defined as in sec. 3.4.1.

6.2.1. Mixture of Generalized Gamma Distribution

Here we assume that the innovation ζ_t is generalized gamma distributed and normalize the distribution so that $\mathbb{E}[\zeta_t] = 1$. The corresponding probability density function (pdf), GG, of ζ_t is given by

$$GG(\zeta_t; \delta, \alpha) = \begin{cases} \dfrac{\delta \zeta_t^{\delta\alpha - 1}}{\beta^{\delta\alpha}\Gamma(\alpha)} \exp\left[-\left(\dfrac{\zeta_t}{\beta}\right)^{\delta}\right], & \text{if } \zeta > 0 \\ 0, & \text{otherwise}, \end{cases} \tag{6.2}$$

where α and δ are shape parameters, $\beta = \dfrac{\Gamma(\alpha)}{\Gamma(\alpha + \frac{1}{\delta})}$ represents the scale parameter and $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt. \tag{6.3}$$

Given the pdf of ζ in *eq.* (6.2) and the fact that ζ_t is a monotonic transformation of x_t, we can easily obtain the pdf, p, of x_t by applying the change of variables technique as previously done in sec. 5.2.1. We have

$$p(x_t; \boldsymbol{\phi}) = \frac{\delta x_t^{\delta\alpha - 1}}{\theta_t^{\delta\alpha}\Gamma(\alpha)} \exp\left[-\left(\frac{x_t}{\theta_t}\right)^{\delta}\right] \tag{6.4}$$

where $\boldsymbol{\phi}$ is the parameter vector and $\theta_t = \dfrac{\beta}{\lambda_t}$ represents the time-varying scale parameter. We can rewrite the mixture density in *eq.* (6.4) as

$$p(x_t | \mathfrak{I}_{t-1}; \boldsymbol{\phi}) = \sum_{i=1}^{n} w_t^i p_i(x_t | M_{t-1} = m^i), \tag{6.5}$$

where

1. $\Im_{t-1}$ denotes the information set available at time $t-1$,
2. $\boldsymbol{\phi} = (m_0, \bar{\lambda}, b, \gamma_1, \delta, \alpha)$ is the parameter vector of the model,
3. n is the number of mixture density components; with $n = 2^k$, where k is the number of intensity components in the model,
4. $w_t^i = Pr(M_t = m^i | x_1, \ldots, x_t)$ represents the ith mixture weight, satisfies $w_t^i \geq 0$, $\sum_{i=1}^{n} w_t^i = 1$, and can be calculated as in *eq.* (2.37)
5. $p_i(x_t | M_{t-1} = m^i)$ is a conditional generalized gamma density given the intensity components or multipliers and is given by

$$p_i(x_t | M_{t-1} = m^i) = \frac{\delta x_t^{\delta\alpha - 1}}{\theta_i^{\delta\alpha}\Gamma(\alpha)} \exp\left[-\left(\frac{x_t}{\theta_i}\right)^{\delta}\right], \tag{6.6}$$

where θ_i which is given by

$$\theta_i = \frac{\beta}{\lambda(m^i)}, \tag{6.7}$$

represents the state-dependent scale parameter.

6.2.2. The Mixture of Burr Distribution

Assuming that the innovation (ζ) is Burr distributed and normalizing the distribution so that $\mathbb{E}[\zeta_t] = 1$, we obtain the pdf that is given by

$$B(\zeta_t | \tau, \kappa) = \frac{\kappa c^{\kappa} \zeta_t^{\kappa - 1}}{(1 + \tau c^{\kappa} x_t^{\kappa})^{1/\tau + 1}}, \tag{6.8}$$

where $0 < \tau < \kappa$ and $c = \dfrac{\Gamma(1 + \frac{1}{\kappa})\Gamma(\frac{1}{\tau} - \frac{1}{\kappa})}{\tau^{(1+\frac{1}{\kappa})}\Gamma(\frac{1}{\tau} + 1)}$.

By applying the change of variables rule we obtain the pdf, p, of financial duration, x_t, as

$$p(x_t; \Omega) = \frac{\kappa \varphi_t^{\kappa} x_t^{\kappa - 1}}{(1 + \tau \varphi_t^{\kappa} x_t^{\kappa})^{1/\tau + 1}}, \tag{6.9}$$

where $\Omega = (m_0, \bar{\lambda}, b, \gamma_1, \kappa, \tau)$ is the parameter vector of the model and φ_t is given by

$$\varphi_t = c\lambda_t. \tag{6.10}$$

The conditional density of x_t given the information set, $\Im_{t-1}$, available at time $t-1$ is then

$$p(x_t | \Im_{t-1}; \Omega) = \frac{\kappa \varphi_t^{\kappa} x_t^{\kappa - 1}}{(1 + \tau \varphi_t^{\kappa} x_t^{\kappa})^{1/\tau + 1}}. \tag{6.11}$$

We can rewrite the pdf in *eq.* (6.11) as we previously did in *eq.* (6.5) and in this case the conditional density of the duration given the latent intensity component M_{t-1}, $p_i(x_t|M_{t-1} = m^i)$, becomes

$$p_i(x_t|M_{t-1} = m^i) = \frac{\kappa\varphi_i^{\kappa} x_t^{\kappa-1}}{(1 + \tau\varphi_i^{\kappa} x_t^{\kappa})^{1/\tau+1}}, \tag{6.12}$$

where $0 < \tau < \kappa$, and φ_i is given by

$$\varphi_i = c\lambda(m^i). \tag{6.13}$$

Unlike the generalized gamma distribution, the Burr distribution is rarely used in duration analysis. This is due to the fact that the Burr distribution requires some restrictions on parameters in order to ensure finite moments. However, compared to exponential or Weibull distribution, the Burr distribution function depends on two parameters that enable the distribution to be more flexible than the two former. The mixture of Burr distribution includes the mixture of Weibull ($\tau \to 0$), and the mixture of exponential ($\tau \to 0 \quad \text{and} \quad \kappa = 1$)

6.3. Statistical Properties

6.3.1. Moments

The moments of MSMD models depend on the moments of ζ_t. By the definition the unconditional expectation (μ_x) and variance (σ_x^2) of x_t are:

$$\mu_x = \mathbb{E}\left[\frac{\zeta_t}{\lambda_t}\right] = \mathbb{E}\left[\lambda_t^{-1}\right], \tag{6.14}$$

and,

$$\begin{aligned} \sigma_x^2 &= \mathbb{E}\left[\left(\frac{\zeta_t}{\lambda_t}\right)^2\right] - \left[\mathbb{E}\left(\frac{\zeta_t}{\lambda_t}\right)\right]^2 \\ &= \mathbb{E}\left(\zeta_t^2\right)\mathbb{E}\left[\left(\frac{1}{\lambda_t}\right)^2\right] - \left[\mathbb{E}\left(\frac{1}{\lambda_t}\right)\right]^2. \end{aligned} \tag{6.15}$$

If for instance, ζ is generalized gamma distributed, then the second moment $\mathbb{E}\left(\zeta_t^2\right)$ is given by

$$\mathbb{E}\left(\zeta_t^2\right) = \frac{\Gamma(\alpha)\Gamma(\alpha + 2/\delta)}{\Gamma^2(\alpha + 1/\delta)}. \tag{6.16}$$

Following the argumentation of Chen et al. (2013), it is clear that the duration processes in the alternative MSMD models are also stationary, ergodic and possess finite moments. The clear proofs of this assertion are:

1. The transition probability matrix in *eq.* (3.24) describes the dynamic process of the intensity components or multipliers M_t^i. Given that $\gamma_i > 0$, it is clear that the processes M_t^i are stationary and ergodic.

2. By the construction of the Markov switching process, it is assumed that M_t^i are independent across k and t and also independent of ξ_t (cf. Calvet and Fisher, 2004a).

3. In sum, it is obvious that the vector process $\left(\xi_t, M_t^i, \ldots, M_t^k\right)$ that determines the duration process in the MSMD models is stationary and ergodic, and thus, the duration process as well.

4. By assuming that the latent multipliers are drawn from Binomial distribution taking m_0 in the high state and $2 - m_0$ in the low state, with $m_0 \in (0, 2]$, one can easily obtain an upper bound as $\lambda_t = \bar{\lambda} m_0^k$, if all intensity components or multipliers are in the high state and a lower bound $\lambda_t = \bar{\lambda}(2 - m_0)^k$, if all intensity components are in the low state, for the mean intensity λ_t. This implies that the expectation of the mean intensity λ_t exists and is finite, and thus, the expectation of the x_t.

Dispersion

The mixture of generalized gamma and Burr representations for financial durations exhibit over- and underdispersion. Compared to the model of Chen et al. (2013) that can only reproduce overdispersion, the capacity of both models to display over- and underdispersion is an additional asset for our models.

$$\begin{aligned}
\sigma_x^2 - \mu_x^2 &= \mathbb{E}\left[\left(\frac{\zeta_t}{\lambda_t}\right)^2\right] - 2\left[\mathbb{E}\left(\frac{\zeta_t}{\lambda_t}\right)\right]^2 \\
&\geq \mathbb{E}\left(\zeta_t^2\right)\mathbb{E}\left[\left(\frac{1}{\lambda_t}\right)^2\right] - 2\left[\mathbb{E}\left(\frac{1}{\lambda_t}\right)^2\right] \\
&= \underbrace{\left[\mathbb{E}\left(\zeta_t^2\right) - 2\right]}_{=h(\alpha,\delta)}\mathbb{E}\left[\left(\frac{1}{\lambda_t}\right)^2\right].
\end{aligned} \tag{6.17}$$

If $\alpha = 1$ and $0 < \delta \leq 1$ or $0 < \alpha \leq 1$ and $\delta = 1$, $h(\alpha, \delta)$ takes positive values. This means that the difference $\sigma_x^2 - \mu_x^2$ is positive, and thus, the MSMD processes display overdispersion. A special case is when $\alpha = \delta = 1$ and $h(1, 1) = 0$ (exponential case). If $\alpha = 1$ and $\delta > 1$ or $\alpha > 1$ and $\delta = 1$, the function $h(\alpha, \delta)$ takes negative values due to the fact that $\mathbb{E}\left(\zeta_t^2\right)$ approaches 1. In this case, the MSMD processes display underdispersion.

In sum, it is clear that there exist combinations of α and δ (by the generalized gamma distribution) or κ and τ (by the Burr distribution) that lead to over- or underdispersion. In the next subsec. 6.3.3 we present a simulation study that confirms our results.

6.3.2. Long Memory Feature

Here we concentrate on the autocorrelation function of financial durations and show that the MSM process can mimic the hyperbolic decay in autocorrelation functions exhibited by financial durations. By the definition the autocorrelation function for a range of lags is

$$\rho(l) = \frac{\mathbb{E}\left[(x_t - \mathbb{E}[x_t])(x_{t+l} - \mathbb{E}[x_{t+l}])\right]}{\mathbb{E}\left[(x_t - \mathbb{E}[x_t])^2\right]^{1/2} \mathbb{E}\left[(x_{t+l} - \mathbb{E}[x_{t+l}])^2\right]^{1/2}}. \tag{6.18}$$

From the duration process we know that

$$\mathbb{E}[x_t] = \mathbb{E}[x_{t+l}]. \tag{6.19}$$

We make use of *eq.* (6.19) to obtain the following reduced form of *eq.* (6.18)

$$\rho(l) = \frac{\mathbb{E}(x_t x_{t+l}) - [\mathbb{E}(x_t)]^2}{\sigma^2_{x_t}}, \tag{6.20}$$

where $\sigma^2_{x_t}$ is calculated as in *eq.* (6.15).

Following Calvet and Fisher (2004a) we obtain the formula for the autocorrelation function of the durations in the models

$$\rho(l) = \frac{\prod_{i=1}^{k}[1 + \eta(1-\gamma_i)^l] - 1}{\varsigma(1+\eta)^k - 1}, \tag{6.21}$$

where $\varsigma = \mathbb{E}[\zeta_t^2]$, and $\eta = \mathbb{E}(M^{-2})[\mathbb{E}(M^{-1})]^{-2} - 1$.

Define $\varrho = \log_b\left(\mathbb{E}(M)/[\mathbb{E}(M^{1/2})]^2\right)$ and consider two arbitrary numbers α_1 and α_2 in the open interval $(0, 1)$ such that $\alpha_1 < \alpha_2$. The set of integers $\Omega_k = \{l : \alpha_1 \log_b(b^k) \leq \log_b l \leq \alpha_2 \log_b(b^k)\}$ contains a considerable number of intermediate lags.[1]

Proposition 1. *The autocorrelation of durations fulfills*

$$\sup_{l \in \Omega_k} \left\| \frac{\ln \rho(l)}{\ln l^{-\varrho}} - 1 \right\| \to 0 \qquad \text{as} \quad k \to +\infty. \tag{6.22}$$

The proposition 1 evidences the capacity of the MSMD processes to reproduce the long memory exhibited by financial duration data. The presence of the high persistence in IBM trade durations has been first pointed out by Engle and Russell (1998) in their seminal paper. Its relevance has been well-documented in details in the volatility literature for a long time and nowadays it is unthought of introducing a model that cannot mimic this feature. In empirical financial duration literature long memory has been only modeled by Jasiak (1998) and later by Deo et al. (2010). So, our MSMD models complete the list of the long memory financial duration models.

6.3.3. Numerical Simulations

Due to the fact that the unconditional moments and autocovariances cannot be computed analytically, we conduct numerical simulations with several sets of parameters (cf. Table 6.1) to gain insight into each model specification. We concentrate on the empirical dispersion, the skewness,

[1] cf. Calvet and Fisher (2004a) for the proof and more information.

and the kurtosis obtained in each model. Table 6.1 presents average results for the dispersion, the skewness, and the kurtosis based on Monte Carlo simulations (100 samples of size 15000). Different parameter scenarios lead to different average values for the dispersion, the skewness, and the kurtosis. It is clear that the mixture of gamma, generalized gamma, Weibull, and Burr models can not only fit data characterized by overdispersion ($\sigma_x/\mu_x > 1$) but also underdispersion[2] ($\sigma_x/\mu_x < 1$). One can also see that by varying the parameter values the empirical moments for skewness and kurtosis also change. These results confirm the fact that the new models offer a lot of flexibility which is missed in the mixture of exponential model proposed by Chen et al. (2013). With simulated samples we check the autocorrelation functions and we find that they exhibit long memory feature for various parameter scenarios. *Fig.* 6.1 depicts the autocorrelation functions in each model specifications (generalized gamma, gamma, Weibull, and Burr) compared to that of exponential. We also see that the shape of the autocorrelation functions for different distributions for innovations remains almost identical (curves shift left). The shift of the autocorrelation functions comes from the term ς in *eq.* (6.21). This suggests that the ability of the MSMD models to reproduce long memory patterns come more from the MSM process than the distributional assumption for the innovation.

6.4. ML Estimation

The estimation of the alternative MSMD models can easily be performed by the ML estimation procedures we adopted for Chen et al. (2013)'s MSMD model in sec. 5.3.2.

6.4.1. Small-Sample Properties

We assess the small-sample properties of the ML estimator by conducting Monte Carlo simulations. The design of the simulation is as follows: For each model specification we use $k = 7$ (the number of intensity components) and Binomial distribution for multipliers. The choice of k is motivated by the research results obtained by Chen et al. (2013); Segnon and Lux (2012). The simulation requires four basic parameters for the multifractal process and one or two additional parameters depending on the distributional assumption for the innovation in the model. The basic parameters are: the Binomial value m_0, the unconditional intensity $\bar{\lambda}$, the frequency growth rate b, and the high-frequency switching probability γ_1. The additional parameters are: α for a gamma distribution, δ for a Weibull, α and δ for a generalized gamma, and τ and κ for a Burr distribution. All simulations use $m_0 = 1.2$, $\bar{\lambda} = 1$, $b = 2$, $\gamma = 0.5$, $\delta = 0.8$, $\alpha = 1.2$, $\tau = 0.5$ and $\kappa = 1.2$ and we consider two sample sizes: $T = 5000$, and $T = 10000$, and there are 100 replications for each sample size.

For each model specification and for each simulation the ML estimation provides a set of parameter estimates. We compute the biases, standard errors (SE) and the root mean-squared errors (RMSE) and the results are reported in Table 6.3. As it turned out, for parameter m_0, results are

[2] Volume durations exhibit underdispersion.

almost identical over models in terms of biases, standard errors and root mean squared errors. The estimators $\hat{\hat{\lambda}}$, $\hat{b}$ and $\hat{\gamma}$ in the exponential MSMD model exhibit larger biases and smaller RMSEs. In other models we observe a tendency towards increasing RMSEs and decreasing biases.

In sum, the parameters are well estimated in all models and the decrease in RMSE with sample size in all models is in harmony with $T^{1/2}$ consistency: proceeding from 10,000 to 5,000, the root mean-squared error increases roughly with factors of about $\sqrt{2}$.

6.5. Model Selection Criteria

To address the issue of the model selection we make use of the three widely applied criteria in the literature, namely the likelihood ratio test, the Akaike information criterion (AIC), and the Bayesian information criterion (BIC).

1. *The Likelihood Ratio Test*:
   The likelihood ratio test is the most popular and often used test in the literature due to its simplicity and the opportunity it offers to test two nested models. The test statistic has the following form
   $$LR = -2\ln\left(\frac{L_s}{L_g}\right), \tag{6.23}$$
   where L_s and L_g are the likelihood functions under the null and alternative hypotheses, respectively. The test statistic is asymptotically chi-squared (χ^2) distributed with degrees of freedom equal to $p_g - p_s$, with p_s and p_g the number of free parameters of specified and generalized models under the null and alternative hypotheses, respectively.

2. *The Akaike Information Criterion (AIC)*:
   The AIC proposed by Akaike (1974) is defined as
   $$AIC = -2\ln(L) + 2p, \tag{6.24}$$
   where $\ln(L)$ is the log-likelihood of the estimated model and p is the number of estimated parameters.

3. *The Schwarz or Bayesian Information Criterion (SIC or BIC)*
   The concept of the BIC has been developed by Schwarz (1978) in a Bayesian framework and is defined as
   $$BIC = -2\ln(L) + p\ln(T), \tag{6.25}$$
   where $\ln(L)$ is the log-likelihood of the estimated model, p is the number of estimated parameters and T is the sample size. Here the BIC works under the assumption that the true model exists and is embedded in the set of aspirant models under consideration.

From *eq.* (6.24), it is clear that as the sample size (T) grows, the penalty term ($2p$) remains constant. In other words, AIC does not weight the penalty term conveniently. As results, the AIC can rapidly overfit, and select a model with a larger number of parameters than it must be.

The overfitting problem of the AIC is overcome by the BIC that re-adjusts the penalty term as the sample size grows. Haughton (1988, 1989) and Nishii (1984) demonstrated that the BIC is a consistent estimator of the model as long as the true model is in the class of candidate models. This means that the probability of selecting the true model approaches 1 as the sample size (T) goes to $+\infty$. However, note that if the true model is not in the class of candidate models and the sample size (T) approaches ∞, the BIC has the tendency to underfit, i.e., it gives more weights to the penalty term than it is necessary, and selects a model with a few number of parameters.

6.6. Empirical Application

6.6.1. Raw Data

The raw data consist of three stocks traded on the New York Stock Exchange (NYSE): Citigroup (C), International Business Machines (IBM), and Ford Motor (F) and three stocks traded on the NASDAQ: Apple (AAPL), Dell (DELL), and Microsoft (MSFT). We define trade and volume durations for C, IBM and F and a price duration for AAPL, DELL and MSFT. Trade durations are defined as time elapsed between two consecutive trades. The price duration ($x_t(\iota_p)$) is the minimal time interval needed to observe a change in the mid-price[3] (p) not less than a threshold (ι_p) that is set to \$0.0156. Mathematically, we define the price duration as:

$$x_t = \inf\left\{x \in \mathbb{R}_+,\ \text{such that } |p_{T_t+x} - p_{T_t}| \geq \iota_p\right\}. \tag{6.26}$$

The volume duration[4] is the minimal time needed to trade a certain amount of shares at least equal to a threshold ι_v that we set to 25000.

$$x_t = \inf\left\{x \in \mathbb{R}_+,\ \text{such that } \sum_{i=t}^{t+x} V_i \geq \iota_v\right\}. \tag{6.27}$$

The volume durations characterize the liquidity of a stock on the market. Long volume durations imply that more time is needed to trade a given amount of shares. A stock is said to be liquid if it is characterized by the short volume durations with small changes in either bid, ask or mid-point price over these volume durations. The sampling period corresponds to July 2004 which has 21 trading days. The data were extracted from the Trade and Quote (TAQ) database[5] available at the NYSE.

[3] Mid-price is used to avoid biases caused by a bid-ask bounce (cf. Roll, 1984).
[4] Volume durations have been introduced by Gouriéroux et al. (1999).
[5] This database consists of two parts: The first reports all trades, while the second lists the best bid and the ask prices posted by market makers.

6.6.2. Seasonal Adjustment

We adjust the raw data as it has already been described above in sec. 5.5.1.1.

6.6.3. Comparison of the MSMD Models

First we estimate all the model parameters and the results are reported in Tables 6.4, 6.6, 6.8, 6.10, 6.12, 6.14. We observe a change in the estimates of the parameter, γ_1, in the different models with different data sets.

- For trade duration data: While the estimates of the transition probability parameter γ_1 in the standard MSMD model (M-Exp) are 0.335, 0.351 and 0.980 for C, IBM and F stocks, respectively, they quickly reach their upper bound value, 0.999, in the generalized MSMD models (M-Weibull, M-Gamma, M-GG and M-Burr). This might indicate a tendency of the generalized models to overestimate the parameter γ_1, and thus, distort the extent of the persistence in trade duration data.

- For price and volume duration data: Except for MSFT and F stocks, the parameters in the standard and generalized MSMD models are well estimated. For MSFT and F stocks the estimates of the rate, b, at which the transition probability increases are closer to 1. One can argue that this may be the reason why the estimates for the parameter, γ_1, for both stocks are so small.

To document the ability of the MSMD models to reproduce the long memory observed in the trade duration data we estimate the MSMD models using Ford trade durations and then, use the estimated parameters to generate data from each model specification. *Figs.* 6.2 depicts the autocorrelation functions of each model specification and that of the Ford trade durations.

The MSMD models can account well for a hyperbolically decaying autocorrelation function (ACF) observed by trade durations. In *Fig.* 6.2 we see that except for the ACF of the mixture of gamma all the other ACFs start at a relatively high first autocorrelation compared to that of the Ford data. The ACF of the mixture of generalized gamma is the closest to the ACF of the Ford data. This suggests that the mixture of generalized gamma model has the best fit as far as the ACF is concerned. This result will be confirmed by the Akaike information criterion (AIC) and Bayesian information criterion (BIC).

In our empirical study, we define the price duration as a time needed to observe a cumulative change in the mid-price not less than a threshold (ι_p) that is set to \$0.0156 and the volume duration as a minimal time needed to trade a given amount of shares at least equal to a threshold (ι_v) that we set to 25000. By decreasing the thresholds the persistence in both types of data becomes higher and one may need models such as MSMDs that can capture the high persistence and provide accurate forecasts.

From sec. 6.2 it clear that on the one hand the mixture of generalized gamma, gamma, Weibull, and exponential are nested and on the other hand the mixture of Burr, Weibull, and exponential are also. So, we can compare their fit using a simple likelihood ratio (LR) test. We test the hypothesis

$$H_0: \text{ the true model = the simple model}$$

against

$$H_a: \text{ the true model = the generalized model.}$$

By doing so, we obtain seven model comparisons. The null hypothesis that the true model is the simple model is strongly rejected at 5%, 2.5% and 1% confidence levels, cf. Table 6.16. This means that the simple model is always rejected in favor of the generalized model. We obtain similar results for different types of financial durations (trade, price and volume) and all stocks used in this study.

The results we obtain by considering the AIC and BIC selection criteria are in conformity with the LR test results. In other words, the AICs and BICs in Tables 6.5, 6.7, 6.9, 6.11, 6.13, 6.15 also speak in favor of the generalized models.

6.7. Conclusion

In this chapter we have introduced the alternative MSMD models, namely the mixture of gamma, Weibull, generalized gamma and Burr representations for financial durations and analyzed their performance to fit the data. The results from simulation studies and empirical application give evidence that with gamma, Weibull, generalized gamma and Burr distributions for the innovation financial duration data (trade, price and volume durations) can be fitted properly. Compared to the mixture of exponential distribution we just observed a slight superiority of the new models in terms of fitting the data. The MSMD models are more appropriate for modeling financial duration data than the ACD models that have ARMA structure, and thus, cannot capture the high persistence observed in the trade duration data. In sum, the generalized versions of the MSMD model are convenient for modeling financial duration data and offer researchers a new tool for empirical investigations. In the next chapter we extend the univariate MSMD model to a bivariate MSMD model that can permit a simultaneous modeling of the price and duration processes.

Table 6.1.: Simulated empirical moments of the Models

	$T = 15000$, $R = 100$						
	α	δ	κ	τ	*Overdispersion*	*Skewness*	*Kurtosis*
M-GG	0.600	0.700			2.212	6.496	83.395
	1.800	1.900			0.694	1.762	7.857
	1.800	0.700			1.330	3.826	29.130
	0.600	1.900			0.967	2.242	11.162
M-Gamma	0.500	1.000			1.683	4.309	35.486
	0.600	1.000			1.549	3.991	31.410
	1.500	1.000			0.999	2.895	17.930
	1.800	1.000			0.995	2.698	15.690
M-Weibull	1.000	0.650			1.881	5.883	76.442
	1.000	0.700			1.732	5.070	53.360
	1.000	1.200			1.001	2.804	16.644
	1.000	1.900			0.815	1.999	9.403
M-Exp	**1.000**	**1.000**			**1.252**	**3.336**	**22.973**
M-Burr			1.500	0.500	1.468	9.761	304.692
			2.500	0.500	0.858	2.944	22.287

Note: R denotes the number of replications and T the sample size. The values in the Table are average results based on Monte-Carlo simulations.

Table 6.2.: Raw Data

Raw data									
	Trade Durations			Price Durations			Volume Durations		
	C	IBM	F	APPL	DELL	MSFT	C	IBM	F
Number of obs.	94520	89162	48295	9877	6423	4861	6268	2862	3689
Overdispersion	1.165	1.146	1.238	1.648	1.479	1.192	0.747	0.655	0.871
Skewness	3.861	3.433	3.378	5.064	4.504	3.078	1.733	1.415	1.863
Kurtosis	32.726	22.214	23.937	53.443	36.256	17.477	8.703	5.977	9.298
Adjusted data									
Overdispersion	1.118	1.079	1.150	1.855	1.317	1.440	0.632	0.528	0.734
Skewness	3.621	3.095	2.880	10.946	3.140	17.083	1.363	0.986	1.505
Kurtosis	27.634	19.871	17.427	256.301	17.986	633.833	7.479	4.438	6.830
Q(10)	2224.9	1631.7	860.316	3478.9	417.280	920.401	1053.6	865.633	882.794
Q(100)	3708.1	4020.5	2034.5	4502.3	606.936	1037.5	1352.6	1113.2	1003.4

Table 6.3.: Monte Carlo MLE results

	Parameters	m_0	$\bar{\lambda}$	b	γ	m_0	$\bar{\lambda}$	b	γ
		T=5000				T=10000			
M-Exp	Bias	-0.011	-0.246	-0.206	-0.082	-0.013	-0.229	-0.155	-0.055
	SE	0.001	0.002	0.022	0.016	0.001	0.001	0.016	0.012
	RMSE	0.008	0.018	0.214	0.160	0.006	0.014	0.163	0.124

	Parameters	m_0	$\bar{\lambda}$	b	γ	δ	α	τ	κ
		T=5000							
M-GG	Bias	-0.009	-0.177	-0.039	-0.038	0.018	-0.009		
	SE	0.002	0.006	0.046	0.030	0.010	0.017		
	RMSE	0.021	0.055	0.459	0.300	0.101	0.165		
		T=10000							
M-GG	Bias	-0.014	-0.161	-0.066	-0.038	0.006	0.004		
	SE	0.002	0.004	0.037	0.026	0.008	0.013		
	RMSE	0.018	0.044	0.368	0.261	0.080	0.132		
		T=5000							
M-Weibull	Bias	-0.010	-0.241	-0.224	-0.089	0.001			
	SE	0.002	0.004	0.042	0.027	0.002			
	RMSE	0.018	0.038	0.420	0.273	0.017			
		T=10000							
M-Weibull	Bias	-0.014	-0.229	-0.191	-0.073	-0.002			
	SE	0.001	0.003	0.026	0.020	0.001			
	RMSE	0.012	0.025	0.261	0.203	0.010			
		T=5000							
M-gamma	Bias	-0.012	-0.188	-0.073	-0.044		-0.005		
	SE	0.0011	0.002	0.022	0.018		0.003		
	RMSE	0.009	0.021	0.217	0.179		0.030		
		T=10000							
M-gamma	Bias	-0.015	-0.167	-0.066	-0.022		-0.006		
	SE	0.001	0.002	0.016	0.012		0.002		
	RMSE	0.005	0.015	0.155	0.126		0.022		
		T=5000							
M-Burr	Bias	-0.011	-0.241	-0.238	-0.088			-0.007	-0.001
	SE	0.002	0.005	0.051	0.033			0.004	0.004
	RMSE	0.023	0.053	0.511	0.328			0.043	0.037
		T=10000							
M-Burr	Bias	-0.013	-0.229	-0.138	-0.044			-0.006	-0.003
	SE	0.001	0.004	0.034	0.026			0.003	0.003
	RMSE	0.015	0.037	0.342	0.262			0.031	0.025

The parameters are: $m_0 = 1.2$, $\bar{\lambda} = 1$, $b = 2$, $\gamma = 0.5$, $\delta = 0.8$, $\alpha = 1.2$, $\tau = 0.5$, $\kappa = 1.2$ and $k = 7$. The table shows average results over 100 replications.

Table 6.4.: Estimates of mixture of exponential, gamma,and Weibull models for trade durations

	C					IBM					F				
	m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1	
M-Exp	1.114	1.142	1.784	0.335		1.104	1.173	2.318	0.351		1.265	0.609	30.148	0.980	
	(0.002)	**(0.009)**	**(0.070)**	**(0.039)**		**(0.002)**	**(0.016)**	**(0.116)**	**(0.044)**		**(0.004)**	**(0.005)**	**(3.339)**	**(0.014)**	
	m_0	$\bar{\lambda}$	b	γ_1	δ	m_0	$\bar{\lambda}$	b	γ_1	δ	m_0	$\bar{\lambda}$	b	γ_1	δ
M-Gamma	1.275	1.808	1.548	0.999	4.358	1.296	1.995	1.448	0.999	6.382	0.717	1.901	1.606	0.999	2.009
	(0.001)	**(0.010)**	**(0.009)**	**(0.000)**	**(0.0715)**	**(0.001)**	**(0.010)**	**(0.007)**	**(0.000)**	**(0.137)**	**(0.003)**	**(0.025)**	**(0.021)**	**(0.000)**	**(0.041)**
	m_0	$\bar{\lambda}$	b	γ_1	α	m_0	$\bar{\lambda}$	b	γ_1	α	m_0	$\bar{\lambda}$	b	γ_1	α
M-Weibull	1.294	1.921	1.479	0.999	2.979	0.600	0.451	7.598	0.999	1.941	0.679	2.304	1.509	0.999	1.808
	(0.001)	**(0.007)**	**(0.007)**	**(0.000)**	**(0.021)**	**(0.002)**	**(0.005)**	**(0.134)**	**(0.000)**	**(0.015)**	**(0.002)**	**(0.028)**	**(0.015)**	**(0.000)**	**(0.027)**

Note we use $k = 7$ for the estimation and the numbers in bold in parentheses are standard errors of the estimations.

Table 6.5.: Likelihood function, AIC and BIC

	$M-Exp$			$M-gamma$			$M-Weibull$		
	C	IBM	F	C	IBM	F	C	IBM	F
Lik	-89486.279	-83717.941	-44107.365	-79073.427	-74132.762	-42555.178	-79525.626	-77588.886	-42598.795
AIC	178980.558	167443.882	88222.731	158154.854	148273.523	85118.357	159059.252	155185.771	85205.589
BIC	179018.385	167481.475	88257.871	158192.680	148311.116	85153.497	159097.078	155223.364	85240.730

Table 6.6.: Estimates of mixture of gen. gamma and Burr models for trade durations

	C						IBM						F					
	m_0	$\bar{\lambda}$	b	γ_1	δ	α	m_0	$\bar{\lambda}$	b	γ_1	δ	α	m_0	$\bar{\lambda}$	b	γ_1	δ	α
M-GG	0.706	1.960	2.462	0.999	1.046	6.432	0.706	2.022	2.439	0.999	0.862	9.574	0.823	1.326	2.448	0.999	0.329	9.688
	(0.001)	**(0.010)**	**(0.007)**	**(0.000)**	**(0.024)**	**(0.247)**	**(0.001)**	**(0.013)**	**(0.007)**	**(0.000)**	**(0.022)**	**(0.368)**	**(0.004)**	**(0.017)**	**(0.059)**	**(0.000)**	**(0.005)**	**(0.292)**
	m_0	$\bar{\lambda}$	b	γ_1	τ	κ	m_0	$\bar{\lambda}$	b	γ_1	τ	κ	m_0	$\bar{\lambda}$	b	γ_1	τ	κ
M-Burr	0.762	1.359	4.598	0.999	2.215	3.526	0.761	1.509	2.668	0.999	1.607	3.534	1.317	2.042	2.366	0.999	0.843	2.436
	(0.007)	**(0.054)**	**(0.327)**	**(0.000)**	**(0.054)**	**(0.065)**	**(0.003)**	**(0.020)**	**(0.018)**	**(0.000)**	**(0.039)**	**(0.058)**	**(0.002)**	**(0.031)**	**(0.017)**	**(0.000)**	**(0.051)**	**(0.048)**

Note we use $k = 7$ for the estimation and the numbers in bold in parentheses are standard errors of the estimations.

Table 6.7.: Likelihood function, AIC and BIC

	$M-GG$			$M-Burr$		
	C	IBM	F	C	IBM	F
Lik	-78721.981	-74028.438	-42209.827	-78959.218	-73999.544	-42509.754
AIC	157453.963	148066.877	84429.654	157928.436	148009.089	85029.507
BIC	157501.246	148113.868	84473.579	157975.719	148056.080	85073.433

Table 6.8.: Estimates of mixture of exponential, gamma and Weibull models for price durations

	AAPL					DELL					MSFT				
	m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1	
M-Exp	1.362	2.669	3.198	0.892		1.249	1.852	1.597	0.558		1.188	1.413	1.009	0.081	
	(0.012)	**(0.333)**	**(0.376)**	**(0.054)**		**(0.007)**	**(0.078)**	**(0.128)**	**(0.106)**		**(0.009)**	**(0.058)**	**(0.109)**	**(0.027)**	
	m_0	$\bar{\lambda}$	b	γ_1	α	m_0	$\bar{\lambda}$	b	γ_1	α	m_0	$\bar{\lambda}$	b	γ_1	α
M-Gamma	1.337	2.738	1.752	0.636	1.113	0.675	2.576	1.676	0.946	1.479	0.725	1.963	2.100	0.995	1.800
	(0.011)	**(0.199)**	**(0.128)**	**(0.108)**	**(0.042)**	**(0.010)**	**(0.137)**	**(0.075)**	**(0.028)**	**(0.071)**	**(0.011)**	**(0.107)**	**(0.206)**	**(0.012)**	**(0.110)**
	m_0	$\bar{\lambda}$	b	γ_1	δ	m_0	$\bar{\lambda}$	b	γ_1	δ	m_0	$\bar{\lambda}$	b	γ_1	δ
M-Weibull	1.338	2.428	1.748	0.752	1.082	1.301	2.321	1.680	0.858	1.158	0.727	1.819	1.859	0.965	1.433
	(0.011)	**(0.144)**	**(0.099)**	**(0.095)**	**(0.032)**	**(0.012)**	**(0.146)**	**(0.118)**	**(0.090)**	**(0.037)**	**(0.009)**	**(0.077)**	**(0.104)**	**(0.026)**	**(0.042)**

Note we use $k = 7$ for the estimation and the numbers in bold in parentheses are standard errors of the estimations.

Table 6.9.: Likelihood function, AIC and BIC

	$M - Exp$			$M - gamma$			$M - Weibull$		
	AAPL	DELL	MSFT	AAPL	DELL	MSFT	AAPL	DELL	MSFT
Lik	-6936.263	-4976.821	-4071.947	-6898.511	-4932.118	-3931.903	-6928.469	-4944.869	-3956.777
AIC	13880.526	9961.642	8151.894	13807.021	9874.236	7873.807	13866.938	9899.738	7923.553
BIC	13909.318	9988.713	8177.850	13843.011	9908.074	7906.252	13902.928	9933.576	7955.998

Table 6.10.: Estimates of mixture of gen. gamma and Burr models for price durations

	AAPL						DELL						MSFT					
	m_0	$\bar{\lambda}$	b	γ_1	δ	α	m_0	$\bar{\lambda}$	b	γ_1	δ	α	m_0	$\bar{\lambda}$	b	γ_1	δ	α
M-GG	1.310	1.995	1.832	0.458	0.741	1.528	1.279	2.110	1.985	0.925	0.553	3.273	0.769	1.643	1.603	0.788	0.773	2.346
	(0.013)	**(0.127)**	**(0.200)**	**(0.119)**	**(0.023)**	**(0.060)**	**(0.013)**	**(0.141)**	**(0.144)**	**(0.068)**	**(0.045)**	**(0.421)**	**(0.010)**	**(0.066)**	**(0.093)**	**(0.087)**	**(0.041)**	**(0.189)**
	m_0	$\bar{\lambda}$	b	γ_1	τ	κ	m_0	$\bar{\lambda}$	b	γ_1	τ	κ	m_0	$\bar{\lambda}$	b	γ_1	τ	κ
M-Burr	0.683	2.030	1.815	0.566	0.410	1.237	1.300	2.198	1.893	0.989	0.169	1.278	1.272	1.967	2.054	0.998	0.299	1.678
	(0.022)	**(0.195)**	**(0.245)**	**(0.230)**	**(0.044)**	**(0.049)**	**(0.010)**	**(0.104)**	**(0.060)**	**(0.005)**	**(0.061)**	**(0.037)**	**(0.010)**	**(0.091)**	**(0.078)**	**(0.001)**	**(0.063)**	**(0.062)**

Note we use $k = 7$ for the estimation and the numbers in bold in parentheses are standard errors of the estimations.

Table 6.11.: Likelihood function, AIC and BIC

	$M - GG$			$M - Burr$		
	AAPL	DELL	MSFT	AAPL	DELL	MSFT
Lik	-6812.893	-4909.979	-3922.296	-6864.715	-4931.991	-3933.951
AIC	13637.786	9831.958	7856.593	13741.430	9875.982	7879.902
BIC	13680.974	9872.564	7895.527	13784.618	9916.588	7918.836

Table 6.12.: Estimates of mixture of exponential, gamma and Weibull models for volume durations

	C					IBM					F				
	m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1		m_0	$\bar{\lambda}$	b	γ_1	
M-Exp	1.046	1.168	1.168	0.055		0.995	1.196	1.380	0.205		1.086	1.265	1.045	0.062	
	(0.009)	**(0.021)**	**(0.225)**	**(0.033)**		**(0.015)**	**(0.022)**	**(2.317)**	**(0.450)**		**(0.010)**	**(0.036)**	**(0.185)**	**(0.038)**	
	m_0	$\bar{\lambda}$	b	γ_1	α	m_0	$\bar{\lambda}$	b	γ_1	α	m_0	$\bar{\lambda}$	b	γ_1	α
M-Gamma	1.109	1.299	1.006	0.061	2.714	1.129	1.346	1.159	0.192	4.966	1.139	1.411	1.228	0.144	2.032
	(0.006)	**(0.030)**	**(0.098)**	**(0.019)**	**(0.051)**	**(0.006)**	**(0.036)**	**(0.104)**	**(0.048)**	**(0.164)**	**(0.008)**	**(0.048)**	**(0.163)**	**(0.057)**	**(0.049)**
	m_0	$\bar{\lambda}$	b	γ_1	δ	m_0	$\bar{\lambda}$	b	γ_1	δ	m_0	$\bar{\lambda}$	b	γ_1	δ
M-Weibull	1.111	1.299	1.108	0.098	1.907	1.132	1.365	1.283	0.331	2.790	1.150	1.471	1.142	0.127	1.643
	(0.005)	**(0.027)**	**(0.129)**	**(0.036)**	**(0.023)**	**(0.005)**	**(0.034)**	**(0.145)**	**(0.106)**	**(0.072)**	**(0.008)**	**(0.052)**	**(0.158)**	**(0.052)**	**(0.026)**

Note we use $k = 7$ for the estimation and the numbers in bold in parentheses are standard errors of the estimations.

Table 6.13.: Likelihood function, AIC and BIC

	$M - Exp$			$M - gamma$			$M - Weibull$		
	C	IBM	F	C	IBM	F	C	IBM	F
Lik	-5235.830	-2365.697	-2927.388	-4023.745	-1322.551	-2547.913	-3904.470	-1247.685	-2465.755
AIC	10479.659	4739.393	5862.775	8057.490	2655.101	5105.827	7818.941	2505.370	4941.509
BIC	10506.632	4763.230	5887.628	8091.206	2684.898	5136.892	7852.657	2535.166	4972.375

Table 6.14.: Estimates of mixture of gen. gamma and Burr models for volume durations

	C						IBM						F					
	m_0	$\bar{\lambda}$	b	γ_1	δ	α	m_0	$\bar{\lambda}$	b	γ_1	δ	α	m_0	$\bar{\lambda}$	b	γ_1	δ	α
M-GG	0.880	1.314	1.038	0.085	2.487	0.678	0.858	1.382	1.001	0.200	4.604	0.524	1.146	1.428	1.109	0.144	2.548	0.532
	(0.005)	**(0.029)**	**(0.144)**	**(0.036)**	**(0.139)**	**(0.052)**	**(0.006)**	**(0.033)**	**(0.311)**	**(0.169)**	**(0.845)**	**(0.111)**	**(0.007)**	**(0.042)**	**(0.171)**	**(0.069)**	**(0.167)**	**(0.044)**
	m_0	$\bar{\lambda}$	b	γ_1	τ	κ	m_0	$\bar{\lambda}$	b	γ_1	τ	κ	m_0	$\bar{\lambda}$	b	γ_1	τ	κ
M-Burr	0.880	1.364	1.409	0.175	0.269	2.418	1.187	0.971	1.831	0.169	0.308	3.103	0.847	1.650	2.041	0.744	0.237	1.674
	(0.012)	**(0.103)**	**(0.503)**	**(0.099)**	**(0.038)**	**(0.056)**	**(0.013)**	**(0.170)**	**(0.461)**	**(0.058)**	**(0.068)**	**(0.104)**	**(0.018)**	**(0.307)**	**(0.323)**	**(0.117)**	**(0.070)**	**(0.049)**

Note we use $k = 7$ for the estimation and the numbers in bold in parentheses are standard errors of the estimations.

Table 6.15.: Likelihood function, AIC and BIC

	$M-GG$			$M-Burr$		
	C	IBM	F	C	IBM	F
Lik	-3890.567	-1238.753	-2431.064	-4016.101	-1294.226	-2443.051
AIC	7793.133	2489.507	4874.129	8044.203	2600.452	4898.100
BIC	7833.592	2525.262	4911.407	8084.662	2636.208	4935.400

Table 6.16.: Results of the likelihood ratio tests

		H_1			
		M-Gamma	M-Weibull	M-GG	M-Burr
	M-exp	0.000	0.000	0.000	0.000
H_0	M-Gamma	0.000	0.000	0.000	0.000
	M-Weibull	0.000	0.000	0.000	0.000

Note: The null hypothesis that the true model is the simple model is rejected at any standard confidence levels for all stocks.

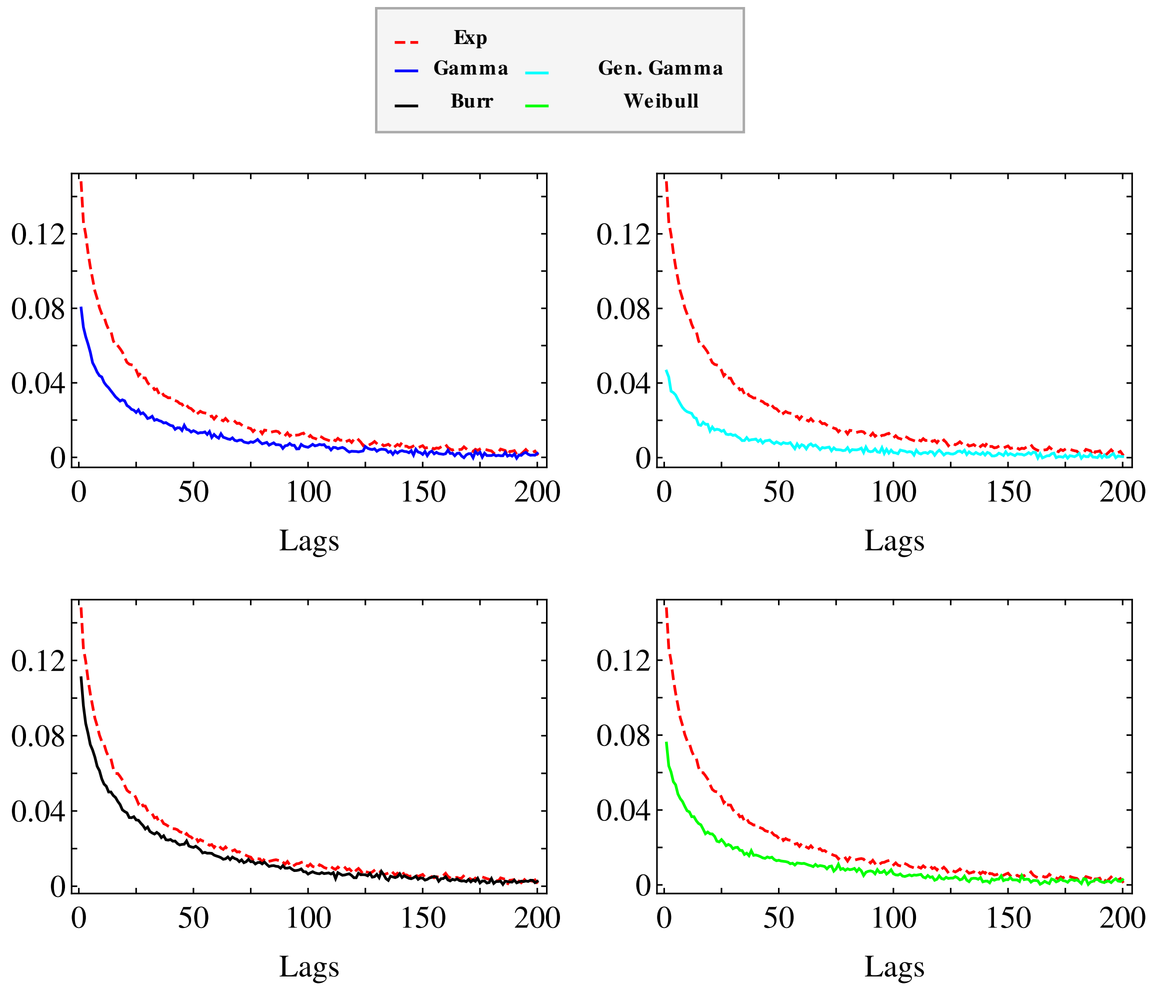

Figure 6.1.: Autocorrelation functions of simulated data sets corresponding to the mixture of exponential, gamma, Weibull, Burr, and generalized gamma specifications with intensity components k that we set to 7. The parameters used for the simulation are: $m_0 = 1.2$, $b = 2$, $\gamma_1 = 0.5$, $\bar{\lambda} = 1$

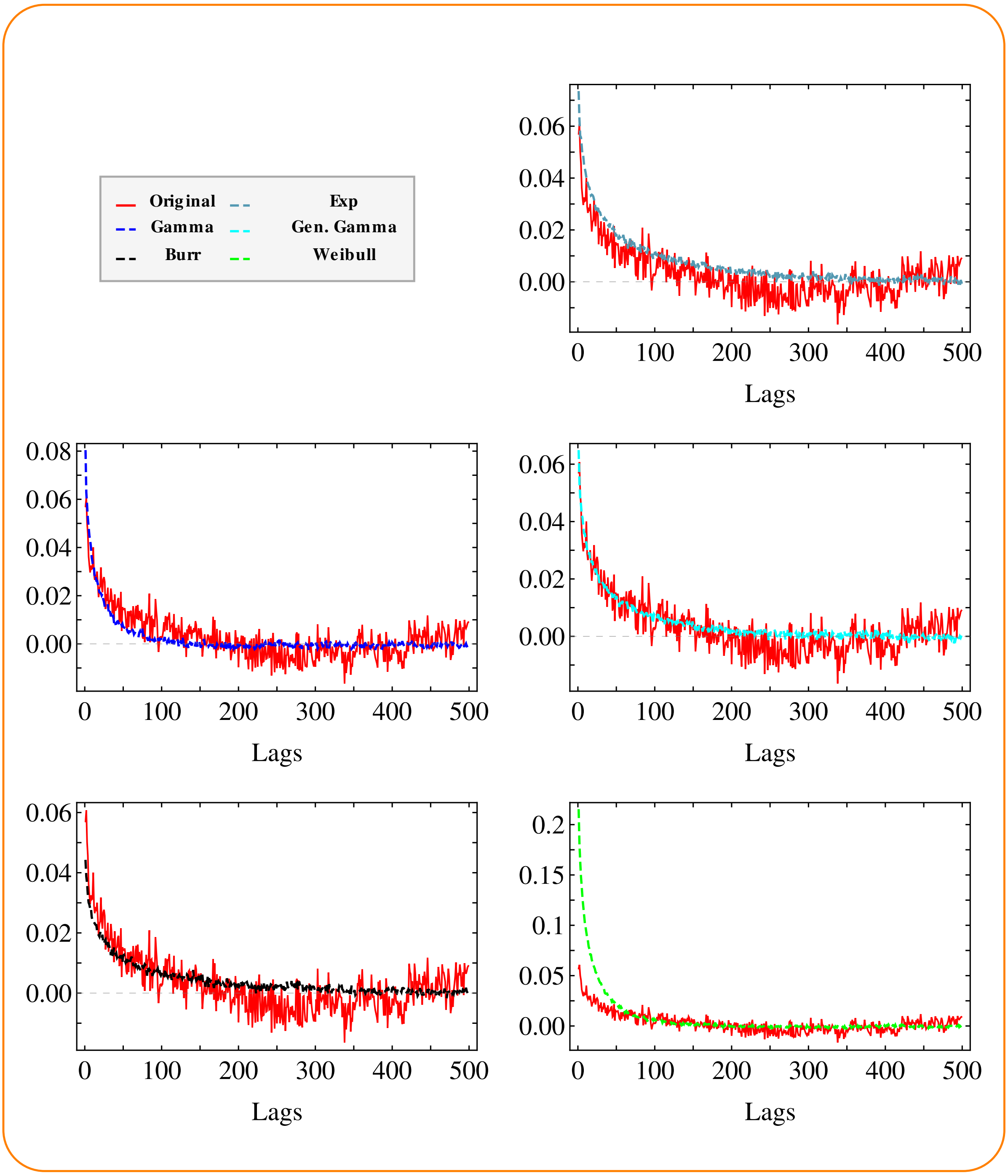

Figure 6.2.: Autocorrelation functions of Ford trade durations and simulated data sets corresponding to the mixture of exponential, gamma, Weibull, Burr, and generalized gamma specifications with intensity components k that we set to 7. The parameters used for the simulation are set equal to their estimated value for the Ford trade duration data.

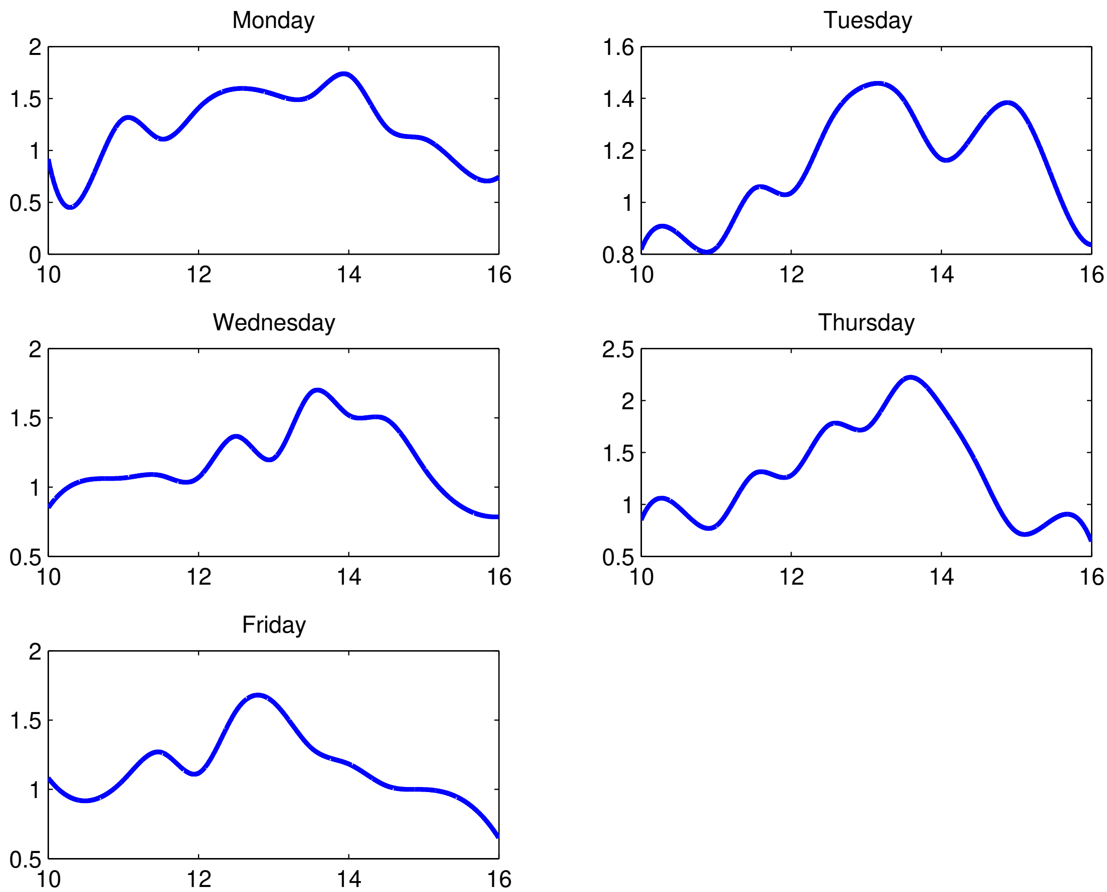

Figure 6.3.: Time-of-the-day function for Ford trade durations

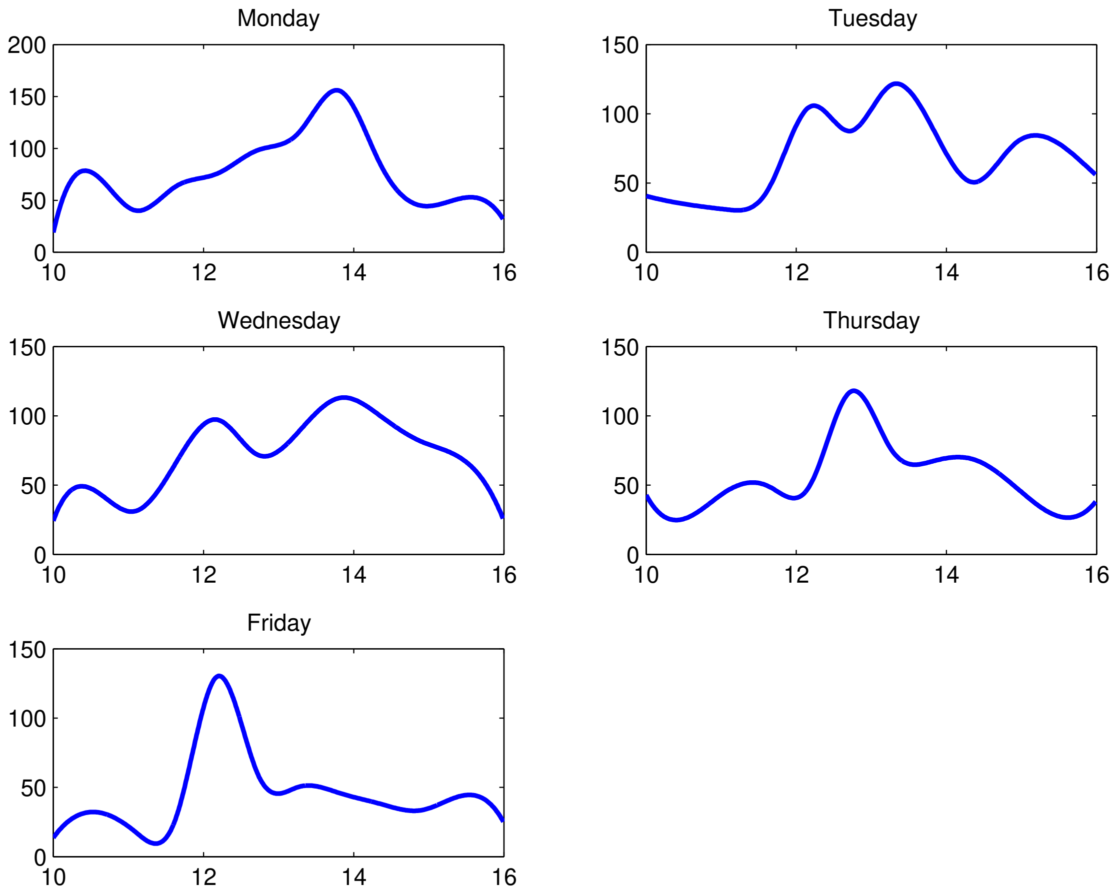

Figure 6.4.: Time-of-the-day function for AAPL price durations

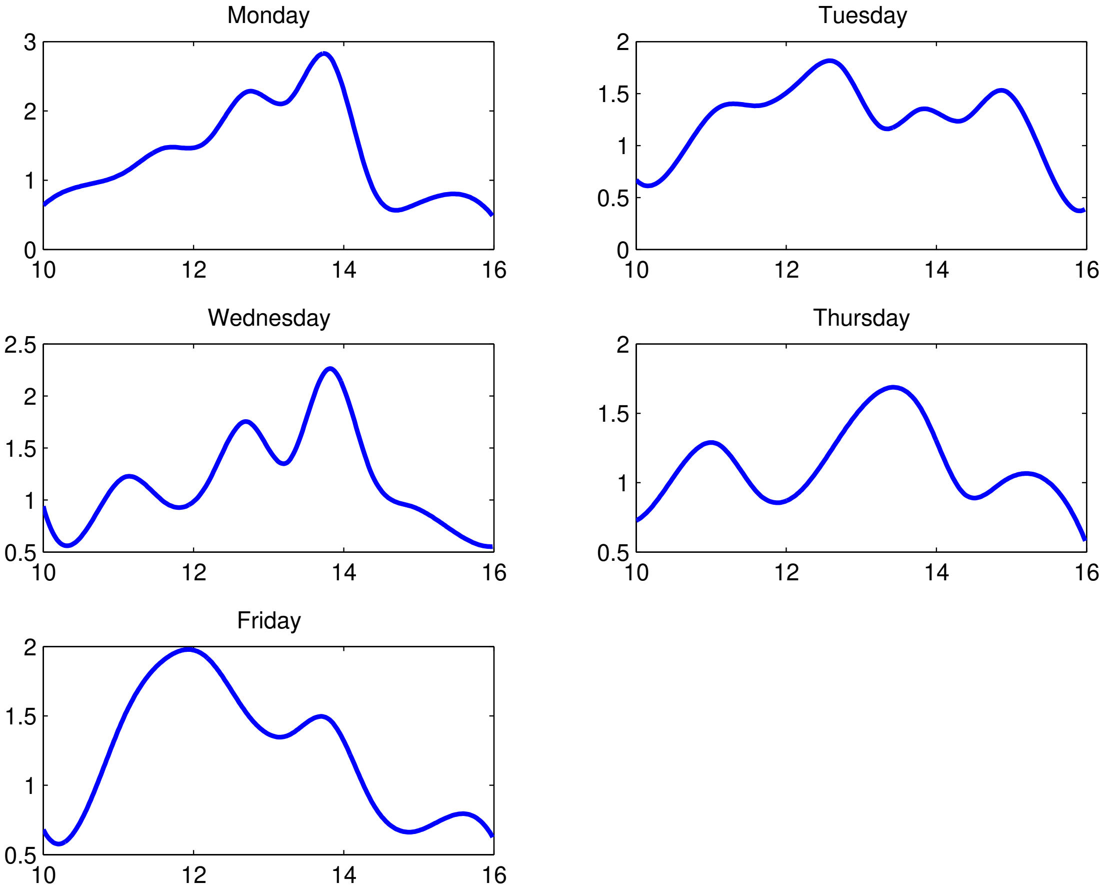

Figure 6.5.: Time-of-the-day function for Citigroup volume durations

7. A Bivariate Markov-Switching Multifractal Duration Model

7.1. Introduction

It is well-documented that financial durations have information content, and therefore, are crucial for the price adjustment process. Since the seminal paper of Engle and Russell (1998), various extensions of the standard ACD model have been proposed in the literature. One issue related to all these models is that they only model the time elapsed between market events, and do not take into account the information given by the price process. This information may be of a great importance for a better understanding of trading activities in the market and may help to answer important empirical questions. It seems worthwhile to extend the univariate models in a way that allows for a joint modeling of the price and the duration processes. Such models can permit researchers, for instance, to analyze the impact of the market activity, measured by durations or the average volume per transaction on the magnitude of the bid-ask spread or trade-to-trade return volatilities. They can also enable the scrutiny of interdependencies between durations, prices and trading volumes, or the co-movement between bid-ask spread volatilities. These research questions receive widespread attention in the microstructure literature. With the availability of high frequency data it seems appropriate to verify whether these theories can be confirmed empirically. In line with this, Russell and Engle (2005) introduced the autoregressive conditional multinomial (ACM) model that permits a simultaneous modeling of the trade duration process and the discrete price changes process. The ACM is a combination of an ACD model with a dynamic multinomial model. Engle (2000) combined an ACD model with a GARCH model to obtain an ACD-GARCH model which allows for a joint modeling of the timing between trades and the volatility related to the price process. Trade durations are modeled by means of the ACD model while volatilities of trade-to-trade returns are captured by the GARCH model conditional on the concurrent trade duration. This model has been extended by Grammig and Wellner (2002) for allowing for a reciprocal relationship between the trade duration process and the volatility process. Hasbrouck (1991) and later Dufour and Engle (2000b) used a vector autoregressive (VAR) system to analyze interdependencies between microstructure economic variables, namely trade durations, prices and volumes. This modeling approach has been extended by Manganelli (2005).

Recent empirical investigations of high-frequency data reveal that many financial quantities exhibit long memory properties and multifractality. For example, Qiu et al. (2012) report strong multifractality in spread returns. Hautsch (2012) accounts high persistency in the bid-ask spreads. Long memory in trade durations and in trading volume has been documented by Jasiak (1998) and Lobato and Velasco (2000), respectively. Chen et al. (2013) report self-similarity properties in inter-trade duration data. It seems essential that we need new models that can reproduce the above-mentioned features as well as permit a simultaneous modeling of the duration and price processes.

In this chapter we introduce a bivariate Markov switching multifractal duration (MSMD) model that is built on the bivariate Markov switching multifractal (MSM) process developed by Calvet et al. (2006). Inspired by the univariate Markov switching multifractal process (cf. Calvet and Fisher, 2001a, 2004a) Calvet et al. (2006) proposed a bivariate MSM process which has been applied for analyzing risk transmission, co-movement of volatilities and volatility spillovers in financial markets (cf. Idier, 2011). The bivariate MSM process also found application in the calculation of value-at-risk (VaR) forecasts for portfolios (cf. Calvet et al., 2006; Liu, 2008). Recently, Liu and Lux (2014) refined the bivariate MSM model by allowing correlations between volatility components to be non-homogeneous with two different parameters that control volatility correlations at high and low frequencies. Our bivariate MSMD model can be used to analyze irregularly spaced as well as regularly spaced data. So far, extant Markov switching multifractal duration (MSMD) models are univariate ones. They have independently been proposed by Chen et al. (2013) and Baruník et al. (2012) and seem to be more appropriate for the analysis of high-frequency financial duration data. However, as we can see above multivariate settings are preferable in empirical research because they allow to answer many important questions. Our motivation is to provide researchers and practitioners a tool that can be used for the joint modeling of the price and duration process as well as for the study of the co-movement in microstructure or trade-related variables at high-frequency level.

The rest of the chapter is organized as follows. Section 7.2 introduces the bivariate Markov switching multifractal duration model. Its statistical properties are presented in Section 7.3. An empirical application is illustrated in Section 7.5 and Section 7.6 concludes.

7.2. A Bivariate MSMD Model

The bivariate MSMD model is defined as:

$$z_t = g[\lambda(M_t)] * \xi_t, \tag{7.1}$$

where $z_t = (z_{1,t}, z_{2,t})'$ is a (2×1) vector of economic variable series, the vector $\xi_t = (\xi_{1,t}, \xi_{2,t})'$ is assumed to follow a bivariate Lognormal distribution such that $\ln \xi_t$ is a bivariate normally distributed random variable with mean vector μ and variance-covariance Σ. $*$ denotes element by element multiplication. The mean intensity function $\lambda(M_t)$ is the vector of the products of multifractal intensity components, i.e. $\lambda(M_t) = [\lambda(M_{1,t}), \lambda(M_{2,t})]'$, where each $\lambda(M_{q,t})$ is defined as the product of intensity components for series q:

$$\lambda(M_{q,t}) = \bar{\lambda}_q \prod_{i=1}^{k} M^i_{q,t}. \tag{7.2}$$

The constant scale parameter $\bar{\lambda}_q$ represents the unconditional mean intensity and $M^i_{q,t}$ is the intensity component at frequency i of series q.

- Two choices of $g(\cdot)$ can be proposed, in particular:
- $g(y) = y$. This functional form is straightforward and has been employed by Baruník et al. (2012) to analyze financial price durations. In this case, we have: $g[\lambda(M_t)] = \lambda(M_t)$.
- $g(y) = 1/y$. This second one has been proposed by Chen et al. (2013) in order to provide a convenient modeling for financial intertrade durations. In this case, we have: $g[\lambda(M_t)] = [\lambda(M_t)]^{-1}$.

The period t intensity state is characterized by a $2 \times k$ matrix $M_t = \left(M^1_t; M^2_t; \ldots; M^k_t\right)$ and the vector of the components at the i^{th} frequency is $M^i_t = (M^i_{1,t}\ M^i_{2,t})$. The intensity vectors M^i_t are persistent, non-negative and satisfy $E[M^i_t] = \mathbf{1}$, where $\mathbf{1} = (1, 1)'$. Economic intuition behind the choice of the dynamics for each vector M^i_t is that intensity arrivals are correlated but not necessarily simultaneous across markets. For this reason Calvet et al. (2006) allow arrivals across series to be characterized by a correlation coefficient ϱ. By considering two random variables $I^i_{1,t}$ and $I^i_{2,t}$ which are equal to 1 if each series $q \in \{1, 2\}$ is hit by an information arrival with probability γ_i, and equal to zero otherwise, Calvet et al. (2006) specified the arrival vector to be *i.i.d.* and its unconditional distribution has to satisfy three conditions. First, the arrival vector is symmetrically distributed: $(I^i_{1,t}, I^i_{2,t}) \overset{d}{=} (I^i_{2,t}, I^i_{1,t})$. Second, the switching probability of a series is equal to an exogenous constant: $Pr(I^i_{2,t} = 1) = \gamma_i$. Third, there exists $\varrho \in [0, 1]$ such that

$$Pr(I^i_{1,t} = 1 | I^i_{2,t} = 1) = (1 - \varrho)\gamma_i + \varrho. \tag{7.3}$$

$\varrho = 0$ signifies that new arrivals are independent and $\varrho = 1$ signifies that they are simultaneous. The above-mentioned three conditions define a unique distribution of $(I^i_{1,t}, I^i_{2,t})$ whose switching probabilities are defined as:

$$\gamma_i = 1 - (1 - \gamma_1)^{b^{i-1}}, \tag{7.4}$$

with parameters $\gamma_1 \in [0, 1]$ and $b \in (1, \infty)$.

The choice of the bivariate Lognormal distribution is motivated by the previous research. In fact, Allen et al. (2008) and later Allen et al. (2009) demonstrated the ability of the univariate Lognormal distribution to fit financial duration data. Furthermore, Lognormal distribution also found application in modeling of the price volatility (cf. Stein and Stein, 1991) and realized volatility (cf. Andersen et al., 2003). The bivariate Lognormal distribution has a closed form conditional density function that allows us to easily apply the maximum likelihood estimation approach. All these reasons determine us to use the bivariate Lognormal distribution for the innovations in the bivariate MSMD model. Note that any other bivariate distribution with positive support can also be used.

For the empirical application we borrow the simple specification in Calvet et al. (2006) which consists in assuming that each M_t^i is drawn from a bivariate binomial distribution $M = (M_1, M_2)'$, with M_1 taking values $m_1 \in (0, 2)$ and $2 - m_1$, and M_2 taking values $m_2 \in (0, 2)$ and $2 - m_2$. In their seminal paper Calvet et al. (2006) allowed for variation of the correlation (ρ_m) between components M_1 and M_2 and reported that the hypothesis of a perfect positive correlation, i.e. $\rho_m = 1$, is never rejected. We also follow Calvet et al. (2006) and set ρ_m to one in our empirical study.

7.3. Statistical Properties of the Model

The mean vector μ and variance-covariance matrix Σ of $\ln \xi_t$ are given by

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}. \tag{7.5}$$

Thus, the mean and variance-covariance of the ξ are given by

$$\mathbb{E}[\xi_t]_i = \exp\left[\mu_i + \frac{1}{2}\Sigma_{ii}\right], \tag{7.6}$$

$$\mathbb{V}ar[\xi_t]_{ij} = \exp\left[\mu_i + \mu_j + \frac{1}{2}\left(\Sigma_{ii} + \Sigma_{jj}\right)\right]\left[\exp\left(\Sigma_{ij}\right) - 1\right] = (d_{ij}). \tag{7.7}$$

Note that by assuming $\mu_i = -\frac{1}{2}\Sigma_{ii}$ we obtain:

$$\mathbb{E}[\xi_t]_i = 1 \quad \text{and} \quad \mathbb{V}ar[\xi_t]_{ij} = \left[\exp\left(\Sigma_{ij}\right) - 1\right] = (d_{ij}). \tag{7.8}$$

This restriction helps to reduce the number of parameters to be estimated in the model.

Let ρ denote the correlation coefficient between $\ln \xi_{1,t}$ and $\ln \xi_{2,t}$, then the corresponding correlation coefficient ς between $\xi_{1,t}$ and $\xi_{2,t}$ is given by

$$\varsigma = \frac{\exp(\rho\sqrt{\sigma_{11}\sigma_{22}}) - 1}{\sqrt{[\exp(\sigma_{11}) - 1][\exp(\sigma_{22}) - 1]}} = h(\rho), \tag{7.9}$$

where $\varsigma \in (-1, 1)$, $h(\rho) = 0$ if $\rho = 0$, $|\varsigma| < \rho$, and $h(\rho) \neq -h(-\rho)$, cf. Mostafa and Mahmoud (1964) for more detail.

The conditional covariance quantifies the co-movement and is given by

$$\mathbb{C}ov(z_{1,t+l}, z_{2,t+l}) = \varsigma\sqrt{d_{11}d_{22}}\,\mathbb{E}\left(g[\lambda(M_{1,t+l})]g[\lambda(M_{2,t+l})]\right) - C, \tag{7.10}$$

and the conditional correlation becomes

$$\mathbb{C}orr(z_{1,t+l}, z_{2,t+l}) = \frac{\varsigma\sqrt{d_{11}d_{22}}\mathbb{E}\left(g[\lambda(M_{1,t+l})]g[\lambda(M_{2,t+l})]\right) - C}{h\left(g[\lambda(M_{1,t+l})], g[\lambda(M_{2,t+l})]\right)} \tag{7.11}$$

where

$$h(\cdot) = \left[\left(\mathbb{E}[g(\lambda(M_{1,t}))]\exp(\sigma_{11}) - a_1^2\right)\left(\mathbb{E}[(\lambda(M_{2,t}))]\exp(\sigma_{22}) - a_2^2\right)\right]^{1/2}, \tag{7.12}$$

and $C = a_1 a_2$ with $a_1 = \mathbb{E}[g(\lambda(M_{1,t}))]$ and $a_2 = \mathbb{E}[g(\lambda(M_{2,t}))]$.

7.4. Estimation Approach

Our bivariate MSMD model requires ten parameters $\left(m_0^1, m_0^2, \bar{\lambda}_1, \bar{\lambda}_2, \sigma_1, \sigma_2, b, \gamma_1, \rho, \varrho\right)$ where m_0^1 and m_0^2 determine the bivariate Binomial distribution of intensity components, γ_1 their transition probabilities, $\bar{\lambda}_1$ and $\bar{\lambda}_2$ are the unconditional mean intensities, σ_1 and σ_2 represent the standard deviations of the Normal innovations, ρ is the correlation between Normal innovations and ϱ the correlation of arrivals across series. To estimate these parameters we use the two-step estimation procedures proposed by Calvet et al. (2006). The main asset of this two-step estimation approach is that it allows for performing the estimation of the bivariate model in two straightforward steps. In the following we briefly describe the two steps:

1. The first step consists in optimizing the sum of the two univariate log-likelihoods

$$L\left(z_{1t}, m_0^1, \bar{\lambda}_1, \sigma_1, b, \gamma_1\right) + L\left(z_{2t}, m_0^2, \bar{\lambda}_2, \sigma_2, b, \gamma_1\right) \tag{7.13}$$

   where L is the log-likelihood of the univariate MSMD. This first step provides the estimates of the parameter vector $\Phi = \left(m_0^1, m_0^2, \bar{\lambda}_1, \bar{\lambda}_2, \sigma_1, \sigma_2, b, \gamma_1\right)$ that are consistent as long as the gradient of the sum of the both univariate log-likelihoods with respect to the true parameters are zero.

2. In the second step we obtain the remaining parameters (ρ, ϱ) via the simulated likelihood method. we use the particle filter described in sec. 2.5.2 to optimize the likelihood function of the bivariate MSMD process. The bivariate pdf in the MSMD model is given by

$$f\left(z_{1t}, z_{2t}|\Im_{t-1}\right) = \sum_{i=1}^{N} f(z_{1t}, z_{2t}|M_{1t} = m_1^i, M_{2t} = m_2^i)Pr(M_{1t} = m_1^i, M_{2t} = m_2^i|\Im_{t-1}). \tag{7.14}$$

   where $f\left(z_{1t}, z_{2t}|M_{1t} = m_1^i, M_{2t} = m_2^i\right)$ is

$$f\left(z_{1t}, z_{2t}|M_{1t} = m_1^i, M_{2t} = m_2^i\right) = \frac{1}{2\pi z_{1t}z_{2t}\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}Q\right], \tag{7.15}$$

with

$$Q = \left[\frac{\ln\left(z_{1t}\lambda(m_1^i)\right) - \mu_1}{\sigma_1}\right]^2 - 2\rho\left[\frac{\ln\left(z_{1t}\lambda(m_1^i)\right) - \mu_1}{\sigma_1}\right]\left[\frac{\ln\left(z_{2t}\lambda(m_2^i)\right) - \mu_2}{\sigma_2}\right] + \left[\frac{\ln\left(z_{2t}\lambda(m_2^i)\right) - \mu_2}{\sigma_2}\right]^2.$$

With simulated draws $\hat{M}_t^i$ from $M_t|\Im_{t-1}$ that we obtain via the particle filter, the Monte Carlo estimate of the conditional density is thus,

$$\hat{f}(z_t|\Im_{t-1}) = \frac{1}{N}\sum_{i=1}^{N} f\left(z_t|M_t = \hat{M}_t^i; \hat{\Phi}\right) \tag{7.16}$$

conditional on the parameter vector $\hat{\Phi}$ obtained in the first step. As explained in Calvet et al. (2006) we simulate each vector M_t^i one-step forward and re-weight using an importance sampler (cf. sec. 2.5.2).

This estimation procedure is a special case of GMM, and thus, provides estimators that are consistent and asymptotically normally distributed (cf. Calvet et al., 2006, for more details.).

7.5. Empirical Application

7.5.1. Data

For the empirical study we use four stocks, namely Citigroup (C), International Business Machine (IBM), Bank of American (BAC) and Coca-Cola (KO), traded on the NYSE. The data were extracted from the Trade and Quote (TAQ) database available on the NYSE. The sample period correspond to July 2004 which has 21 trading days. The TAQ database is composed of two sections: The first section reports all trades and the second contains the best bid-ask prices posted by the market-makers. Here we use the data from the second section for computing the bid-ask spread. The bid-ask spread serves as an indicator for the market liquidity and is positively correlated with the transaction cost (cf. Hautsch, 2012). The role of the bid-ask spread in financial markets has been described in detail in information- or inventory-based models developed in the market microstructure literature. In fact, all these models consider the bid-ask spread as the only one instrument that gives the market maker a margin to avoid losses when trading with informed traders. Hautsch (2003) finds that the width of the bid-ask spread posted at the beginning of a spell can be used to predict a market-makers' assessment of liquidity risk.

A growing body of research devoted to the dynamical properties of the bid-ask spreads demonstrates how informational they are for the markets. Farmer et al. (2004), and later Mike and Farmer (2008) demonstrate that the pdf of the bid-ask spreads follows a power law with the exponent around 3. Long memory properties in the bid-ask spread series have been documented by Qu et al. (2007) and Mike and Farmer (2008). It is also well-documented and reported by Hautsch (2012) that the bid-ask spreads exhibit high persistence and long-range dependence. Research by

Qu et al. (2007) pointed out that the bid-ask spreads are characterized by monofractality, a feature that is different from multifractal structure found by a number of financial quantities (e.g. financial returns). Recently, the analysis of bid-ask spread returns and volatilities provides results that speak in favor of a strong multifractality in the spread returns and a presence of long-range cross-correlations between spread volatilities of different stocks (cf. Qiu et al., 2012). In this chapter we investigate the relationship between bid-ask spreads of different stocks by means of our bivariate MSMD model.

To do this we first compute the irregularly spaced bid-ask spreads s_t and rescale them as an average bid-ask spread $\varnothing s_t$ in time interval $\Delta t = 1$ (cf. Plerou et al., 2005):

$$\varnothing s_t = \frac{1}{\tau} \sum_{t=1}^{\tau} s_t, \tag{7.17}$$

where τ is the total number of quotes posted in the time interval $\Delta t = 1$.

7.5.2. Data Adjustment

High-frequency data exhibit strong seasonality due to the different trading activities observed during each trading day. These seasonal patterns have also been found in bid-ask spread (cf. Chung et al., 1999; Qu et al., 2007) and bid-ask spread volatilities (cf. Qiu et al., 2012). McInish and Wood (1992) pointed out that the bid-ask spreads display U-shaped patterns over the trading day. In other words, bid-ask spreads are higher at the opening and closing time than the rest of the trading day and these seasonal patterns are the opposite of that observed by the intertrade durations. It is well-known that such intraday patterns can lead to spurious results and have to be removed before using the data for any estimation. Different methods have been applied to remove the intraday patterns: A $\vartheta-$time scale method by Dacorogna et al. (1993), the maximal overlap discrete wavelet transform (MODWT) by Dacorogna et al. (2001). Here we adopt the methodologies recently proposed by Liu et al. (1999) that consist in segmenting the data set for each trading day that extends from 10 : 00 a.m. to 16 : 00 p.m. (360 minutes) into 360 consecutive 1 min intervals and then averaging over the total number of trading days. Formally, we have:

$$a_s(t') = \frac{1}{N} \sum_{i=1}^{N} s_i(t') \tag{7.18}$$

where $a_s(t')$ denotes the intra-day pattern of the bid-ask spread at time t' in the 360 continuous working minutes on the NYSE, i is the i^{th} trading day, and N is the number of trading days. The estimated seasonal pattern displays two kinds of shapes. The first one matches predicted trading patterns by information models that claim that bid-ask spreads are high at the opening time and decline throughout the day (cf. Glosten and Milgrom, 1985; Easley and O'Hara, 1987; Madhavan, 1992). This trading pattern can be explained by the diminution of the adverse selection problem the market-makers face throughout the trading day. The second ones correspond to that predicted by market-power models which show a crude reverse J-shape (cf. Stoll and Whaley, 1990; McInish and Wood, 1992) intra-day pattern (cf. *Fig.* 7.5).

7.5.3. Results

We first scrutinize the bid-ask spread data for the four stocks. We compute the empirical first, second, third and fourth moments for bid-ask spread data of each stock. The results are reported in Tables 7.1 and 7.2. The values for standard deviations are smaller than 1 and thus, indicate a presence of a low dispersion in the data. The bid-ask spread data for all four stocks exhibit high and positive skewness and excess kurtosis. Positive excess kurtosis indicates a peaked distribution while positive values for skewness show that the data are right skewed (cf. *Figs.* 7.8 through 7.11). *Figs.* 7.3 and 7.4 display the autocorrelation functions of the raw and adjusted data. To see whether the data for the bid-ask spread of the four stocks display long memory we employ the detrended fluctuation analysis to estimate the Hurst exponent (H) (cf. *Figs.* 7.3 and 7.4). The Hurst exponent belongs to the classical statistical methods for gauging the extent of persistency in a time series. It was introduced by English hydrologist Hurst (1951) in his seminal paper to address the problem of reservoir control near Nile River Dam in depth. In finance, the Hurst exponent (H) is popular due to the fact that it allows for classifying time series into different types and gaining insights into their dynamic properties. The Hurst exponent can only takes values in the interval (0, 1). A Hurst exponent around 0.5 corresponds to Brownian time series. A Hurst exponent value between 0 and 0.5 refers to time series that exhibit anti-persistent behavior and that between 0.5 and 1 is indicative of persistent behavior. The estimates of the Hurst exponent for the four stocks reported in Tables 7.1 and 7.2 are markedly greater than 0.5 and smaller than 1 (roughly speaking are located in interval (0.5, 1)), and thus, indicate the presence of persistence in the data used in this study.

We also compute for each stock the tail index (α) using the Hill's method (1975a)

$$\gamma_H = \frac{1}{\alpha} = \frac{1}{m}\sum_{i=1}^{m}\left[\ln(x_{(i)}) - \ln(x_{(m)})\right], \qquad (7.19)$$

where γ_H is a consistent estimate of the inverse of α. T is the sample size, m the number of observations located in the distribution's tail and the observations in the sample are put in descending order: $x_{(1)} \geq x_{(2)} \geq \cdots \geq x_{(m)} \geq \cdots \geq x_{(T)}$.

The results are presented in Tables 7.1 and 7.2. The tail index provides information about the behavior of the tails of all possible distributions and helps to categorize them in three classes. We distinguish between *thin-tailed* distributions that possess finite moments, a cumulative distribution function which declines exponentially in the tails and a tail index that approaches to ∞ ($\alpha \to \infty$), *fat-tailed* distributions whose cumulative distribution function declines with a power in the tails and exhibit finite and positive tail index ($\alpha > 0$), and *bounded* distributions that have no tails and whose tail index is negative ($\alpha < 0$) (cf. Dacorogna et al., 2001). In this study each of the four stocks has a tail index between 3.3 and 4.7 (roughly between 3 and 5). The estimates of the tail indexes are finite, positive and greater than 2. This is an evidence that the distribution of the bid-ask spread data used in this study belongs to *Fat-tailed* distributions and that the distribution would converge under aggregation to the Gaussian. As stressed by Racheva and Samorodnitsky (2003) when $H = 1/\alpha$, depending on the value of α the data exhibit either persistence or no memory. In this study it is clear that $1/\alpha$ takes values in the interval $(0, 0.5)$ and that the values obtained for H are greater than $1/\alpha$ which means that the data exhibit long-range dependence.

The estimates of the parameters in the bivariate MSMD model are reported in Table 7.3. We make use of *eq.* (7.9) to calculate the correlation coefficient ς and its standard error is obtained using the delta method (cf. Appendix *A*.5 for more details about the delta method). Note that for our empirical studies we use the specification of Chen et al. (2013). The estimates of the bivariate MSMD model for bid-ask spread pairs (BAC, C), (IBM, C) and (BAC, IBM) seem precisely estimated and are high significant. We do not obtain similar results for the last both parameters for the bid-ask spread pair (C, KO). Although the estimated values for ϱ are small, we can infer that the first three bid-ask spread pairs (cf. Table 7.3) comove or are dependent. It is clear that the intensity of the dependence is low, but significant and thus, cannot be neglected. This may be important for the market-maker to know stocks whose bid-ask spreads are simultaneously affected by arrival of new information in the market, and therefore, facilitate the risk management of their portfolio.

7.6. Conclusion

In this chapter we have proposed a bivariate MSMD model that can be used to model duration and price processes simultaneously. This model can also be utilized to analyze the covariation in microstructure variables. The new model is an extension of the univariate MSMD models developed by Chen et al. (2013) and Baruník et al. (2012), independently to bivariate settings. By analyzing tick-by-tick bid-ask spread data of four stocks traded on NYSE we find that the data exhibit long range dependence, fat tails, and self-similarity properties. Our new model can properly capture all these features and helps us to identify covariation in bid-ask spreads of different stocks. The market-makers or market participants have to manage huge portfolios over the trading day. So, these results can help them to better manage the market risk.

Table 7.1.: Descriptive statistics of raw bid-ask spread data

	KO	BAC	IBM	C
Min	0.010	0.010	0.010	0.010
Max	0.273	0.177	0.386	0.262
Mean	0.018	0.023	0.026	0.017
Std	0.010	0.011	0.013	0.008
Skewness	6.851	3.410	5.266	7.179
Kurtosis	107.846	29.878	93.523	142.330
Hurst Exponent(H)	0.754	0.711	0.754	0.821
Tail Index(α)	3.595	4.116	3.990	3.341
Q(10)	2842.2	1464.5	1822.4	2354.5
Q(100)	4044	2156.5	4090.6	5775

Table 7.2.: Descriptive statistics of adjusted bid-ask spread data

	KO	BAC	IBM	C
Min	0.268	0.315	0.270	0.364
Max	7.332	6.203	9.685	9.540
Mean	1.002	1.001	1.001	1.000
Std	0.467	0.460	0.462	0.412
Skewness	3.791	2.374	2.892	4.108
Kurtosis	34.405	16.033	29.147	43.963
Hurst Exponent(H)	0.741	0.712	0.747	0.823
Tail Index(α)	4.191	4.681	4.336	3.913
Q(10)	2649.1	1601.2	2062	2490.8
Q(100)	3921.3	2379.8	4981.6	7307.4

Table 7.3.: Two-Step Bivariate MSMD Parameters estimation

	BAC-C	IBM-C	BAC-IBM	C-KO
		Estimates		
$\bar{\lambda}_1$	1.126***	1.118***	1.126***	0.925***
	(**0.002**)	(**0.005**)	(**0.002**)	(**0.002**)
$\bar{\lambda}_2$	1.093***	1.089***	1.123***	1.066***
	(**0.002**)	(**0.004**)	(**0.003**)	(**0.003**)
m_1	1.086***	1.076***	1.086***	1.097***
	(**0.002**)	(**0.003**)	(**0.002**)	(**0.002**)
m_2	1.029***	1.011***	1.087***	1.156***
	(**0.002**)	(**0.004**)	(**0.003**)	(**0.003**)
b	1.710***	2.210***	1.710***	3.833***
	(**0.001**)	(**0.002**)	(**0.001**)	(**0.001**)
γ	0.783***	0.777***	0.783***	0.564***
	(**0.002**)	(**0.007**)	(**0.003**)	(**0.004**)
σ_1	0.261***	0.282***	0.261***	0.251***
	(**0.008**)	(**0.019**)	(**0.009**)	(**0.008**)
σ_2	0.225***	0.238***	0.259***	0.254***
	(**0.007**)	(**0.017**)	(**0.010**)	(**0.006**)
ς	0.102***	0.141***	0.127***	0.111
	(**0.039**)	(**0.058**)	(**0.082**)	(**0.129**)
ϱ	0.191***	0.203***	0.148***	0.111
	(**0.018**)	(**0.055**)	(**0.053**)	(**0.124**)

Note that we use for the estimation $k = 6$ and standard errors in parentheses are computed as described in Calvet *et al.* (2006). *** indicate that the parameters are significant at the 1% level. The numbers in bold in parentheses are standard errors of the estimations.

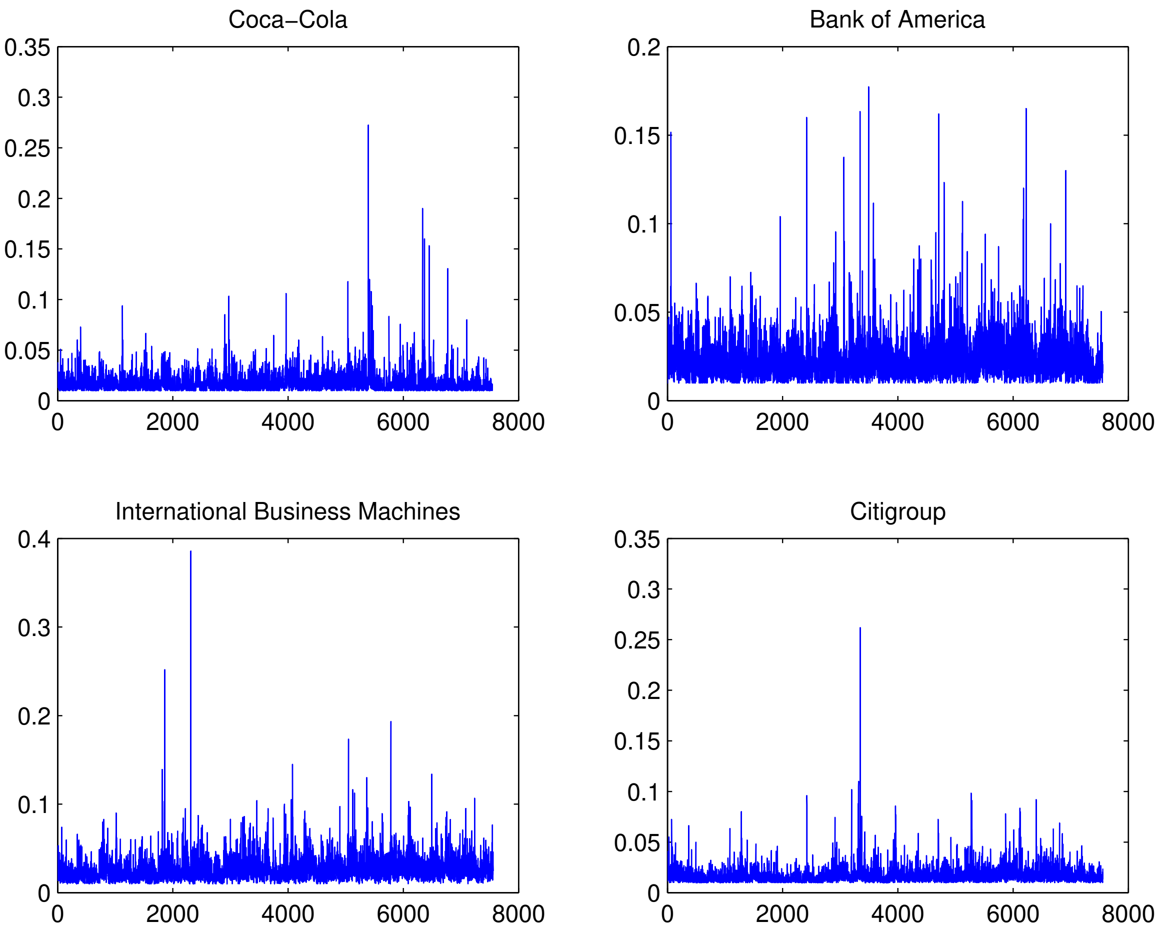

Figure 7.1.: Plot of raw bid-ask spread data for the four stocks

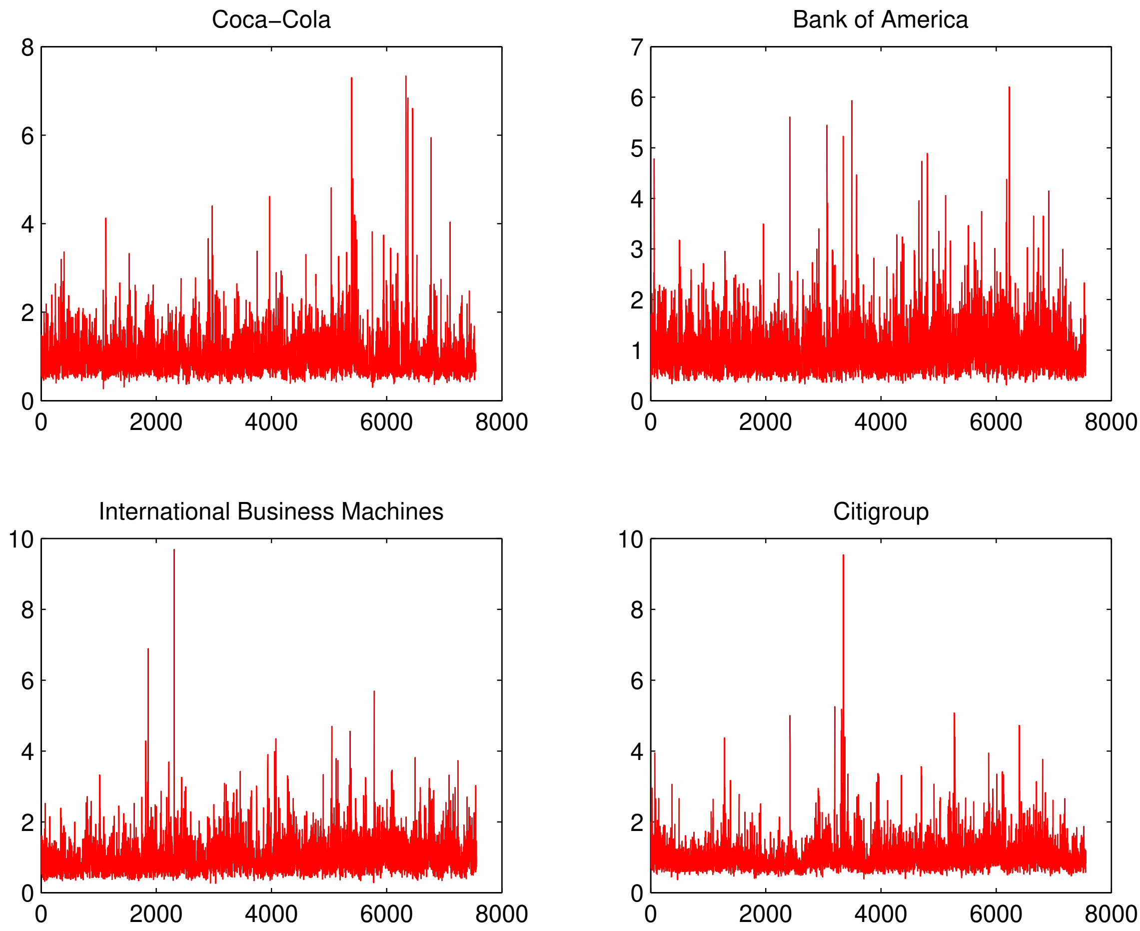

Figure 7.2.: Plot of adjusted bid-ask spread data for the four stocks

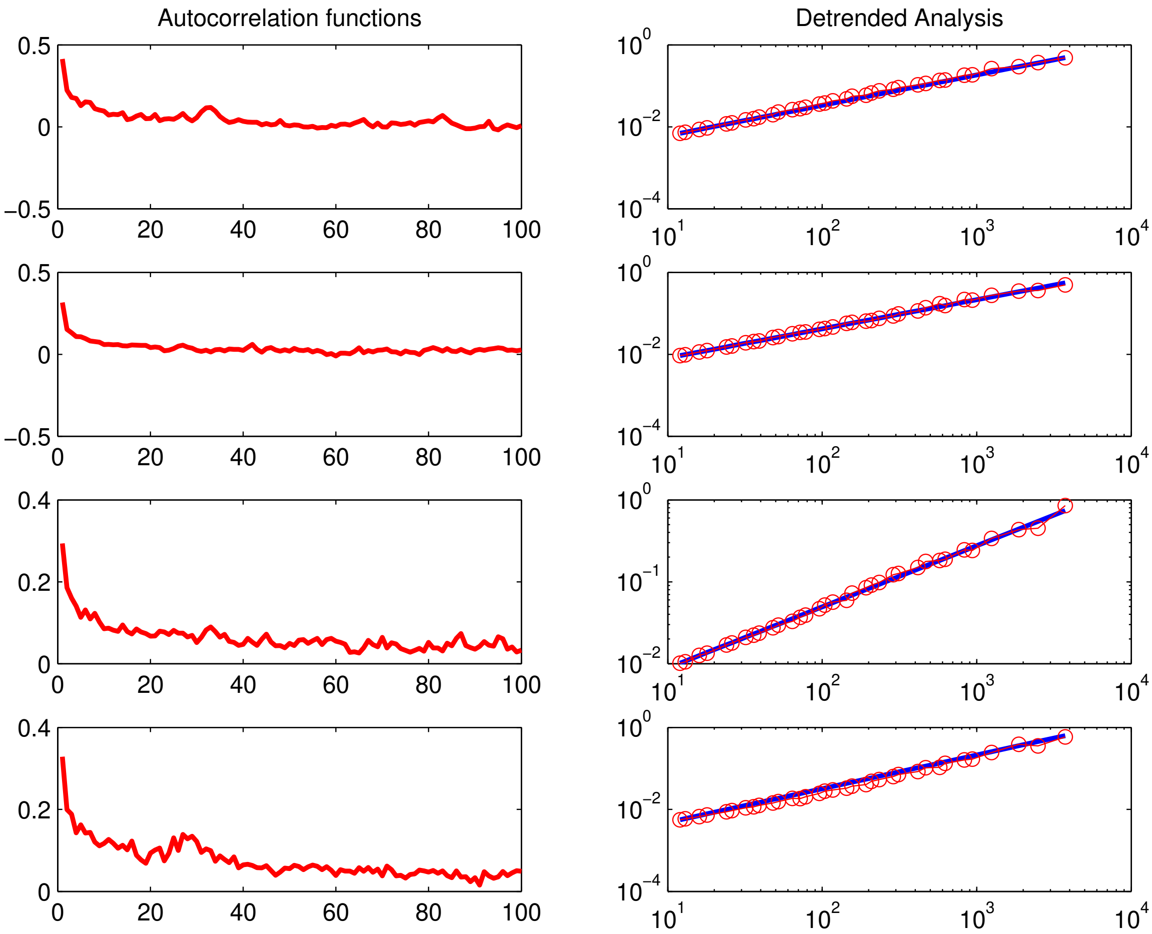

Figure 7.3.: Illustration of the long-term dependence observed in the bid-ask spread raw data for Coca-Cola (left upper panel), Bank of America (left first central panel), International Business Machines (left second central panel) and Citigroup (lower left panel). The determination of the corresponding Hurst exponent H is displayed in the right-hand panels.

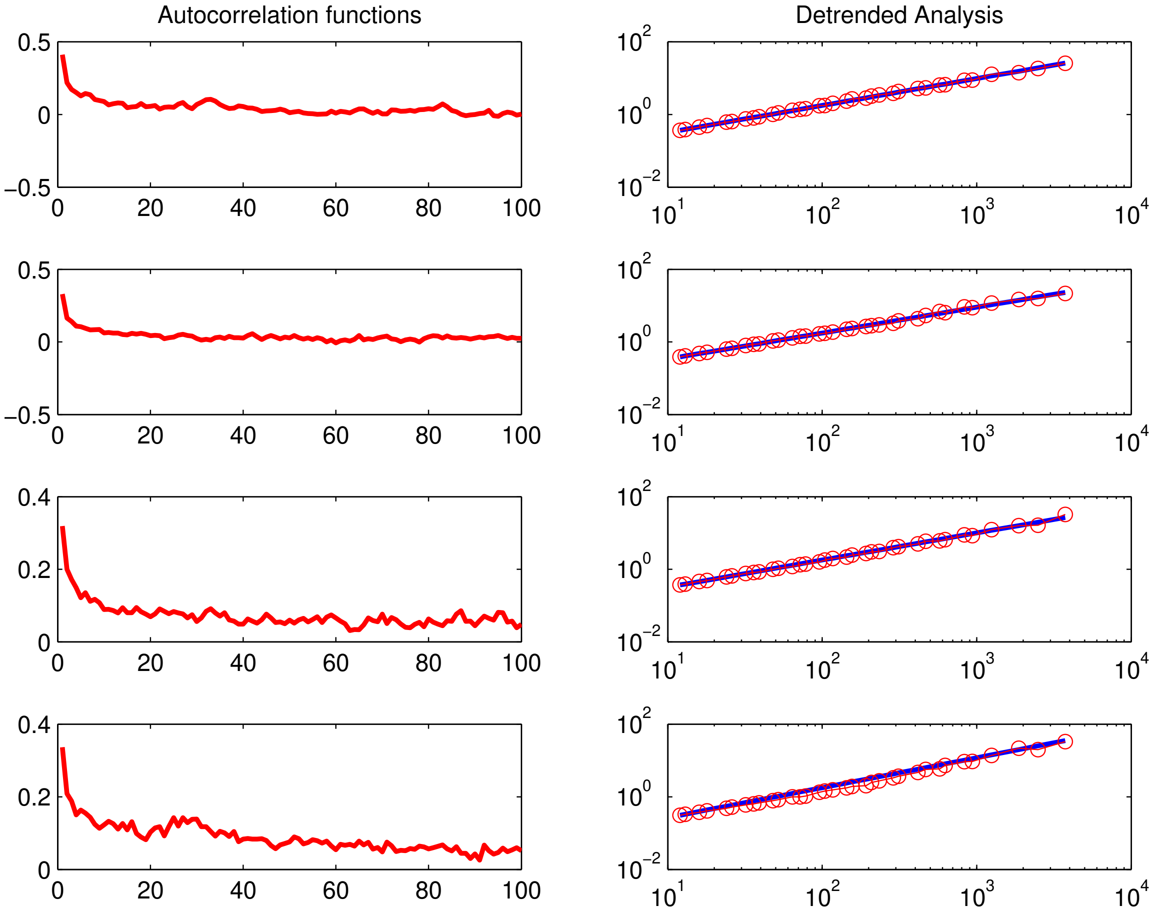

Figure 7.4.: Illustration of the long-term dependence observed in the bid-ask spread adjusted data for Coca-Cola (left upper panel), Bank of America (left first central panel), International Business Machines (left second central panel) and Citigroup (lower left panel). The determination of the corresponding Hurst exponent H is displayed in the right-hand panels.

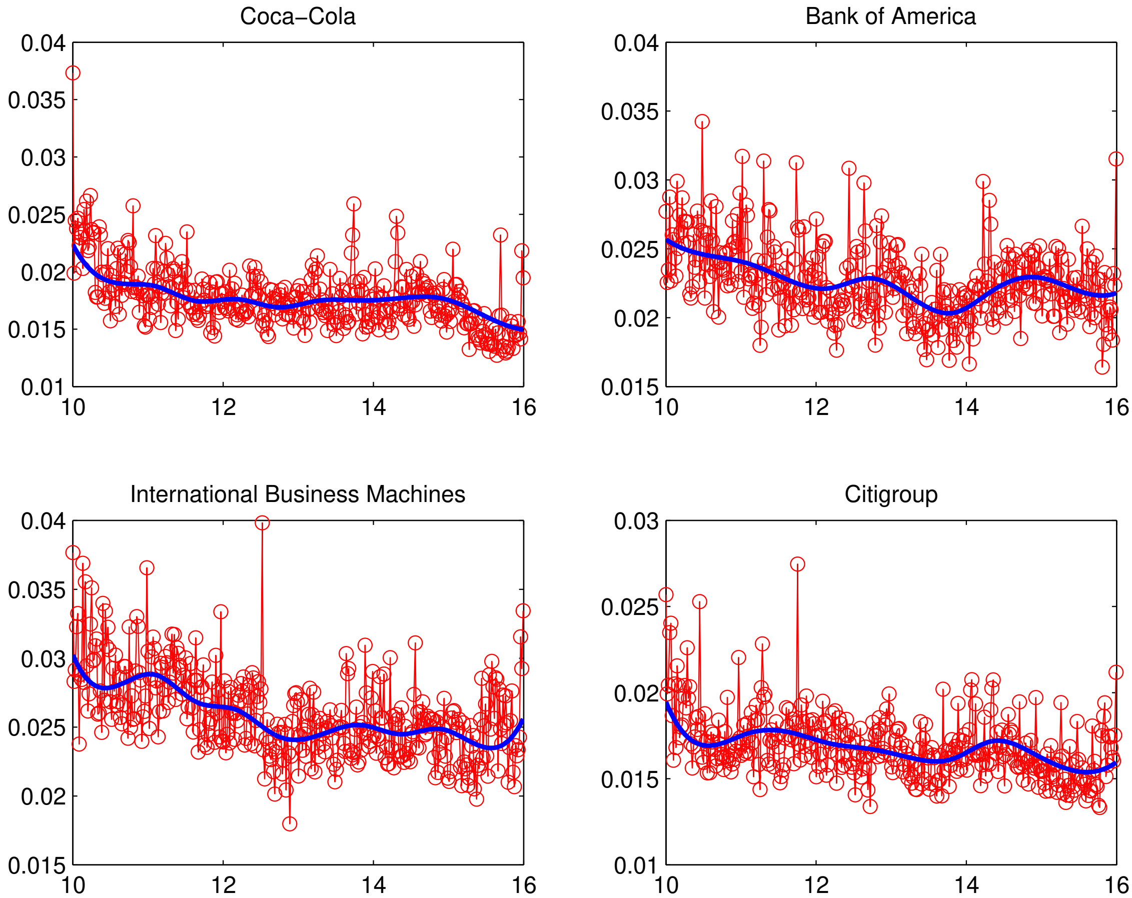

Figure 7.5.: Intraday pattern of bid-ask spread

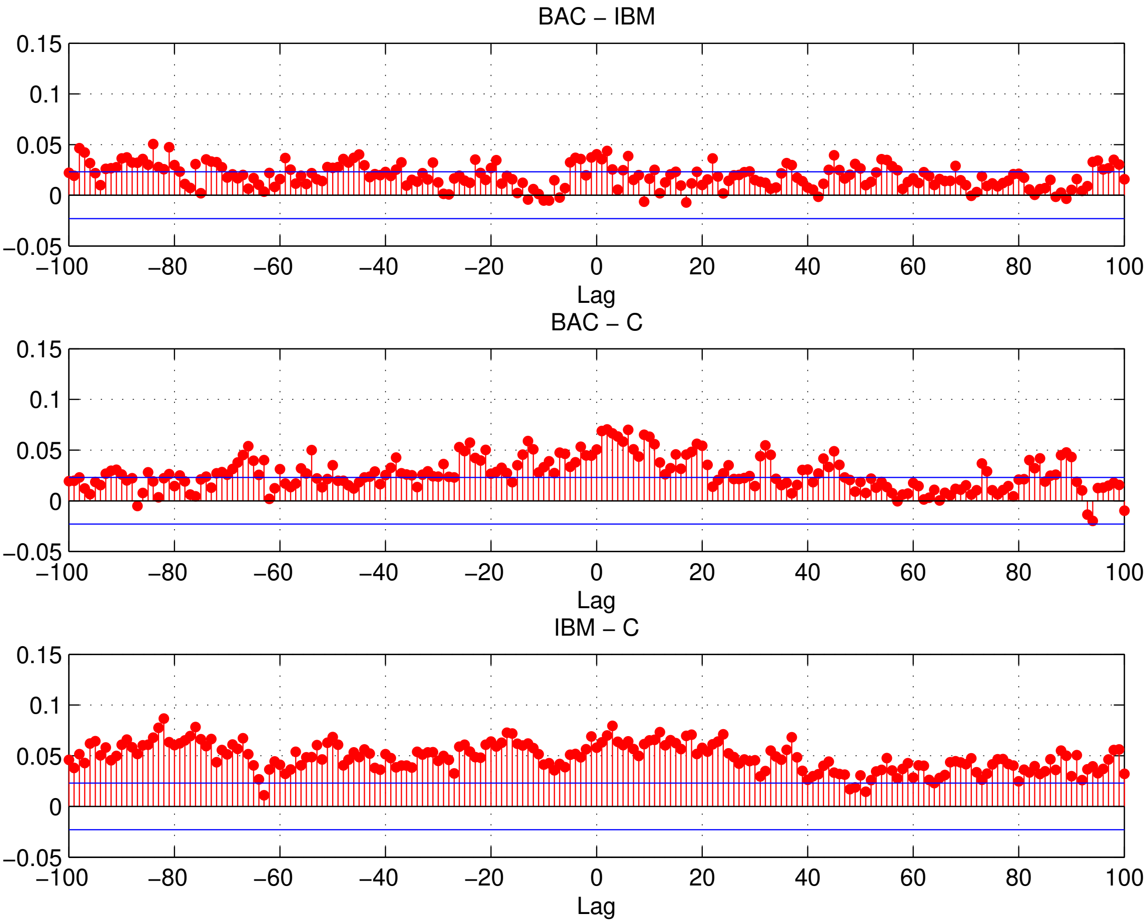

Figure 7.6.: Plot of sample cross-correlation

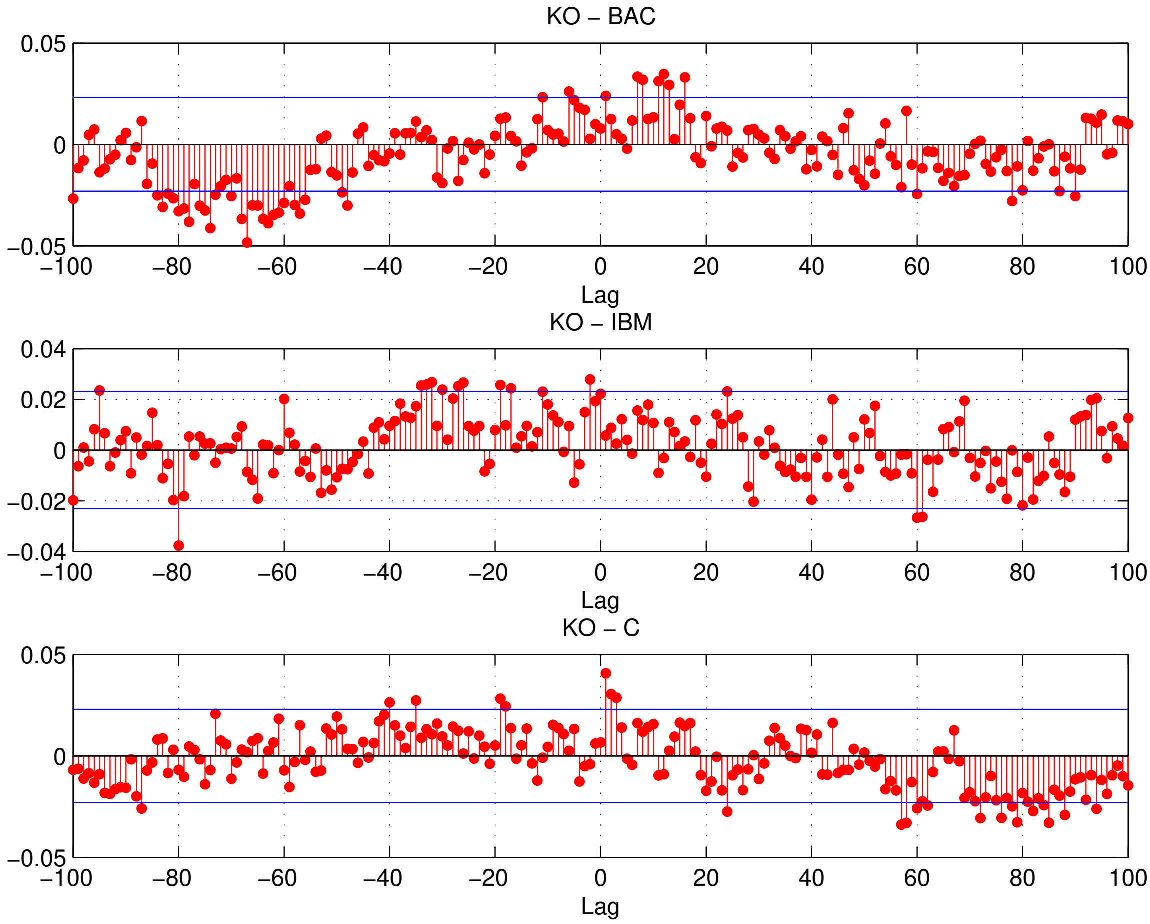

Figure 7.7.: Plot of sample cross-correlation

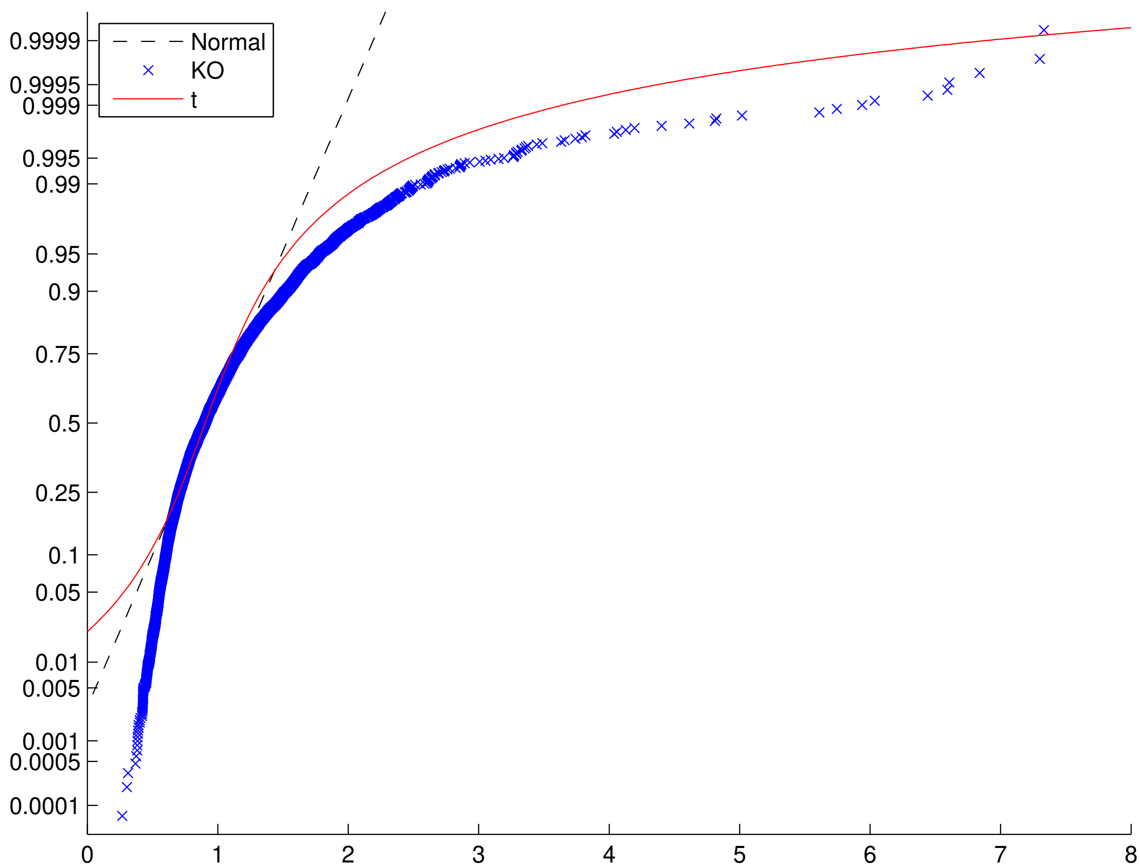

Figure 7.8.: Probability Plot of KO-spread compared to Normal and Student distributions

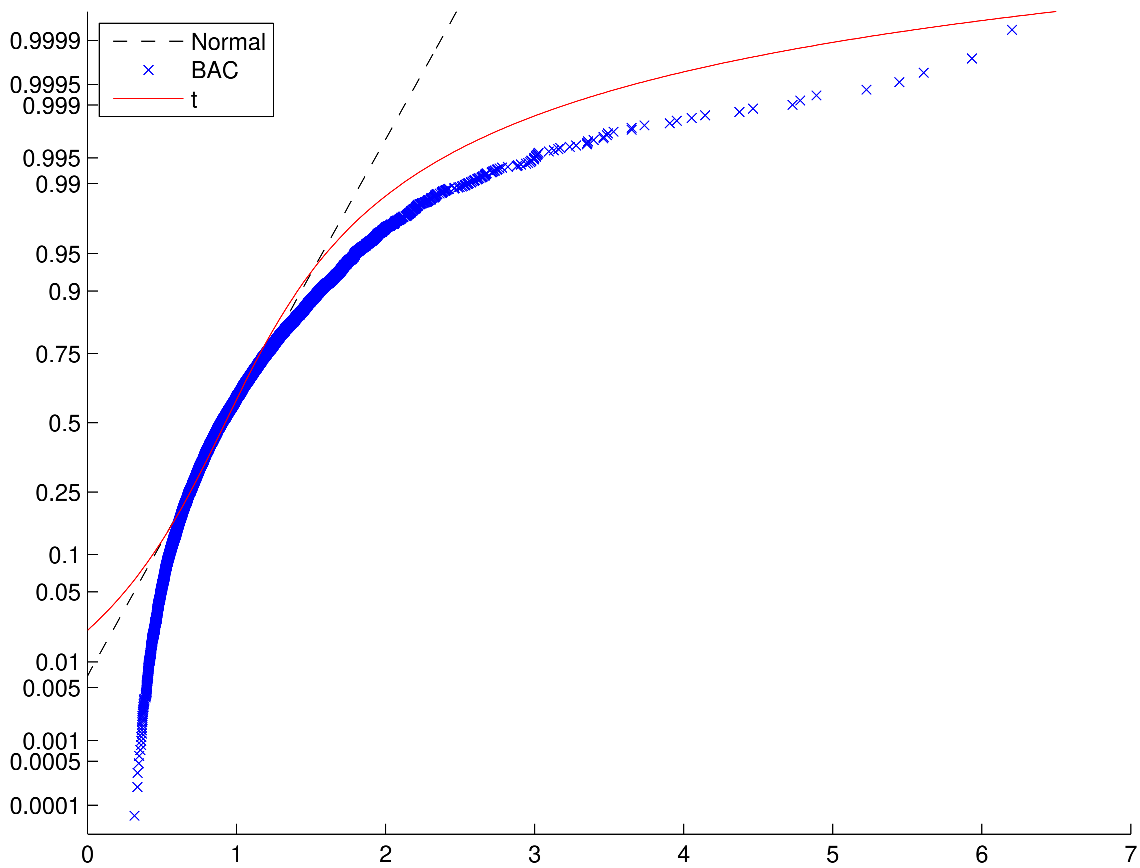

Figure 7.9.: Probability Plot of BAC-spread compared to Normal and Student distributions

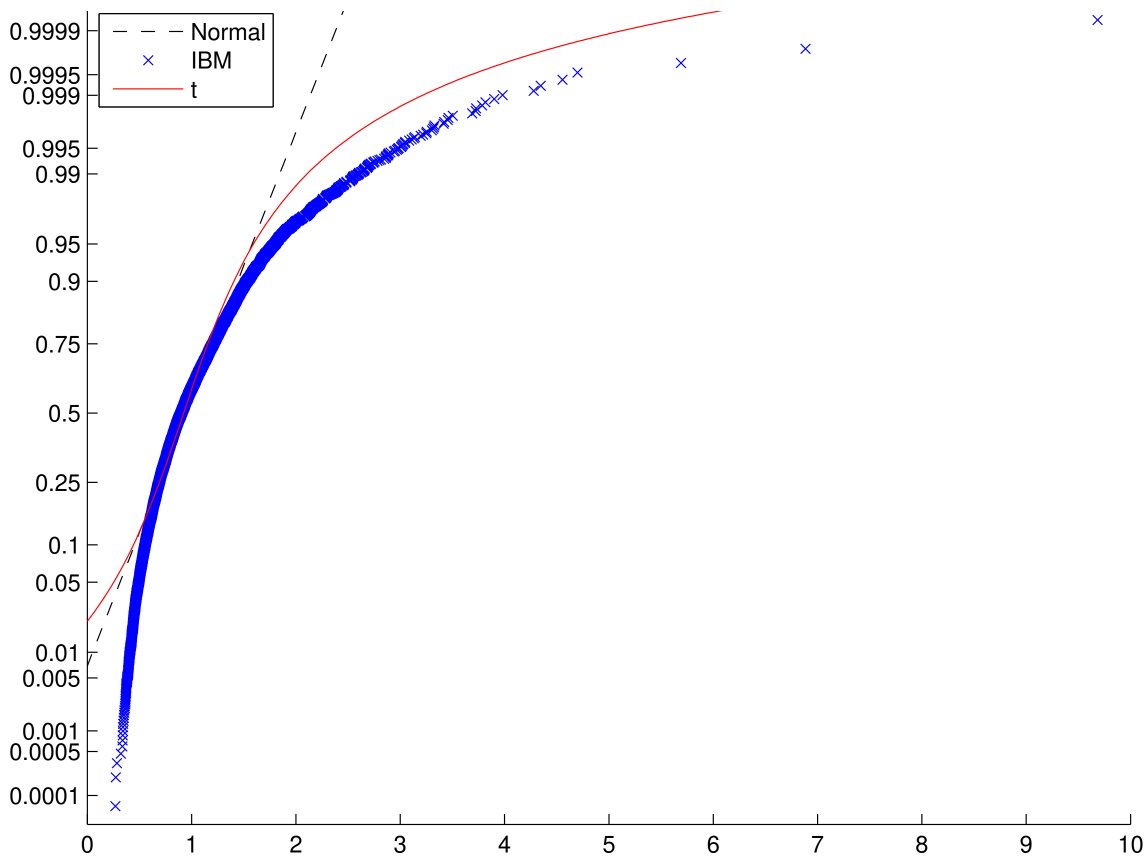

Figure 7.10.: Probability Plot of IBM-spread compared to Normal and Student distributions

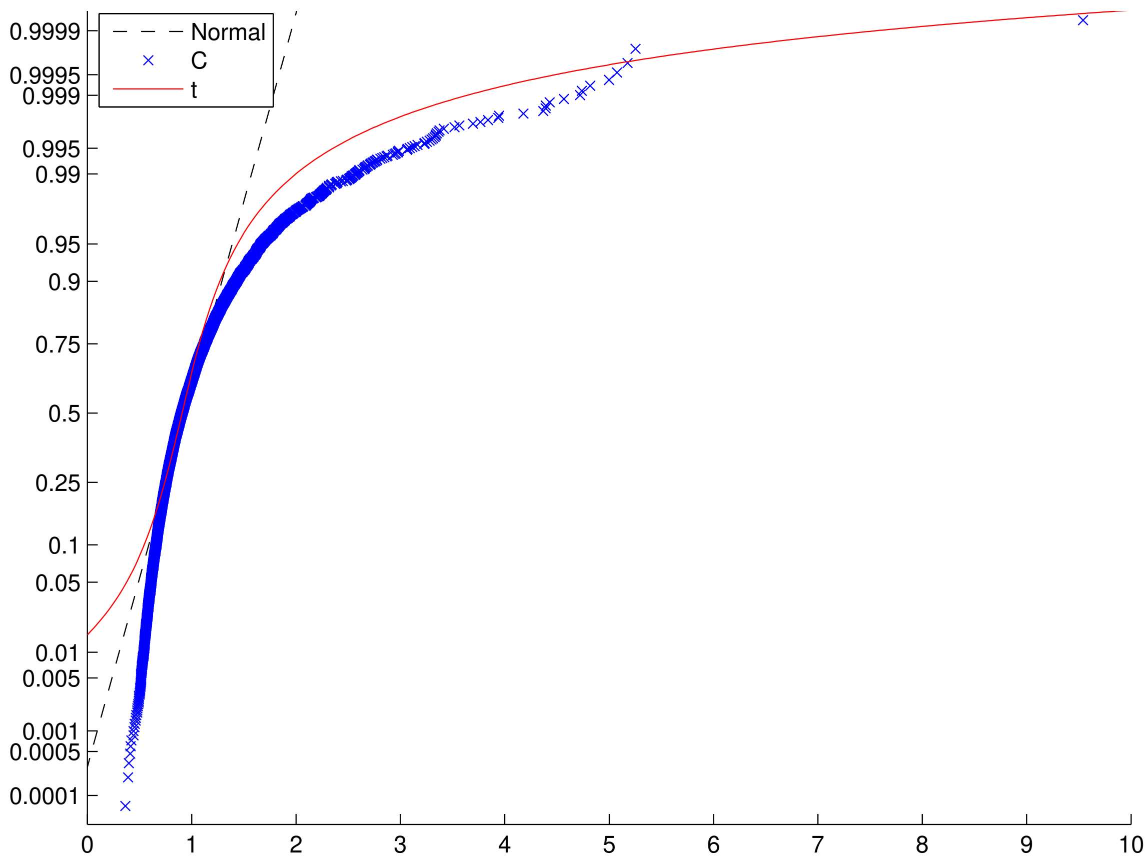

Figure 7.11.: Probability Plot

8. Forecasting Intraday Value-at-Risk Using Markov-Switching Multifractal Duration Model

8.1. Introduction

Over the last decade the interest to forecasting irregularly spaced intraday value-at-risk (ISIVaR) has been grown due to the rapid change in the trading environment. The automatization of financial markets and the improvement in information technology (IT) allow active market participants to execute transactions at fine time intervals. Traditional trading strategies such as *buy-and-hold* have been abandoned in favor of day trading and now the market is dominated by high frequency traders. In this new environment where prices of equities, commodities, exchange rates and interest rates interruptedly change, and thus, continuously cause a change in market risk, risk measurement plays a central role. Indeed, capital adequacy rules are determined by risk levels (cf. Basel II and III). It becomes important for the market participants, especially high-frequency traders and financial institutions whose investment horizons are less than 5 or 10 minutes, to be able to estimate and control their exposure to market risk.

The use of value-at-risk as a tool for financial risk assessment in the market is popular by regulators and owners of financial institutions, because it helps to quantify the maximal amount to be lost on a portfolio over a given period of time, at a certain confidence level. Different value-at-risk models and sophisticated statistical methodologies for their assessment have been proposed in the literature (cf. Christoffersen and Pelletier, 2004; Engle and Manganelli, 2004; Giacomini and Komunjer, 2005; Haas, 2005; Berkowitz et al., 2011), but unfortunately, however, all these models and techniques are not appropriate for analyzing ISIVaR. This is due to the fact that financial data used for computing ISIVaR are irregularly spaced.

Until now, less research has been done to develop tools that can allow to assess the market risk at intraday time horizons. Fixed interval models as such Normal GARCH, Student GARCH, and RiskMetrics can be used to forecast intraday value-at-risk (IVaR) for a given time interval, for instance 15 or 30 minutes (cf. Giot, 2005), but they are not convenient for ISIVaR, because they cannot capture the irregularly spacing feature of financial duration data. Meanwhile high-frequency models have been developed in the literature, for instance the standard ACD model (cf. Engle and Russell, 1998) and its extensions, cf. sec. 3.3 and Pacurar (2008), albeit the non-existence of an appropriate backtesting makes it difficult to test the performance of these ISIVaR models.

Giot (2005) was the first, to our knowledge, working on the forecasting intraday market risk in a conditional value-at-risk framework. To estimate the conditional intraday volatility and compute a conditional parametric ISIVaR, he applied the Log-ACD model to price durations. Unfortunately, however, the performance of the Log-ACD model compared to that of Normal GARCH or Student GARCH in the fixed interval framework is poor. One can argue that the poor performance of the Log-ACD model is due to the fact that he used an average of the ISIVaRs on a given time interval as a regularly spaced intraday VaR for backtesting purposes. Another work on the same issue of intraday market risk measurement has also been performed by Dionne et al. (2009). Their methodology consists in combining a Log-ACD-ARMA-EGARCH model with an intraday Monte Carlo simulation. The lack of an appropriate backtesting for ISIVaR models obliges them to make use of an average of ISIVaRs in 15 minutes interval as a regularly spaced intraday VaR for testing the performance of their model.

In this chapter we forecast the ISIVaR in a semi-parametric framework using the Markov-switching multifractal duration (MSMD) model and the generalized gamma autoregressive conditional duration (GGACD) model. We choose the GGACD model because findings by Bauwens et al. (2004) and in sec. 5.5.2.2 show that the GGACD properly fits price durations and outperforms the more complicated models such as the stochastic conditional duration (SCD) model of Bauwens and Veredas (2004), the stochastic volatility duration (SVD) model of Ghysels et al. (2004) and the FIACD model of Jasiak (1998). To evaluate and compare their forecasting abilities we employ a GMM duration-based test recently developed by Candelon et al. (2011). In contrast to the duration-based approach of Christoffersen and Pelletier (2004) that ignored the discrete nature of the problem and that has been designed using the continuous Weibull distribution, the GMM duration-based test is implemented based on the geometric distribution. Haas (2005) demonstrated throughout Monte Carlo simulations that the use of the continuous Weibull distribution instead of an appropriate discrete distribution would negatively affect the power of the test. It is important to note that we can directly apply the GMM duration-based test to the ISIVaRs.

Papers by Baruník et al. (2012) and Segnon and Lux (2012) shed light on the ability of the MSMD model to reproduce financial price durations. We think that the MSMD model can be of great importance for financial institutions in terms of avoiding the under- or overestimation of the risk.

The rest of the chapter is organized as follows. We present in Section 8.2 intraday volatility associated to price duration models. We define in Section 8.3 the irregularly spaced intraday VaR and Section 8.4 describes the GMM duration-based backtesting procedures. An empirical application is presented in Section 8.5 and we conclude in Section 8.6.

8.2. Intraday Volatility

Here we do not present again the Markov switching multifractal duration (MSMD) and the standard autoregressive conditional duration model. We refer the reader to sec. 3.4 and 3.3.1, respectively.

8.2.1. Instantaneous Price Changes Volatility

In their seminal work Engle and Russell (1998) formally showed how the instantaneous intraday volatility can be linked to the conditional hazard rate function of price durations. Following them we first define the conditional intensity function as

$$\vartheta\left[T|N(T), T_1, \ldots, T_{N(T)}\right] = \lim_{\Delta T \to 0} \frac{Pr\left[N(T+\Delta T) > N(T)|N(T), T_1, \ldots, T_{N(T)}\right]}{\Delta T}, \tag{8.1}$$

$Pr\left[N(T+\Delta T) > N(T)|N(T), T_1, \ldots, T_{N(T)}\right]$ is the conditional probability of an event in $(T, T+\Delta T)$ given the history $\Im_T = \left[N(T), T_1, \ldots, T_{N(T)}\right]$ of events up to time T, and $N(T)$ is the number of events that occurred at time T.

We can rewrite *eq.* (8.1) as

$$\lim_{\Delta T \to 0} \frac{Pr\left[\Delta N_{(T,T+\Delta T)} > 0|\Im_T\right]}{\Delta T} = \lim_{\Delta T \to 0} \frac{Pr\left[x_t \in (T, T+\Delta T) \mid x_t > T, \Im_T\right]}{\Delta T}, \tag{8.2}$$

where x_t is the waiting time for the t^{th} event conditional on $\Im_T$. Once again, we can rewrite *eq.* (8.2) as

$$\lim_{\Delta T \to 0} \frac{Pr(\Delta N_{(T,T+\Delta T)} > 0|\Im_T)}{\Delta T} = \lim_{\Delta T \to 0} \frac{F(T+\Delta T \mid \Im_T) - F(T \mid \Im_T)}{1 - F(T \mid \Im_T)}. \tag{8.3}$$

By passing the limit we obtain

$$\vartheta\,(T|\Im_T) = \frac{f(T \mid \Im_T)}{1 - F(T \mid \Im_T)}, \tag{8.4}$$

where F is the cumulative distribution function and f the probability density function of the waiting time x_t conditional on $\Im_T$. The right-hand side of the last line in *eq.* (8.3) corresponds to the definition of the conditional hazard function. It is clear that the conditional intensity function and the hazard function are interchangeable.

Engle and Russell (1998) defined the instantaneous volatility as

$$\sigma^2(T) = \lim_{\Delta T \to 0} E\left\{\frac{1}{\Delta T}\left[\frac{P(T+\Delta T) - P(T)}{P(T)}\right]^2\right\}, \tag{8.5}$$

with $P(T)$ a stock price related to the arrival time T. So, we can define the conditional instantaneous intraday volatility as

$$\sigma^2(T|T_{N(T)},\ldots,T_1) = \left(\frac{\iota_p}{P(T)}\right)^2 \vartheta(T|T_{N(T)},\ldots,T_1), \tag{8.6}$$

where $P(T)$ represents the bid-ask midpoint, ι_P is a constant, and $\vartheta(T|T_{N(T)},\ldots,T_1)$ is the conditional intensity function or the conditional hazard function.

8.2.2. Conditional Hazard Functions

In the following we derive the pertinent conditional hazard functions associated with the Binomial MSMD model and the GGACD(1,1) model.

- *The Binomial MSMD model*:

  The innovation in the MSMD model is standard exponential distributed. Using a transform of random variables technique (cf. Appendix A.1) the conditional density function of a price duration x_t given a state intensity M_t, $f(x_t|M_t = m^i)$, follows an exponential distribution. Formally, we have

$$f(x_t|M_t = m^i) = \lambda_t(m^i)\exp[-\lambda_t(m^i)x_t]. \tag{8.7}$$

  So, one can obtain the conditional hazard function as follows. We proceed in two steps. In the first step we provide the mathematical formula of the conditional density function of a duration x_t given the past history $\Im_{t-1}$. We have:

$$\begin{aligned} f(x_t|\Im_{t-1}) &= \sum_{i=1}^{n} f(x_t|M_t = m^i)Pr\left(M_t = m^i \mid \Im_{t-1}\right) \\ &= \sum_{i=1}^{n} \lambda_t(m^i)\exp[-\lambda_t(m^i)x_t]Pr\left(M_t = m^i \mid \Im_{t-1}\right), \end{aligned} \tag{8.8}$$

  where $Pr\left(M_t = m^i \mid \Im_{t-1}\right)$ represents the probability of M_t conditional on the past history and satisfies this condition: $\sum_{i=1}^{n} Pr\left(M_t = m^i \mid \Im_{t-1}\right) = 1$.

  In the second step we derive the corresponding cumulative probability function that is given by

$$\begin{aligned} F(x_t|\Im_{t-1}) &= \int_0^{x_t} f(u_t|\Im_{t-1})du_t \\ &= \int_0^{x_t} \sum_{i=1}^{n} f(u_t|M_{t-1} = m^i)Pr(M_{t-1} = m^i|\Im_{t-1})du_t. \end{aligned} \tag{8.9}$$

  $f(u_t|M_{t-1} = m^i) = \lambda_t(m^i)\exp[-\lambda_t(m^i)u_t]$ is Lebesgue integrable. This allows us to apply the linearity rule of integration, i.e., the integral of a sum of functions is the sum of the integrals of the functions. By applying this rule, $F(x_t|\Im_{t-1})$ becomes

$$\begin{aligned} F(x_t|\Im_{t-1}) &= \sum_{i=1}^{n} Pr(M_{t-1} = m^i|\Im_{t-1}) \int_0^{x_t} \lambda_t(m^i) \exp[-\lambda_t(m^i)u_t]du_t \\ &= \sum_{i=1}^{n} Pr(M_{t-1} = m^i|\Im_{t-1}) \left(1 - \exp[-\lambda_t(m^i)x_t]\right) \\ &= \sum_{i=1}^{n} Pr(M_{t-1} = m^i|\Im_{t-1}) - \sum_{i=1}^{n} Pr(M_{t-1} = m^i|\Im_{t-1}) \exp[-\lambda_t(m^i)x_t] \\ &= 1 - \sum_{i=1}^{n} Pr(M_{t-1} = m^i|\Im_{t-1}) \exp[-\lambda_t(m^i)x_t]. \end{aligned} \tag{8.10}$$

Given the conditional cumulative distribution function $F(x_t|\Im_{t-1})$ the conditional hazard function $g(x_t|\Im_{t-1})$ is defined as

$$g(x_t|\Im_{t-1}) = \frac{\sum_{i=1}^{n} \alpha_i \lambda_t(m^i) \exp[-\lambda_t(m^i)x_t]}{\sum_{i=1}^{n} \alpha_i \exp[-\lambda_t(m^i)x_t]}, \tag{8.11}$$

with $\alpha_i = Pr(M_{t-1} = m^i|\Im_{t-1})$.

Thus, the conditional instantaneous intraday volatility in the MSMD model becomes

$$\sigma^2(x_t|\Im_{t-1}) = c(x_t) \frac{\sum_{i=1}^{n} \alpha_i \lambda_t(m^i) \exp[-\lambda_t(m^i)x_t]}{\sum_{i=1}^{n} \alpha_i \exp[-\lambda_t(m^i)x_t]}, \tag{8.12}$$

where $c(x_t)$ is a time varying scaling factor $\left(c(x_t) = (\iota_p/P(x_t))^2\right)$.

- *The generalized gamma ACD model*:

  Here we assumed that the innovation is generalized gamma distributed and its probability density function is given by

$$GG(\xi_t; \eta, \alpha) = \begin{cases} \frac{\eta(\xi_t)^{\eta\alpha-1}}{\theta^{\eta\alpha}\Gamma(\alpha)} \exp\left[-\left(\frac{\xi_t}{\theta}\right)^{\eta}\right] & \text{if } \xi \geq 0 \\ 0 & \text{if } \xi \leq 0 \end{cases} \tag{8.13}$$

  where $\eta, \alpha > 0$, $\Gamma(\cdot)$ denotes the gamma function and $\theta = \Gamma(\alpha)/\Gamma\left(\alpha + \frac{1}{\eta}\right)$.

  Using a transform of random variables technique (cf. Appendix *A*.1) the corresponding probability density function $f(x_t|\Im_{t-1})$ of the duration x_t is given by

$$f(x_t|\Im_{t-1}) = \frac{\eta(x_t)^{\eta\alpha-1}}{\theta_t^{\eta\alpha}\Gamma(\alpha)} \exp\left[-\left(\frac{x_t}{\theta_t}\right)^{\eta}\right], \tag{8.14}$$

  where $\theta_t = \Psi_t \frac{\Gamma(\alpha)}{\Gamma(\alpha+1/\eta)}$ is the time-varying scale parameter. Ψ_t is defined as in *eq.* (5.1).

The associated conditional survivor function involves the incomplete gamma function and is defined as

$$S(x_t|\Im_{t-1}) = 1 - I\left[\alpha, \left(\frac{x_t}{\theta_t}\right)^{\eta}\right], \tag{8.15}$$

where $I(\alpha, z) = \int_z^{\infty} t^{\alpha-1}\exp(-t)dt$, with $z = \left(\frac{x_t}{\theta_t}\right)^{\eta}$.

eqs. (8.14) and (8.15) lead to the hazard (or intensity) function that can be formalized as

$$g(x_t|\Im_{t-1}) = \frac{\frac{\eta(x_t)^{\eta\alpha-1}}{\theta_t^{\eta\alpha}\Gamma(\alpha)}\exp[-(x_t/\theta_t)^{\eta}]}{1 - I(\alpha, (x_t/\theta_t)^{\eta})}. \tag{8.16}$$

8.3. Irregularly Spaced Intraday VaR

Irregularly spaced intraday VaR (ISIVaR) is an extension of VaR to irregularly spaced intraday returns (ISIR) that are computed as $r(t) = \ln(P_t) - \ln(P_{t-1})$ using average prices[1] at which quotes are posted by market makers. we use the term *irregularly spaced* returns due to the fact that quotes are recorded continuously so that the observed bid and ask prices are no longer equidistantly time-spaced data. As in Colletaz et al. (2007) we defined the irregularly spaced intraday VaR (ISIVaR) for a shortfall probability α as a couple $(\lambda_{t|t-1}, \text{ISIVaR}_{t|t-1}(\alpha))$ that gives simultaneously two main information, namely, the expected duration for t^{th} price change, $1/\lambda_{t|t-1}$, and the corresponding level of risk $\text{ISIVaR}_{t|t-1}(\alpha)$ as such

$$Pr[r_t < -\text{ISIVaR}_{t|t-1}(\alpha)] = \alpha, \quad \forall t \in \mathbb{Z}. \tag{8.17}$$

Without loss of generality let us formalize the irregularly spaced intraday returns as

$$r_t = \sigma(x_t|\Im_{t-1})\xi_t, \tag{8.18}$$

where $\sigma(x_t|\Im_{t-1})$ is the instantaneous price change volatility and ξ_t an *i.i.d.* innovation with zero mean and unit variance.

The market risk for the t^{th} price variation can be obtained as

$$\text{ISIVaR}_{t|t-1}(\alpha) = -F^{-1}(\alpha)\frac{\iota_p}{P(x_t)}[g_{t|t-1}(x_t)]^{1/2}, \tag{8.19}$$

where $F(\cdot)$ is the cumulated distribution function of variable ξ_t and ι_p is the size of the cumulative absolute price change and exogenously fixed.

For the MSMD model the 1-ahead out-of-sample ISIVaR forecast for price change number $(\tau + 1)$ can easily be generated using the information contained in the τ price changes as follows.

[1] P_t the bid-ask mid-point

1. The first step consists in estimating the MSMD model with the adjusted duration $\{x_t\}_{t=1}^{\tau}$. With the estimated parameters, we calculate the 1-ahead out-of-sample $\hat{f}(x_{\tau+1})$ and $1 - \hat{F}(x_{\tau+1})$, and the conditional 1-ahead hazard function as

$$\hat{g}(x_{\tau+1}) = \frac{\hat{f}(x_{\tau+1})}{1 - \hat{F}(x_{\tau+1})} \tag{8.20}$$

and the forecast value of volatility is given by

$$\hat{\sigma}^2(x_{\tau+1}|\mathfrak{I}_\tau) = \hat{g}(x_{\tau+1})\left[\frac{\iota_p}{P(T_\tau)}\right]^2. \tag{8.21}$$

2. With the results of the above MSMD model, compute the series of in-sample standardized series of returns as

$$\hat{\xi}_t = \frac{r_t}{\hat{\sigma}(x_t|\mathfrak{I}_{t-1})}, \tag{8.22}$$

where $\hat{\sigma}(x_t|\mathfrak{I}_{t-1})$ are the series of computed in-sample volatilities.

3. Compute the empirical $\alpha-$quantile q of in-sample standardized series of returns as follows

$$q = percentile\left(\{\xi_t\}_{t=1}^{\tau}, 100\alpha\right), \tag{8.23}$$

and at the end calculate the value of ISIVaR for the next price change with a given shortfall probability α as

$$\text{ISIVaR}_{\tau+1}(\alpha) = -q\hat{\sigma}(x_{\tau+1}|\mathfrak{I}_\tau). \tag{8.24}$$

We applied the algorithm described above to the standard generalized gamma ACD (1,1) model. Note that it can also be applied to the MSMD model of Baruník et al. (2012), the ACD models with different distributional assumptions for the innovations and its extensions without changing any step.

8.4. GMM Duration-Based Test Approach

One important question remains how to test the predictive ability of the ISIVaR models. Here we adopt the GMM duration-based test approach developed by Candelon et al. (2011). The Test approach is developed based on orthonormal polynomials associated to the geometric and exponential distribution in the GMM framework and follows the basic idea of Bontemps and Meddahi (2005, 2012). The benefit of the GMM duration-based test is that it allows by means of the choice of moments conditions to test the unconditional coverage *(uc)*, independence *(ind)*, and conditional coverage *(cc)* hypotheses separately. These options have been missed by the extant duration-based tests in the literature. As in Christoffersen (1998), we define the hit-no-hit variable, I_t, as

$$I_t(\alpha) = \begin{cases} 1, & \text{if} \quad r_t < -\text{ISIVaR}_{t|t-1}(\alpha) \\ 0, & \text{else} \end{cases} \tag{8.25}$$

which imparts when a price change happens, if the observed return is lower or higher than the ex-ante level of ISIVaR. Following Christoffersen (1998) ISIVaR forecasts are valid if and only if the sequences of hit-no-hit variables $\{I_t\}$ fulfill the following two hypotheses:

1. The probability of an ex post irregularly spaced return exceeding the ISIVaR forecast must be equal to the coverage rate (α). This first hypothesis is termed *the unconditional coverage hypothesis*. Formally, we have

$$Pr\left[I_t(\alpha) - 1 = 0\right] = \alpha. \tag{8.26}$$

2. The second is the independence hypothesis which requires that ISIVaR violations observed at two different dates for the same coverage rate must be distributed independently. In other words, the correlation between hit-no-hit variables $I_t(\alpha)$ and $I_{t-k}(\alpha)$ at time t and $t-k$ $\forall k \neq 0$ is zero.

The simultaneous fulfillment of the both hypotheses *uc* and *ind* leads to say that the ISIVaR forecasts have a correct conditional coverage. We see that a correct ISIVaR forecast imposes that the hit-no-hit variable I_t has to follow a martingale process:

$$\mathbb{E}\left[I_t(\alpha) - \alpha | \Im_{t-1}\right] = 0. \tag{8.27}$$

This, in turn, implies that the sequence $\{I_t(\alpha)\}$ are *i.i.d.* Bernoulli distributed with a success probability equals to α. Indeed, the duration (D) between consecutive two violations is geometric distributed as it has been showed by Kupiec (1995). Formally, we define D as:

$$D_t = T_t - T_{t+1} \tag{8.28}$$

where, T_t is the time at which the t^{th} hit occurs. The probability density function of the duration D is given by

$$Geo(d; p) = p(1-p)^{d-1}, \quad d = 1, 2, 3 \ldots, \tag{8.29}$$

with p the success probability.

Despite the discrete nature of the problem, many papers adopt rather the continuous approximation of the geometric[2] distribution to implement appropriate test statistics. Since the work of Haas (2005) who to our knowledge, was the first to use the discrete Weibull in the backtesting procedures, it becomes clear that the use of the continuous approximation of the geometric distribution can have negative consequences for the power of the duration-based backtests in finite

[2] The geometric distribution has the nice property of the memoryless-ness.

samples. Following Haas (2005), Candelon et al. (2011) proposed a GMM duration-based test that is implemented using the geometric distribution. They provide orthonormal polynomials and moment conditions associated with the geometric distribution and its continuous analogous, i.e. the exponential distribution. In the following we briefly describe the orthonormal polynomials and moment conditions that will be used to test the performance of our models.

8.4.1. Orthonormal Polynomials and Moment Conditions

The classical orthonormal polynomials can be classified in two categories: Orthonormal polynomials associated with discrete variables or discrete orthonormal polynomials (Charlier, Meixner, Hahn, ...) and orthonormal polynomials associated with continuous variables or continuous orthonormal polynomials (Jacobi, Laguerre, Hermite, ...). Discrete orthonormal polynomials are on a linear lattice and the continuous ones on the real line. These orthonormal polynomials have the nice property that their expectation is equal to zero (cf. Appendix *A*.6). Bontemps and Meddahi (2012) derived moment conditions using the Hermite orthonormal polynomials associated to the Normal distribution to test for normality. In this chapter we employ orthonormal polynomials related to the geometric and exponential distribution. The first one can be viewed as a special case of the *Meixner* orthonormal polynomials related to a negative Binomial (*Pascal*) distribution and the second one is known as the Laguerre orthonormal polynomial. In the following we briefly present both orthonormal polynomials:

1. *Geometric distribution*: Let assume that the stochastic variable z, $\forall z \in \mathbb{N}^*$, is *geometric* distributed with a success probability θ. The associated orthonormal polynomials are defined as

$$M_{j+1}(z;\theta) = \frac{(1-\theta)(2j+1)+\theta(j-z+1)}{(j+1)\sqrt{1-\theta}} M_j(z;\theta) - \left(\frac{j}{j+1}\right) M_{j-1}(z;\theta) \tag{8.30}$$

   with $j = 1, \dots, p$, $M_{-1}(z;\theta) = 0$, and $M_0(z;\theta) = 1$.

   These orthonormal polynomials have been used in Candelon et al. (2011). In Appendix *A*.7 we explain how they can be obtained as a special case of the *Meixner* orthonormal polynomials related to a negative Binomial.

2. By assuming that the stochastic variable z, $\forall z \in \mathbb{R}^+$, is *exponential* distributed with a parameter θ, the associated Laguerre orthonormal polynomials satisfy the following recurrence relation

$$L_{j+1}(z;\theta) = \frac{1}{j+1}\Big[(2j+1-\theta z)L_j(z;\theta) - jL_{j-1}(z;\theta)\Big], \quad \forall j \geq 1 \tag{8.31}$$

   and

$$L_0(z;\theta) = 1, \quad L_1(z;\theta) = 1 - \theta z. \tag{8.32}$$

If the true distribution of Z is a *geometric* distribution with a success probability θ or *exponential* one with a parameter rate θ, then the moment conditions $E[M_j(z;\theta)] = 0$, or $E[L_j(z;\theta)] = 0$ are valid $\forall j > 1$ and can be tested, individually or jointly.

In the empirical application we apply the orthonormal polynomial associated with the geometric distribution and the Laguerre orthonormal polynomial to derive moment conditions and test the forecasting performance of the models. With the hit-no-hit variables $\{I_t(\alpha)\}_{t=1}^T$ we compute a sequence of N durations, $\{z_1, \ldots, z_N\}$, between ISIVaR violations. Under the null of correct conditional coverage (*cc*), the durations z_i, $i = 1, \ldots, N$, are *i.i.d.* and have a geometric (one can also assume that z_i are exponentially *i.i.d.*)[3] distribution. The null *cc* hypothesis can be formalized as

$$H_{0,cc} : \mathbb{E}[M_j(z_i;\alpha)] = 0 \quad j = 1, \ldots, k \tag{8.33}$$

or

$$H_{0,cc} : \mathbb{E}[L_j(z_i;\alpha)] = 0 \quad j = 1, \ldots, k \tag{8.34}$$

where k is the number of moment conditions considered and α is the coverage rate.

Interesting is that the moment conditions also permit to test the unconditional coverage and independence hypothesis separately. The correct unconditional coverage hypothesis implies that the probability of an ISIVaR violation occurring has to be equal to the coverage rate, α, and the independence hypothesis requires that ISIVaR violations happened at two different dates for the same coverage rate must be independently distributed. This offers the opportunity to check whether both hypotheses are simultaneously fulfilled or not.

The null *uc* hypothesis is defined as

$$H_{0,uc} : \mathbb{E}[M_1(z_i;\alpha)] = 0 \ \text{ or } \ \mathbb{E}[L_1(z_i;\alpha)] = 0, \tag{8.35}$$

and that of the *ind* hypothesis can be expressed as

$$H_{0,ind} : \mathbb{E}[M_j(z_i;\theta)] = 0 \ \text{ or } \ \mathbb{E}[L_j(z_i;\theta)] = 0, \tag{8.36}$$

which means that the duration between two consecutive violations is *geometric* or *exponential* distributed and if $\theta \neq \alpha$ the correct *uc* is not valid.

8.4.2. Empirical Test Method

Within the GMM framework the implementation of the test procedure is very easy due to the fact that the asymptotic variance-covariance matrix of the orthonormal polynomials are known (cf. Bontemps and Meddahi (2012) for Hermite orthonormal polynomials). Here under the assumption that the sequence of N durations $\{z_1, \ldots, z_N\}$ are *i.i.d.*, the moments associated with the orthonormal polynomials are asymptotically independent with unit variance (cf. Candelon et al.,

[3] This has been the common approach before the work of Haas (2005). Here we also consider this case and formulate the corresponding hypothesis.

2011). As result, the test statistic J_{cc} under the null hypothesis of correct conditional coverage related to the k first orthonormal polynomials is given by

$$J_{cc} = \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\delta(z_i;\alpha)\right)'\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\delta(z_i;\alpha)\right)\xrightarrow[N\to\infty]{d}\chi^2(k) \tag{8.37}$$

where $\delta(z_i;\alpha)$ is a vector of dimension $(k,1)$ whose entries are the orthonormal polynomials $M_j(z_i;\alpha)$, for $j = 1,\ldots,k$, and α is the coverage rate. Under the null hypothesis of the correct unconditional coverage (uc), the test statistic, J_{uc}, is obtained as a particular case of the J_{cc} when k is equal to one ($k = 1$).

At the finish, the test statistic for the independence (J_{ind}) hypothesis is given by

$$J_{ind} = \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\delta(z_i;\theta)\right)'\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\delta(z_i;\theta)\right)\xrightarrow[N\to\infty]{d}\chi^2(k) \tag{8.38}$$

where $\delta(z_i;\theta)$ is a vector of dimension $(k,1)$ whose entries are the orthonormal polynomials $M_j(z_i;\theta)$, for $j = 1,\ldots,k$, and θ is a success probability at which the orthonormal polynomials $M_j(z_i;\theta)$ are computed.

As stressed by Candelon et al. (2011) the true ISIVaR violations rate θ is unknown and may be different from the coverage rate α predefined by the risk manager. So, the test statistic for independence has to be computed by replacing θ by its consistent estimator $\hat{\theta}$. It is well known in the literature that such a substitution may change the asymptotic distribution of the test statistic. Nevertheless, Bontemps and Meddahi (2012) demonstrated that the asymptotic distribution does not change if the moments can be formalized as a projection onto the orthogonal of the score. This condition is satisfied by the orthonormal polynomials related with geometric or exponential distribution. As consequence, the test statistic becomes

$$J_{ind} = \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\delta(z_i;\hat{\theta})\right)'\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\delta(z_i;\hat{\theta})\right)\xrightarrow[N\to\infty]{d}\chi^2(k-1) \tag{8.39}$$

with the adjusted degrees of freedom $k - 1$.

8.5. Empirical Study

8.5.1. Data

For the empirical analysis we consider two stocks traded on the New Stock Exchange during the period from 1 to 30 July 2004 that includes 21 trading days: Boeing (BA), and Coca-Cola (KO). The data were extracted from the Trade and Quote (TAQ) database available at the NYSE. For each stock we define the irregularly spaced intraday returns as $r(t) = \ln(p_t) - \ln(p_{t-1})$, where p represents the bid-ask mid-point and the price duration ($x_t(\iota_p)$) is the minimal time interval needed to observe a change in the mid-price[4] (p) not less than a threshold (ι_p) that is set to \$0.0156.

8.5.2. Data Adjustment

We adjust the raw data as it has already been described above in sec. 5.5.1.1.

8.5.3. Results of Backtesting

We first estimate the parameters in both models using the total sample of observations. The results for the estimation of the models are reported in Table 8.2.

For each stock we first split the total sample of observations T in two sub-samples: An estimation sample S_e of size τ which contains adjusted durations x_t with $t = \tau_0 \ldots \tau$ ($\tau_0 = 1$) and a forecast sample S_f. To generate out-of-sample ISIVaR we use the algorithm described in sec. 8.3 and two different forecasting schemes, namely a fixed and a rolling forecast scheme with a fixed window width.

1. A fixed scheme: We estimate the parameters in each model specification (GGACD, MSMD) only once using data of the estimation sample of size τ and with the estimated parameters we generate all the forecasts for the out-of-sample period stretching over $\tau + 1$ to T.

2. A rolling scheme: To explain how we proceed by the rolling forecast scheme, let us define τ_0 as $\tau_0 = \tau - L$ where L is the length of the window used for estimating the model. Indeed, as τ is increased, new durations are included in estimation sample but older are removed. By each estimation we produce the ISIVaR forecast for the next observation of the forecast sample.

Figs. 8.3 and 8.4 depict the forecasts for intraday volatilities and expected durations for Boeing and Coca-Cola stocks. The expected duration is the inverse of the mean intensity and can be interpreted as the liquidity risk. Small mean intensities imply higher expected durations which in turn mean high liquidity risk. Clustering effects can be observed and give evidence that the model can capture the flow of information in the market.

In order to examine the performance of both models (MSMD and GGACD models) to produce accurate forecasts for ISIVaR we generate the hit-no-hit variables I_t by comparing the observed irregularly spaced returns to out-of-sample ISIVaR forecasts. *Fig.* 8.5 displays the plots for ISIVaRs obtained using rolling scheme in red and irregular spaced intra-day returns in blue for both stocks, Boeing and Coca-Cola. With the hit-no-hit variables I_t we produce the duration variables D_t and then apply the test statistic for validation purpose. Tables 8.4 through 8.6 present the results of the backtesting for both stocks at different confidence levels.

For both stocks (Boeing and Coca-Cola) the null hypothesis of a correct unconditional coverage, i.e. the proportion of hits is not statistically different from the coverage rate, α, cannot be rejected

[4] Mid-price is used to avoid biases caused by a bid-ask bounce.

at all levels $(5\%, 2.5\%, 1\%)$ in both models by considering both forecasting schemes. This gives evidence that both models perform well.

Concerning the conditional coverage test statistics (J_{cc}) we consider higher order orthonormal polynomials ($k = 2, 3$). In the following we present the results of the comparison between both models:

1. *At* 5% *confidence level with fixed forecast scheme*, the null hypothesis of a correct conditional coverage is rejected in both models using orthonormal polynomials associated with a geometric distribution. This is due the violation of the independence assumption that can be confirmed with the results in Table 8.6. However, the null hypothesis is accepted when instead of the geometric distribution one employs its continuous counterpart, i.e. the exponential distribution. The results obtained under the assumption that the duration variable D is exponential distributed show that although the consecutive hit-no-hit variables I_t are not independent, the null hypothesis of correct conditional coverage is accepted leading to wrong inferences. This confirms the findings of Haas (2005) who warned against using the continuous distribution that can deteriorate the sensitivity of the test.

2. *At* 5% *confidence level with rolling forecast scheme*, we obtain similar results in both models (GGACD and MSMD) for the Boeing stock (BA) as described above (cf. Tables 8.5 and 8.6). Interestingly, we see that using the rolling scheme the conditional coverage property is reached by the Coca-Cola stock (KO) in the MSMD model (cf. Table 8.5). This means that the null hypothesis of a correct conditional coverage is accepted for KO. It is obvious that the MSMD model outperforms the GGACD model.

3. *At* 2.5% *and* 1% *confidence level with fixed forecast scheme*, except for the Coca-Cola stock (KO) the null hypothesis is overall accepted. The results show that both models (GGACD and MSMD) perform well for both stocks (Boeing and Coca-Cola).

4. *At* 2.5% *and* 1% *confidence level with rolling forecast scheme*, the null hypothesis of a correct conditional coverage is accepted in both models for BA stock. For KO stock both models have some problems to provide accurate forecasts. However, we observe a superior performance of the MSMD model compared to that of the GGACD model.

In sum, both models perform well when using the fixed forecast scheme and exhibit similar performance. By employing the rolling scheme we find that the MSMD model dominates the GGACD model. We hope that these results can motivate practitioners and researchers to use the MSMD model.

8.6. Conclusion

In this chapter we have shown how the Markov-switching multifractal duration model can be used to compute the market risk at intraday level. We compare the forecasting performance of the

Table 8.1.: Information on the raw data

	Price durations	
	Boeing stock (BA)	Coca-Cola stock (KO)
Number of observations	8586	6906
Minimum value	1	1
Maximum value	800	2341
Mean value	52.726	65.400
Overdispersion	1.322	1.513
Skewness	3.456	6.134
Kurtosis	21.733	84.163

MSMD model to that of the standard GGACD model via the GMM duration-based test procedures developed by Candelon et al. (2011). The results of the Backtesting are quite satisfactory and prove that both models (GGACD and MSMD) are adequate to forecasting ISIVaR. Interestingly, we find the forecasting performance of the MSMD model when using the rolling forecasting scheme is superior to that of the GGACD model. We conclude that the MSMD model considerably dominates the GGACD model. Another important finding is that we can empirically confirm the fact that using the continuous distribution in the implementation of the backtesting methodology can lead to make wrong inferences. We recommend the market participants to use the MSMD model when forecasting the intraday market risk.

For future research, one can apply the MSMD model to calculate a market liquidity risk. It is clear that the average of times elapsed between intertrade durations quantifies the speed of trading activity and is a natural indicator of market liquidity (cf. Ghysels et al., 2004). They define the *Time at Risk* (TaR) at level α as $P_t(x_{t+1} > \text{TaR}) = \alpha$, where P_t is the conditional distribution at time t of the one step ahead duration x_{t+1}, and TaR(α) defines the minimal time without a trade that may happen with probability α.

Table 8.2.: Adjusted data

	Price durations	
	Boeing stock (BA)	Coca-Cola stock (KO)
Number of observations	8586	6906
Minimum value	0.006	0.006
Maximum value	21.784	41.497
Mean value	1.001	1.003
Overdispersion	1.269	1.610
Skewness	3.536	9.652
Kurtosis	29.794	189.739

Table 8.3.: Estimation of MSMD and GGACD models

MSMD				
	Boeing		Coca-Cola	
	Estimates	St. Error	Estimates	St. Error
m_0	1.210	0.009	1.227	0.011
$\bar{\lambda}$	1.584	0.073	1.702	0.082
b	1.407	0.183	1.659	0.195
γ	0.335	0.120	0.737	0.159
GGACD				
	Boeing		Coca-Cola	
	Estimates	St. Error	Estimates	St. Error
ω	0.092	0.019	0.107	0.027
β	0.155	0.019	0.131	0.019
δ	0.746	0.035	0.751	0.043
α	4.219	0.705	4.491	0.859
η	0.421	0.038	0.389	0.039

Note: In the MSMD model we set the intensity components to seven (k=7).

Table 8.4.: Fixed scheme backtesting Results

	k	GGACD		MSMD	
Boeing (BA)					
$\alpha = 5\%$					
Hits Freq.		0.144		0.159	
		Discrete	Cont.	Discrete	Cont.
J_{uc}	1	0.258	0.295	0.241	0.282
J_{cc}	2	**0.001**	0.280	**<0.001**	0.068
	3	**0.002**	0.063	**<0.001**	**0.003**
$\alpha = 2.5\%$					
Hits Freq.		0.092		0.095	
J_{uc}	1	0.278	0.301	0.270	0.294
J_{cc}	2	0.116	0.462	0.048	0.337
	3	0.206	0.644	0.085	0.533
$\alpha = 1\%$					
Hits Freq.		0.044		0.041	
J_{uc}	1	0.328	0.339	0.350	0.361
J_{cc}	2	0.594	0.534	0.451	0.363
	3	0.791	0.711	0.661	0.550
Coca-Cola (KO)					
$\alpha = 5\%$					
Hits Freq.		0.150		0.167	
J_{uc}	1	0.255	0.294	0.256	0.300
J_{cc}	2	**0.002**	0.291	**0.001**	0.395
	3	**0.002**	**0.031**	**0.003**	0.098
$\alpha = 2.5\%$					
Hits Freq.		0.086		0.095	
J_{uc}	1	0.220	0.241	0.246	0.270
J_{cc}	2	**<0.001**	**0.003**	**0.004**	0.081
	3	**<0.001**	**0.006**	**0.008**	0.168
$\alpha = 1\%$					
Hits Freq.		0.047		0.048	
J_{uc}	1	0.294	0.306	0.278	0.291
J_{cc}	2	0.473	0.567	0.293	0.430
	3	0.652	0.768	0.380	0.594

Note: The hit empirical frequency is the ratio of ISIVaR violations to the forecast sample size. The length of the forecast sample is 4294 for Boeing stock and 3453 for Coca-Cola. The data correspond to the period from 15 to 30 July, 2004. The Table contains the p-values for the backtesting using orthonormal polynomials related to the geometric and exponential distributions. The p-values in bold indicate that the null hypothesis of the correct conditional coverage is rejected.

Table 8.5.: Rolling scheme backtesting Results

	k	GGACD		MSMD	
		Boeing (BA)			
		$\alpha = 5\%$			
Hits Freq.		0.152		0.159	
		Discrete	Cont.	Discrete	Cont.
J_{uc}	1	0.248	0.288	0.238	0.279
J_{cc}	2	**<0.001**	0.142	**<0.001**	0.051
	3	**<0.001**	**0.012**	**<0.001**	**0.002**
		$\alpha = 2.5\%$			
Hits Freq.		0.088		0.086	
J_{uc}	1	0.281	0.303	0.289	0.311
J_{cc}	2	0.156	0.492	0.276	0.576
	3	0.197	0.699	0.355	0.776
		$\alpha = 1\%$			
Hits Freq.		0.040		0.038	
J_{uc}	1	0.348	0.358	0.356	0.366
J_{cc}	2	0.475	0.393	0.414	0.337
	3	0.679	0.561	0.622	0.511
		Coca-Cola (KO)			
		$\alpha = 5\%$			
Hits Freq.		0.147		0.151	
J_{uc}	1	0.272	0.310	0.277	0.317
J_{cc}	2	**0.037**	0.558	0.062	0.606
	3	**0.039**	0.067	0.129	0.227
		$\alpha = 2.5\%$			
Hits Freq.		0.083		0.088	
J_{uc}	1	0.240	0.260	0.248	0.270
J_{cc}	2	**0.003**	0.040	**0.011**	0.093
	3	**0.007**	0.093	**0.012**	0.185
		$\alpha = 1\%$			
Hits Freq.		0.044		0.042	
J_{uc}	1	0.239	0.249	0.241	0.251
J_{cc}	2	0.029	0.070	0.038	0.083
	3	**0.001**	**0.002**	0.013	0.016

Note: The hit empirical frequency is the ratio of ISIVaR violations to the forecast sample size. The length of the forecast sample is 4294 for Boeing stock and 3453 for Coca-Cola. The data correspond to the period from 15 to 30 July, 2004. The Table contains the p-values for the backtesting using orthonormal polynomials related to the geometric and exponential distributions. The p-values in bold indicate that the null hypothesis of the correct conditional coverage is rejected.

Table 8.6.: Independence tests for Boeing and Coca-Cola using both fixed and rolling schemes

	k	GGACD	MSMD	GGACD	MSMD
		Fixed scheme		Rolling scheme	
		Boeing (BA)		Boeing (BA)	
		5%		5%	
J_{ind}	2	**<0.001**	**<0.001**	**<0.001**	**<0.001**
	3	**0.001**	**<0.001**	**<0.001**	**<0.001**
		2.5%		2.5%	
J_{ind}	2	0.039	0.026	0.055	0.111
	3	0.104	0.038	0.097	0.199
		1%		1%	
J_{ind}	2	0.309	0.209	0.225	0.187
	3	0.596	0.455	0.473	0.417
		Coca-Cola (KO)		Coca-Cola (KO)	
		5%		5%	
J_{ind}	2	**<0.001**	**<0.001**	**0.011**	0.051
	3	**0.001**	**0.001**	**0.017**	0.062
		2.5%		2.5%	
J_{ind}	2	**<0.001**	**0.001**	**0.001**	**0.002**
	3	**<0.001**	**0.003**	**0.003**	**0.005**
		1%		1%	
J_{ind}	2	0.234	0.130	0.011	0.014
	3	0.457	0.228	**0.001**	0.010

Note: The hit empirical frequency is the ratio of ISIVaR violations to the forecast sample size. The length of the forecast sample is 4294 for Boeing stock and 3453 for Coca-Cola. The data correspond to the period from 15 to 30 July, 2004. The Table contains the p-values for the backtesting using the geometric and exponential distributions. The p-values in bold indicate that the null hypothesis of the correct independence is rejected.

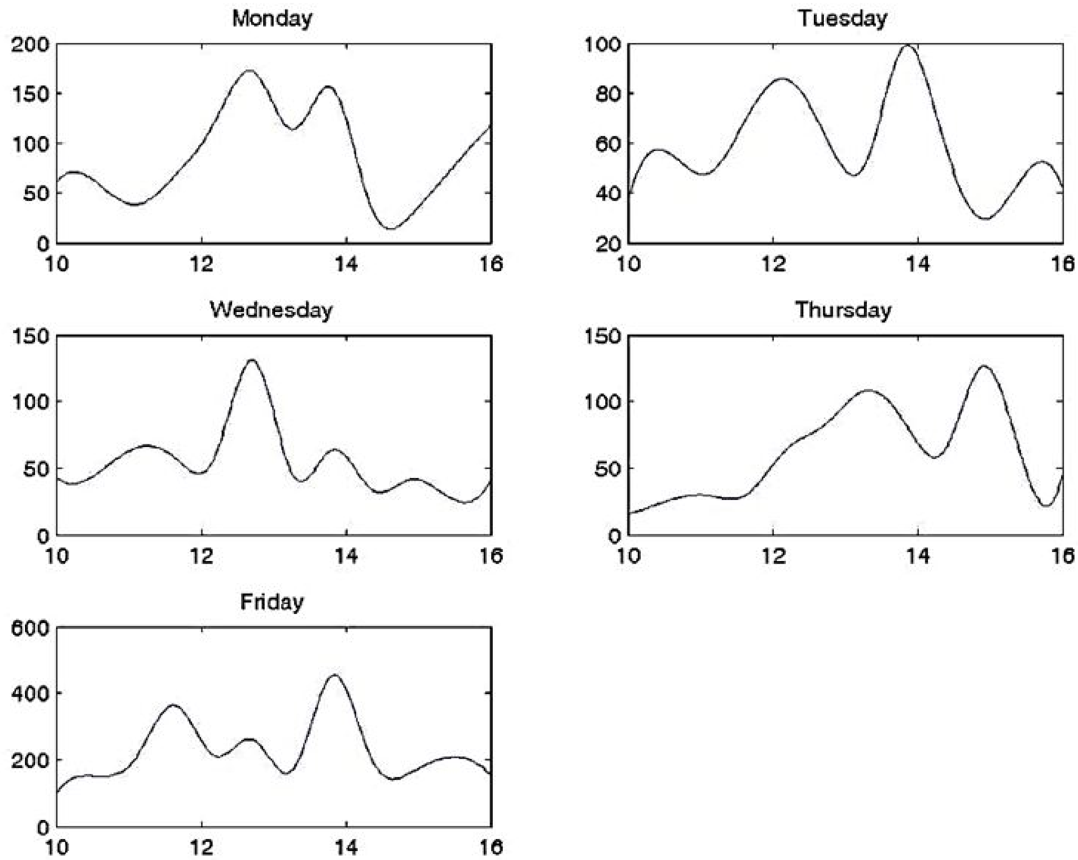

Figure 8.1.: Estimated time-of-the-day effects for Boeing

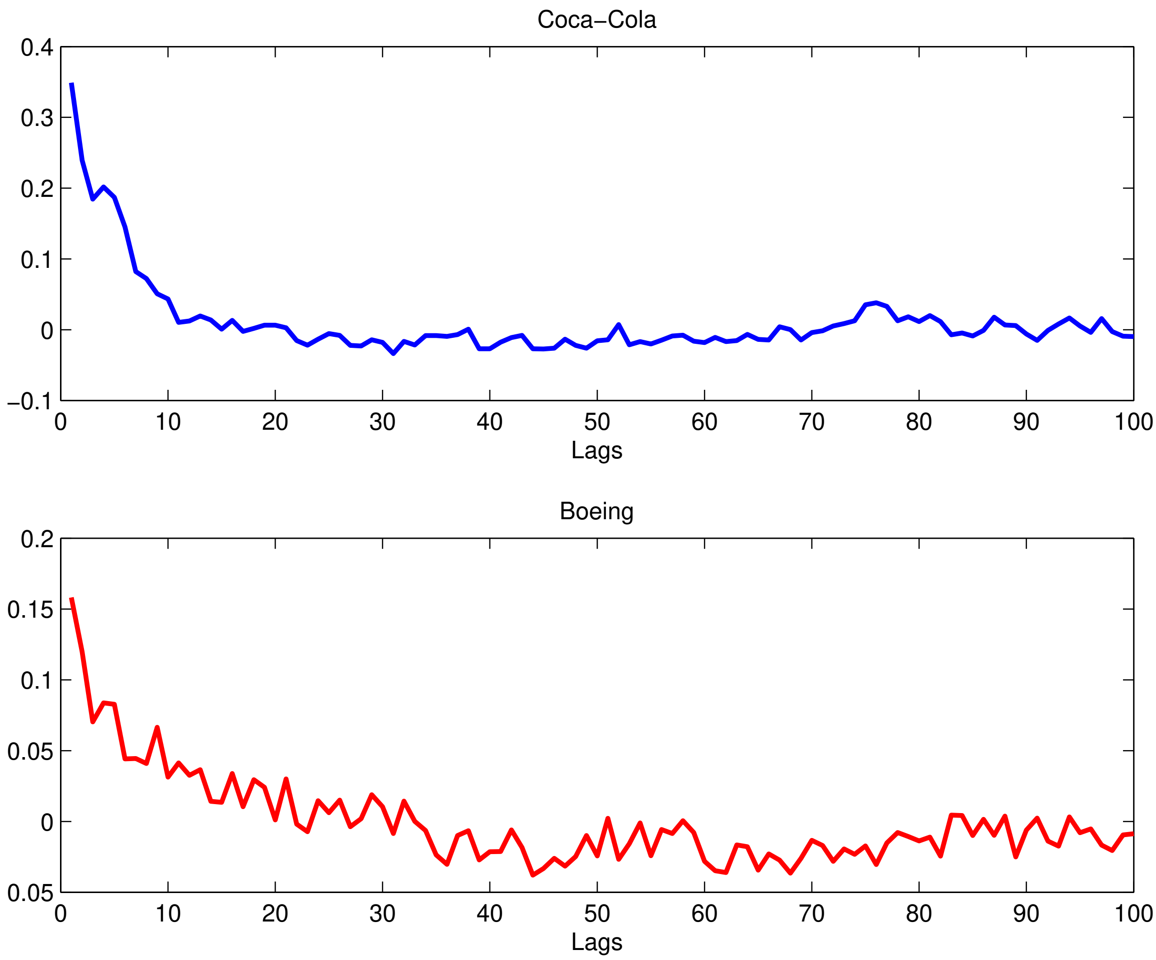

Figure 8.2.: Autocorrelation functions for both stocks (Boeing and Coca-Cola)

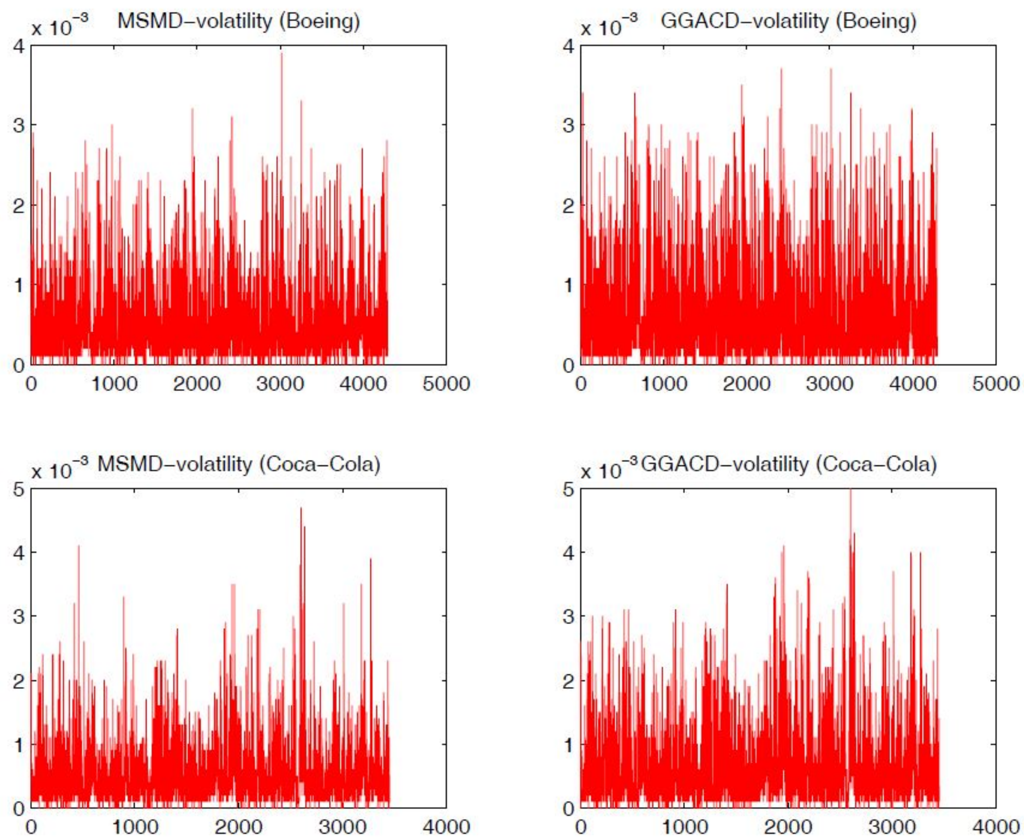

Figure 8.3.: Conditional volatility for price events (Boeing and Coca-Cola)

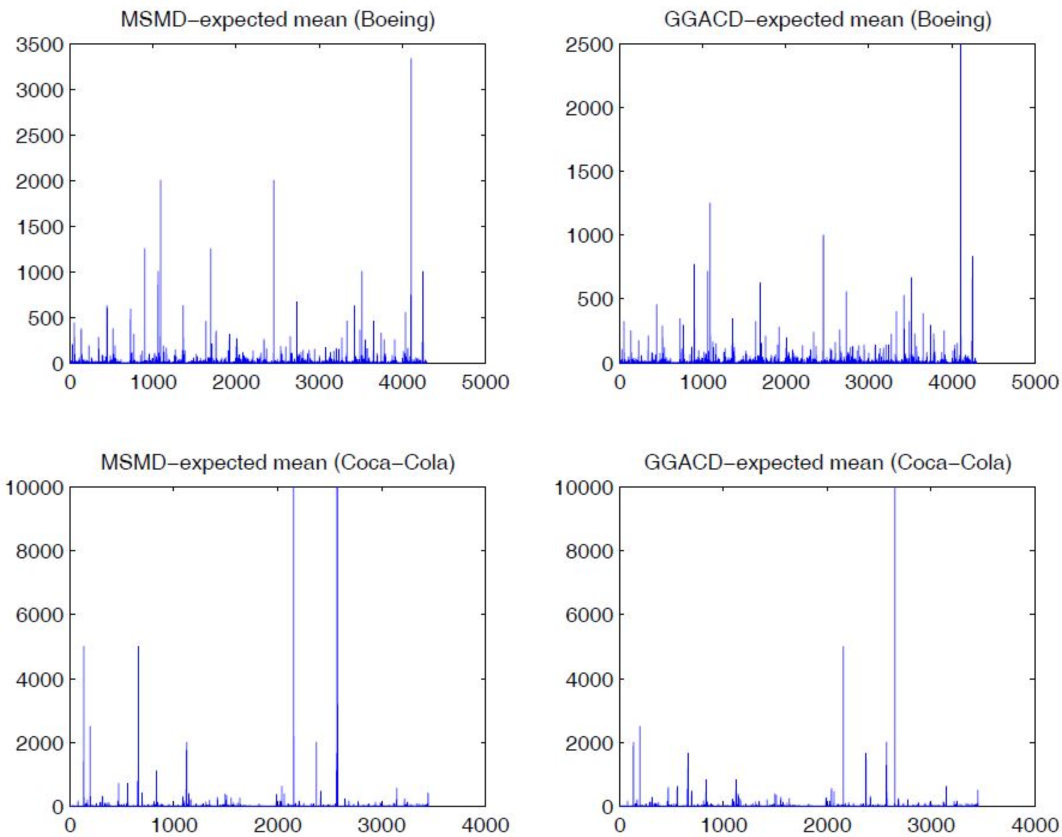

Figure 8.4.: Conditional expected mean for both stocks (Boeing and Coca-Cola)

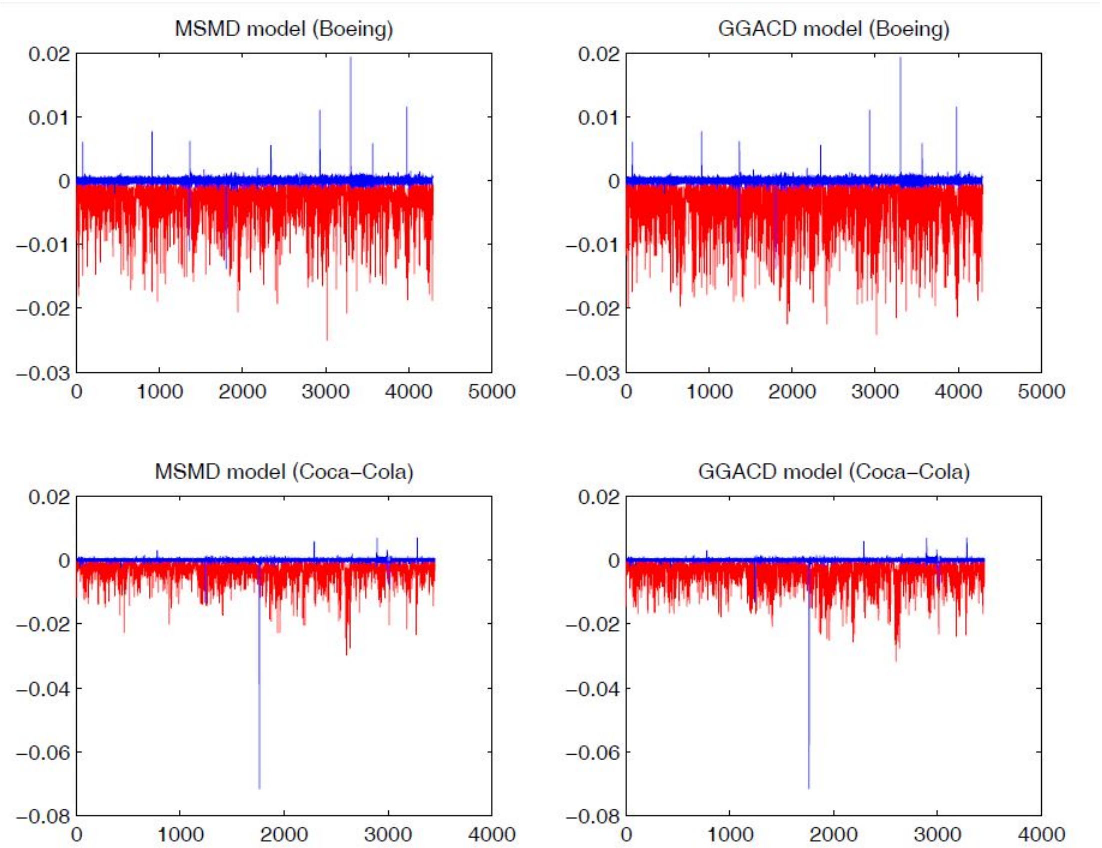

Figure 8.5.: Plot of ISIVaRs obtained using the rolling scheme in red and Boeing and Coca-Cola irregular spaced intra-day returns in blue.

Part IV.

General Conclusions And Outlooks

9. General Conclusions and Outlooks

9.1. General Conclusion

The development of multifractal measures and processes started in the earlier seventies with Mandelbrot (1974)'s works and achieved a lot of success in modeling of energy in turbulent dissipation in statistical physics. Although the ability of the processes to reproduce the most universal characteristics of asset returns (long memory, fat tails, multifractality and scaling behavior), they met with disapproval by econometricians and did not find successful applications in empirical finance. The reasons for this negligence are twofold: (i) The first generation models are non-causal nature and (ii) the extant estimation procedures are not familiar to economists.

Nowadays their acceptance and application in quantitative finance is due to the development of iterative time series models (Calvet and Fisher, 2001a, 2004a) and appropriate econometric tools that allow statistical inferences (Calvet and Fisher, 2004a; Calvet et al., 2006; Lux, 2008). A huge number of studies have already assessed their forecast performance and the empirical results indicate that multifractal models outperform the GARCH, MS-GARCH and FIGARCH models in terms of fitting and forecasting asset return volatility at long horizons.

In this thesis we have provided evidence of the capacity and robustness of the multifractal processes to model high frequency financial intertrade duration, bid-ask spreads, and oil price volatility. Our empirical results confirmed once again that the multifractal models outperform the traditional ACD models and the GARCH-type models in terms of fitting and forecasting financial data.

Chapter 4 has shown that the multifractal model can be used for forecasting oil price volatility. It seems impossible to demonstrate the superiority of the multifractal model over the GARCH-type models across six different loss functions that are used as criteria to evaluate our portfolio of models. However, based on the standard loss functions we observe that the multifractal model mostly cannot be outperformed by the GARCH-type models. Furthermore, we found that long memory GARCH models and the multifractal model can be combined to obtain accurate volatility forecasts.

We have found in chapters 5 and 6 that the Markov switching multifractal duration models can properly reproduce the long memory properties, the fat tails and the clustering effects observed in high frequency financial duration data (trade, price and volume durations). Using adequate statistical procedures we obtained the empirical results that witness the superiority of the MSMD models over the traditional ACD models with flexible distributions for innovations and over the exponential FIACD (EFIACD) model. We also found that flexible distributions (generalized gamma and Burr) did not enhance the forecast performance of the MSMD model as it is the case in the ACD models. This pointed out that the ability of the MSMD model to fit high frequency financial durations stems in large part from the multifractal processes.

In chapter 7 we extended the univariate MSMD model to a bivariate setting. The bivariate MSMD model offers the opportunity to model not only the duration processes but also the primary information available by financial transactions and to obtain more accurate forecasts. In addition, it also allows to analyze the interdependence between trade related variables. Our empirical results provided new insights into the bid-ask spreads of different stocks. Bid-ask spreads represent the crucial instruments for market makers in the financial markets and offer them the possibility to offset the costs they incur by trading with informed traders. We found that the bid-ask spreads of sector-specific or cross-sector stocks may be move together. This results will be of important interest for market makers and portfolio managers and will help them to better control their exposure to market risk.

In chapter 8 we have proved that the MSMD model is able to forecast accurately irregularly spaced intraday value-at-risk (ISIVaR). We have assessed the forecasting performance of the MSMD via a GMM duration-based test and compared it to that of generalized gamma ACD model. The empirical results that we obtained at different confidence levels (5%, 2.5% and 1%) are robust and clearly speak for the MSMD model.

In sum, the MF models are appropriate tools for the measurement and management of the market risk. The market risk is a vital input in portfolio optimization, asset allocation, derivative pricing and hedging, trading and conducting effective monetary policy. Therefore, we find that multifractal models can help financial institutions, portfolio managers, and regulators to exactly quantify market risk and avoid an over- or underestimation of the market risk. Because an overestimation of the market volatility would lead to an increase of the regulatory capital requirement and an underestimation can cause a collapse of the financial institution leading to a banking meltdown. Such a collapse can destabilize the whole financial sector and affect the real economy. We recommend policy makers to use the MF models when forecasting volatility because compared to other models, MF models are the only ones to our knowledge to provide accurate and robust forecasts over the long term.

9.2. Outlooks

There are many research questions that are very interesting and that we do not pursue in this thesis. These questions can be avenues for future research. The Markov switching multifractal duration models of Chen et al. (2013) or of Baruník et al. (2012) can be exploited to test some market microstructure hypotheses. Following Bauwens and Giot (2000) one can utilize the MSMD model to analyze the way the market maker revise their beliefs relative to bid-ask prices. Additional explicative variables relative to the trades can be introduced in the MSMD models in order to see whether these variables may improve the forecast performance of the models. One can also apply the models to study the concept of excess volume durations introduced by Hautsch (2003). Future research can extend the MSMD models to asymmetric ones using the asymmetric MSM process proposed by Leövey (2013) in his PhD thesis.

The bivariate MSMD model can also be used to analyze the interdependence between trading volumes and bid-ask spreads. The bivariate MSMD model can be generalized to multivariate setting. Multivariate MSMD model would provide possibilities to model duration and price processes simultaneously and better understand the price formation process.

Appendices

A. Supplement to the Thesis

A.1. Transformation of Random Variables

Given a random variable X whose probability density function $f_X(x)$ is known, the transformation of random variable method allows to derive the probability density function of another random variable Y related to X by the $Y = h(X)$ where the function $h(.)$

$$h : \mathbb{R} \to \mathbb{R}.$$

The inverse image of a set Ω is given by

$$h^{-1}(\Omega) = \{x \in \mathbb{R}; h(x) \in \Omega\}.$$

So, the mapping

$$\Omega \mapsto Pr\{h(X) \in \Omega\} = Pr\left\{X \in h^{-1}(\Omega)\right\}$$

fulfils the axioms of a probability and we have the following theorem (cf. Casella and Berger, 2002, chap. 2).

Theorem A.1.1 (Transformation Theorem) *Let X have pdf $f_X(x)$ and $Y = h(X)$, where h is a monotone function. Suppose that $f_X(x)$ is continuous on $h^{-1}(\Omega)$ and that $h^{-1}(Y)$ has a continuous derivative on Ω. Then, the pdf of Y is given by*

$$f_Y(y) = f_X\left[h^{-1}(y)\right]\left|\frac{d}{dy}h^{-1}(y)\right|, \quad y \in \Omega. \tag{A.1}$$

Proof In order to compute the probability density function of $Y = h(X)$ in terms of the probability density function of X, let start with the cumulative distribution function of Y and assume that h is an increasing function. We have

$$F_Y(y) = Pr(Y \leq y) = Pr[h(X) \leq y] = Pr\left[X \leq h^{-1}(y)\right] = F_X\left[h^{-1}(y)\right]. \tag{A.2}$$

By applying the chain rule to *eq.* (A.2), we obtain the pdf of Y as

$$f_Y(y) = \frac{d}{dy}F_X\left[h^{-1}(y)\right] = f_X\left[h^{-1}(y)\right]\frac{d}{dy}h^{-1}(y). \tag{A.3}$$

If h is a decreasing function on the range of X, then we rewrite the cumulative distribution function as

$$F_Y(y) = Pr(Y \leq y) = Pr[h(X) \leq y] = Pr\left[X \geq h^{-1}(y)\right] = 1 - F_X\left[h^{-1}(y)\right], \tag{A.4}$$

and the pdf of Y is given by

$$f_Y(y) = -\frac{d}{dy}F_X\left[h^{-1}(y)\right] = -f_X\left[h^{-1}(y)\right]\frac{d}{dy}h^{-1}(y). \tag{A.5}$$

A.2. Transformation of Random Variables in 2-D case

In the bivariate case one has to use the Jacobian matrix of the transformation (cf. also Casella and Berger, 2002, chap. 4).

Definition Suppose that x and y are two independent variables that are related to another two independent variables u and v by $x = h(u, v)$, $y = g(u, v)$. The Jacobian, $J(u, v)$ of x and y with respect to u and v is given by

$$J(u, v) = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \tag{A.6}$$

The joint distribution function of x and y is

$$f(x, y) = f(g(u, v), h(u, v))\,|J(u, v)|\,. \tag{A.7}$$

A.3. Uni- and Bivariate Lognormal Distribution Function

By assuming that a positive random variable ξ follows Lognormal distribution, then $\ln\xi$ follows Normal distribution with mean μ and standard deviation σ. The probability density function of the random variable ξ is given by

$$f(\xi) = \frac{1}{\xi\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln(\xi) - \mu}{\sigma}\right)^2\right], \ \xi > 0. \tag{A.8}$$

Its cumulative distribution function can be expressed as

$$F(\xi) = \Phi\left(\frac{\ln(\xi) - \mu}{\sigma}\right) \tag{A.9}$$

where Φ is the cumulative distribution function of the standard Normal distribution.

The joint distribution function of two positive correlated continuous lognormally distributed random variables ξ_1 and ξ_2 can be obtained via the Jacobian of the transformation. The joint distribution function is known in the literature as the bivariate Lognormal distribution function and can be formalized as

$$f(\xi_1, \xi_2) = \frac{1}{2\pi\xi_1\xi_2\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}Q\right], \tag{A.10}$$

where

$$Q = \left(\frac{\ln(\xi_1)-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{\ln(\xi_1)-\mu_1}{\sigma_1}\right)\left(\frac{\ln(\xi_2)-\mu_2}{\sigma_2}\right) + \left(\frac{\ln(\xi_2)-\mu_2}{\sigma_2}\right)^2, \tag{A.11}$$

and μ_1, σ_1 and μ_2, σ_2 are the means and standard deviations of $\ln \xi_1$ and $\ln \xi_2$, respectively (cf. Yerel and Konuk, 2009, for more detail on the uni- and bivariate Lognormal distribution functions).

A.4. The Joint Probability Density Function of the Bivariate MSMD Model

The bivariate MSMD model can be formalized as

$$\begin{cases} z_{1t} &= \dfrac{\xi_{1t}}{\lambda_{1t}} \\ z_{2t} &= \dfrac{\xi_{2t}}{\lambda_{2t}}. \end{cases} \tag{A.12}$$

The joint probability density function, $h(z_{1t}, z_{2t})$ of the bivariate MSMD model is given by

$$\begin{aligned} h(z_{1t}, z_{2t}) &= f(z_{1t}\lambda_{1t}, z_{2t}\lambda_{2t}) \begin{vmatrix} \frac{\partial \xi_{1t}}{\partial z_{1t}} & \frac{\partial \xi_{1t}}{\partial z_{2t}} \\ \frac{\partial \xi_{2t}}{\partial z_{1t}} & \frac{\partial \xi_{2t}}{\partial z_{2t}} \end{vmatrix} \\ &= f(z_{1t}\lambda_{1t}, z_{2t}\lambda_{2t}) \begin{vmatrix} \lambda_{1t} & 0 \\ 0 & \lambda_{2t} \end{vmatrix} \\ &= \lambda_{1t}\lambda_{2t} f(z_{1t}\lambda_{1t}, z_{2t}\lambda_{2t}), \end{aligned} \tag{A.13}$$

where f is the bivariate Lognormal distribution function defined in A.3

A.5. The Delta Method

Let denote ρ the correlation coefficient between $\ln \xi_{1,t}$ and $\ln \xi_{2,t}$, then the corresponding correlation coefficient ς between $\xi_{1,t}$ and $\xi_{2,t}$ is given by

$$\varsigma = \frac{\exp(\rho\sqrt{\sigma_{11}\sigma_{22}}) - 1}{\sqrt{[\exp(\sigma_{11}) - 1][\exp(\sigma_{22}) - 1]}} = h(\rho), \tag{A.14}$$

where $\varsigma \in (-1, 1)$, $h(\rho) = 0$ if $\rho = 0$, $|\varsigma| < \rho$, and $h(\rho) \neq -h(-\rho)$.

Using the two-steps estimation approach described in Sec. 7.5.1 we obtain the estimate $\hat{\rho}$ that is asymptotically Normal distributed. This means

$$\sqrt{T}\,(\hat{\rho} - \rho_0) \to^d N(0, \sigma^2_{\rho_0}) \tag{A.15}$$

where $\sigma^2_{\rho_0}$ is the asymptotic variance of the estimate $\hat{\rho}$ and ρ_0 is the "true" unknown parameter.

In fact, we are interested in the limiting distribution of ς and its asymptotic variance. The delta method allows us to derive the limiting distribution of ς (cf. Weisberg, 2001). Let $h : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function such that $h'(\rho_0) \neq 0$. We can rewrite h as Taylor series of the form:

$$h(X) \equiv h(\mu) + (X - \mu)h'(\mu), \tag{A.16}$$

where

$$h'(\mu) = \frac{\partial h(X)}{\partial X}\,|_{X=\mu} \tag{A.17}$$

and

$$\begin{aligned} \mathbb{V}ar\,[h(X)] &= \mathbb{V}ar\,(X - \mu)\,[h'(\mu)]^2 \\ &= \sigma^2\,[h'(\mu)]^2\,, \end{aligned} \tag{A.18}$$

where σ^2 is the variance of X.

From the *eq.* (A.16) we have

$$\sqrt{T}\,(h(\hat{\rho}) - h(\rho_0)) \to^d N(0, \sigma^2_{\rho_0}[h'(\rho_0)]^2). \tag{A.19}$$

The delta method estimator of the variance of the ς is obtained by using the estimators of ρ and $\sigma^2_{\rho_0}$ as above-described

$$\mathbb{V}ar[\varsigma] = \hat{\sigma}^2\,[h'(\hat{\rho})]^2\,. \tag{A.20}$$

A.6. Classical Discrete Orthogonal Polynomials

As defined in Arvesú et al. (2003) orthogonal polynomials $\{p_n : n = 0, 1, 2, \dots\}$ in account with a positive measure μ on the real line fulfill the conditions

$$\int p_n(x)x^j d\mu(x) = 0, \quad j = 0, 1, \dots, n - 1. \tag{A.21}$$

In the case of discrete orthogonal polynomials the corresponding discrete measure μ can be expressed as a linear combination of Dirac measures on the $N + 1$ points $x_0, \dots, x_N$. Formally, we have

$$\mu = \sum_{k=0}^{N} \rho_k \delta_{x_k}, \quad \rho_k > 0, x_k \in \mathbb{R} \text{ and } N \in \mathbb{N} \cup \{+\infty\}. \tag{A.22}$$

The orthogonality conditions of a discrete orthogonal polynomial p_n on the set $\{x_k = k : k = 0, 1, \ldots, N\}$ are formalized as (cf. Arvesú et al., 2003)

$$\sum_{k=0}^{N} p_n(k)(-k)_j \rho_k = 0 \quad j = 0, 1, \ldots, n-1, \tag{A.23}$$

with $(a)_j = a(a+1)\ldots(a+j-1)$ if $j > 0$ and $(a)_0 = 1$ (cf. the Pochhammer symbol).

Meixner Orthogonal Polynomials:

The monic[1] discrete orthogonal polynomials associated with a negative Binomial (NB) distribution (Pascal distribution) on $\mathbb{N}$ satisfy the following orthogonality conditions

$$\sum_{k=0}^{+\infty} M_n(k; \beta, c)(-k)_j \frac{(\beta)_k}{k!} c^k = 0, \quad j = 0, 1, \ldots, n-1, \tag{A.24}$$

with $\beta > 0$ and $0 < c < 1$.

$\beta > 0$ and $0 < c < 1$.

As demonstrated in Filipuk and Van Assche (2013) the Meixner orthonormal polynomial M_n satisfies the following recurrence relation

$$xM_n(x; \beta, c) = a_{n+1} M_{n+1}(x; \beta, c) + b_n M_n(x; \beta, c) + a_n M_{n-1}(x; \beta, c), \tag{A.25}$$

with $n \geq 0$, where the recurrence coefficients are given by

$$a_n^2 = \frac{cn(\beta + n - 1)}{(1-c)^2}, \quad b_n = \frac{n + (\beta + n)c}{(1-c)}, \tag{A.26}$$

and initials conditions are $M_0 = 1$ and $M_{-1} = 0$.

We can rewrite *eq.* (A.25) as

$$M_{n+1}(x; \beta, c) = \frac{x - b_n}{a_{n+1}} M_n(x; \beta, c) - \frac{a_n}{a_{n+1}} M_{n-1}(x; \beta, c), \tag{A.27}$$

or equivalently,

$$M_{n+1}(x; \beta, c) = h(x; \beta, c) M_n(x; \beta, c) - g(x; \beta, c) M_{n-1}(x; \beta, c), \tag{A.28}$$

where

$$g(x; \beta, c) = \frac{a_n}{a_{n+1}}, \tag{A.29}$$

[1] In algebra, a univariate polynomial is said to be monic, if its leading coefficient is equal to 1.

and

$$h(x;\beta,c) = \frac{x - b_n}{a_{n+1}}. \tag{A.30}$$

A.7. A Special Case of Meixner Orthonormal Polynomials

The geometric distribution is a special case of the negative Binomial distribution. By setting $\beta = 1$ and $c = 1 - \alpha$ we obtain the geometric distribution (Geo($d;\alpha$)=NB($x;1,1-\alpha$)). The *eq.* (A.28) becomes

$$M_{n+1}(d;\alpha) = k(d;\alpha)M_n(d;\alpha) - q(d;\alpha)M_{n-1}(d;\alpha), \tag{A.31}$$

where

$$q(d;\alpha) = g(x;\beta,c) = \frac{n}{n+1}, \tag{A.32}$$

and

$$k(d;\alpha) = -h(x;1,1-\alpha) = \frac{(1-\alpha)(1+2n) + \alpha(n-x+1)}{(1+n)\sqrt{1-\alpha}}, \tag{A.33}$$

with $x = d - 1$ and initials conditions do not change: $M_0 = 1$ and $M_{-1} = 0$.

Bibliography

Admati, A. R. and P. Pfleiderer (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies 1*, 3–40.

Aiolfi, M. and A. Timmermann (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics 135*, 31–53.

Akaike, H. (1974). A new look at the statistical identification model. *IEEE Trans. Automat. Control 19*, 716–723.

Allen, D., F. Chan, M. McAleer, and S. Peiris (2008). Finite sample properties of the QMLE for the Log-ACD model: Application to australian stocks. *Journal of Econometrics 147*, 163–185.

Allen, D., Z. Lazarov, M. McAleer, and S. Peiris (2009). Comparison of alternative ACD models via density and interval forecasts: Evidence from the Australian stock market. *Mathematics and Computers in Similation 79*, 2535–2555.

Amihud, Y. and H. Mendelson (1980). Dealer ship market market-making with inventory. *Journal of Financial Economics 8*, 31–53.

Andersen, T., T. Bollerslev, P. Christoffersen, and F. Diebold (2006). Volatility and correlation forecasting. *Handbook of Economic Forecasting 1*, 777–878.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2001). The distribution of realized stock return volatility. *Journal of Financial Econometrics 63*, 43–76.

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Ané, T. and H. Geman (2000). Order flows, transaction clock, and Normality of asset returns. *Journal of Finance 55*, 2259–2284.

Arbeiter, M. and N. Patzschke (1996). Random self-similar multifractals. *Mathematische Nachrichten 181*, 5–42.

Arneodo, A., J.-F. Muzy, and D. Sornette (1998). "direct" causal cascade in the stock market. *European Physical Journal B 2*, 277–282.

Arshed, M., M. T. Rasool, and M. I. Ahmad (2003). Anderson Darling and modified Anderson Darling tests for generalized Pareto distribution. *Pakistan Journal of Applied Sciences 3*, 85–88.

Arvesú, J., J. Coussement, and W. Van Assche (2003). Some discrete multiple orthogonal polynomials. *Journal of Computational and Applied Mathematics 153*, 19–45.

Ausloos, M. and K. Ivanova (2000). Introducing false EUR and false EUR exchange rates. *Physica A: Statistical Mechanics and its Applications 286*, 353–366.

Ausloos, M., N. Vandewalle, P. Boveroux, A. Minguet, and K. Ivanova (1999). Applications of statistical physics to economic and financial topics. *Physica A: Statistical Mechanics and its Applications 274*, 229–240.

Bachelier, L. (1900). Théorie de la spéculation. *Annales de l'Ecole Normale Supérieure 3*.

Backus, D. K. and M. J. Crucini (2000). Oil prices and the terms of trade. *Journal of International Economics 50*, 185–213.

Bacry, E., J. Delour, and J.-F. Muzy (2000). A multivariate multifractal model for return fluctuations.

Bacry, E., J. Delour, and J.-F. Muzy (2001). Multifractal random walk. *Physical Review E 64*, 026103–026106.

Bacry, E., L. Duvernet, and J.-F. Muzy (2012). Continuous-time skewed multifractal processes as a model for financial returns. *Journal of Applied Probability 49*, 482–502.

Bacry, E., A. Kozhemyak, and J.-F. Muzy (2008). Continuous cascade model for asset returns. *Journal of Economic Dynamics and Control 32*, 156–199.

Bacry, E., A. Kozhemyak, and J.-F. Muzy (2013). Lognormal continuous cascades: Aggregation properties and estimation. *Quantitative Finance 13*, 795–818.

Bacry, E. and J.-F. Muzy (2003). Log-infinitely divisible multifractal processes. *Communications in Mathematical Physics 236*, 449–475.

Bacry, E. and J.-F. Muzy (2010). Multifractal models for asset prices. in Encyclopedia of quantitative finance, Wiley.

Bagehot, W. (1971). The only game in town. *Financial Analysts Journal 22*, 12–17.

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

Ball, C. A. and W. N. Torous (1984). The maximum likelihood estimation of security price volatility: Theory, evidence and application to option pricing. *Jounal of Business 57*, 97–113.

Barber, B. M. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies 21*, 785–818.

Barndorff-Nielsen, O. E. and K. Prause (2001). Apparent scaling. *Finance and Stochastics 5*, 103–113.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society Series B 64*, 253–280.

Barral, J. (1999). Moments, continuité, et analyse multifractale des martingales de mandelbrot. *Probability Theory Related Fields 113*, 535–569.

Barral, J. and B. B. Mandelbrot (2001). Multifractal products of cylindrical pulses. Cowles Foundation Discussion Paper 1287, Cowles Foundation for Research in Economics, Yale University.

Barral, J. and B. B. Mandelbrot (2002). Multifractal products of cylindrical pulses. *Probability Theory Related Fields 124*, 409–430.

Barton, C. and P. La Pointe (1995). *Fractals in Petroleum Geology and Earth Processes*. Springer.

Baruník, J., N. Shenai, and F. Žikeš (2012). Modeling and forecasting persistent financial durations. Working Paper.

Bauwens, L. (2006). Econometric analysis of intra-daily trading activity on the Tokyo Stock Exchange. *Monetary and Economic Studies*.

Bauwens, L. and F. Galli (2009). Efficient importance sampling for ML estimation of SCD models. *Computational Statistics and Data Analysis 53*, 1974–1992.

Bauwens, L., F. Galli, and P. Giot (2008). The moments of Log-ACD models. *Quantitative and Qualitative Analysis in Social Sciences*, 1–28.

Bauwens, L. and P. Giot (2000). The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks. *Annals d'Economie et de Statistique 60*, 117–149.

Bauwens, L. and P. Giot (2001). *Econometric Modelling of Stock Market Intraday Activity*. Dordrecht: Kluwer Academic Publishers.

Bauwens, L. and P. Giot (2003). Asymmetric ACD models: Introducing price information in ACD models. *Empirical Economics 28*, 709–731.

Bauwens, L., P. Giot, J. Grammig, and D. Veredas (2004). Comparison of financial duration models via density forecasts. *International Journal of Forecasting 20*, 589–609.

Bauwens, L. and D. Veredas (2004). The stochastic conditional duration model: A latent variable model for the analysis of financial durations. *Journal of Econometrics 119*, 381–412.

Behr, A. and U. Pötter (2009). Alternatives to the normal model of stock returns: Gaussian mixture, generalised logF and generalised hyperbolic models. *Annals of Finance 5*, 49–68.

Berkowitz, J. (2001). Testing density forecasts, with application to risk management. *Journal of Business and Economic Statistics 12*, 465–474.

Berkowitz, J., P. Christoffersen, and D. Pelletier (2011). Evaluating Value-at-Risk models with desk-level data. *Journal of Management Science 57*, 2213–2227.

Bhatti, C. R. (2010). The Birnbaum-Saunders autoregressive conditional duration model. *Mathematics and Computers in Simulation 80*, 2062–2078.

Bierens, H. J. (1990). A consistent conditional moment test of functional form. *Econometrica 58*, 1443–1458.

Black, F. (1976). Studies of stock market volatility changes. *1976 Proceedings of the American Statistical Association, Business and Economic Section*, 177–181.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–659.

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. *Review of Economics and Statistics 72*, 498–505.

Bollerslev, T., R. F. Engle, and D. Nelson (1994). *Handbook of Econometrics*, Volume 4, Chapter ARCH models, pp. 2961–3038. Elsevier Science BV, Amsterdam.

Bontemps, C. and N. Meddahi (2005). Testing normality: A GMM approach. *Journal of Econometrics 124*, 149–186.

Bontemps, C. and N. Meddahi (2012). Testing distributional assumptions: A GMM approach. *Journal of Applied Econometrics 27*, 978–1012.

Bouchaud, J. P. and M. Potters (2004). *Theory of Financial Risks and Derivative Pricing*. Cambridge University Press.

Breidt, F. J., N. Crato, and P. de Lima (1998). On the detection and estimation of long memory in stochastic volatility. *Journal of Econometrics 83*, 325–348.

Breymann, W., S. Ghashghaie, and P. Talkner (2000). A stochastic cascade model for FX dynamics. *International Journal of Theoretical and Applied Finance 3*, 357–360.

Brockwell, P. and R. Dahlhaus (2004). Generalized Levinson-Durbin and Burg algorithms. *Journal of Econometrics 118*, 129–144.

Brockwell, P. and R. Davis (1991). *Time Series: Theory and Methods*. Berlin: Springer.

Burr, W. (1942). Cumulative frequency functions. *Annals of Mathematical Statistics 13*, 215–232.

Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business 12*, 309–316.

Calvet, L., M. Fearnley, A. Fisher, and M. Leippold (2013). What's beneath the surface? option pricing with multifrequency latent states. Working Paper.

Calvet, L. and A. Fisher (2001a). Forecasting multifractal volatility. *Journal of Econometrics 105*, 27–58.

Calvet, L. and A. Fisher (2001b). Forecasting multifractal volatility. *Journal of Econometrics 105*, 27–58.

Calvet, L. and A. Fisher (2002). Multifractality in asset returns: Theory and evidence. *Review of Economics and Statistics 83*, 381–406.

Calvet, L. and A. Fisher (2004a). Regime-switching and the estimation of multifractal processes. *Journal of Financial Econometrics 2*, 44–83.

Calvet, L. and A. Fisher (2004b). Regime-switching and the estimation of multifractal processes. *Journal of Financial Econometrics 2*, 44–83.

Calvet, L., A. Fisher, and B. B. Mandelbrot (1997). Large deviations and the distribution of price changes. Cowles Foundation Discussion Papers 1165, Cowles Foundation for Research in Economics, Yale University.

Calvet, L., A. Fisher, and S. Thompson (2006). Volatility comovement: A multifrequency approach. *Journal of Econometrics 31*, 179–215.

Calvet, L., A. Fisher, and L. Wu (2013). Staying on top of the curve: A cascade model of term structure dynamics. Working Paper.

Candelon, B., G. Colletaz, C. Hurlin, and S. Tokpavi (2011). Backtesting value at risk: A GMM duration-based test. *Journal of Financial Econometrics 9*, 314–343.

Carius, S. and G. Ingelman (1990). The log-normal distribution for cascade multiplicities in hadron collisions. *Physics Letters B 252*, 647–652.

Casella, G. and R. L. Berger (2002). *Statistical Inference* (2nd ed.). Cengage Learning.

Castaing, B., Y. Gagne, and E. J. Hopfinger (1990). Velocity probability density functions of high Reynolds number turbulence. *Physica D 46*, 177–200.

Chatfield, C. (1993). Calculating interval forecasts. *Journal of Business and Economics Statistics 11*, 121–135.

Chen, F., F. Diebold, and F. Schorfheide (2013). A Markov switching multifractal intertrade duration model, with application to U.S. equities. *Journal of Econometrics 177*, 320–342.

Chen, W. W. and R. S. Deo (2004). A generalized portmanteau goodness-of-fit test for time series models. *Econometric Theory 20*, 382–416.

Chen, Y. and C. Hsieh (2010). Generalized moment tests for autoregressive conditional duration models. *Journal of Financial Econometrics 8*, 345–391.

Chen, Z., P. C. Ivanov, K. Hu, and H. E. Stanley (2002). Effect of nonstationarities on detrended fluctuation analysis. *Physical Review E 65*, 041107.

Cheong, C. W. (2009). Modeling and forecasting crude oil markets using ARCH-type models. *Energy Policy 37*, 2346–2355.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review 39*, 841–862.

Christoffersen, P. F. and D. Pelletier (2004). Backtesting value-at-risk: A duration based approach. *Journal of Financial Econometrics 2*, 84–108.

Chung, K. H., B. F. V. Ness, and R. A. Ness (1999). Limit orders and the bid-ask spread. *Journal of Financial Economics 53*, 255–287.

Cioczek-Georges, R. and B. B. Mandelbrot (1995). A class of micropulses and antipersistent fractional Brownian motion. *Stochastic Processes and their Applications 60*, 1–18.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica 41*, 135–156.

Colletaz, G., C. Hurlin, and S. Tokpavi (2007). Irregularly spaced intraday value at risk (ISIVaR) models: Forecasting and predictive abilities. Unpublished Paper, University of Orléans.

Conover, W. J. (1999). *Practical nonparametric statistics* (Third ed.). John Wiley and Sons, Inc. New York.

Conrad, C. (2010). Non-negativity conditions for the hyperbolic GARCH model. *Journal of Econometrics 157*, 441–457.

Copeland, T. and D. Galai (1983). Information effects and the bid-ask spread. *Journal of Finance 38*, 1457–1469.

Corradi, V. (2000). Reconsidering the continuous time limit of the GARCH(1,1) process. *Journal of Econometrics 96*, 145–153.

Corwin, S. A. and J. F. Coughenour (2008). Limited attention and the allocation of effort in securities trading. *Journal of Finance 63*, 3031–3067.

Cowles, A. (1933). Can stock market forecasters forecast? *Econometrica 1*, 309–324.

Cowles, A. and H. Jones (1937). Some a posteriori probabilities in stock market action. *Econometrica 5*, 280–294.

Crato, N. and P. J. de Lima (1994). Long-range dependence in the conditional variance of stock returns. *Economics Letters 45*, 281–285.

Dacorogna, M. M., R. Gençay, U. A. Müller, R. B. Olsen, and O. V. Pictet (2001). *An Introduction to High Frequency Finance*. Academic Press, San Diego, CA.

Dacorogna, M. M., U. A. Müller, R. J. Nagler, R. B. Olsen, and O. V. Pictet (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance 12*, 413–438.

Dacorogna, M. M., U. A. Müller, R. B. Olsen, and O. V. Pictet (1998). *Nonlinear Modelling of High Frequency Financial Time Series*, Chapter Modelling short-term volatility with GARCH and HARCH models, pp. 161–176. John Wiley & Sons, Inc.

Darolles, S., C. Gouriéroux, and G. L. Fol (2000). Intraday transaction price dynamics. *Annales d'Economie et de Statistique 60*, 207–238.

Davidian, M. and J. Carroll (1987). Variance function estimation. *Journal of American Statistics Association 82*, 1079–1091.

Davidson, J. (2004). Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business and Economic Statistics 22*, 16–29.

Davis, S. J. and J. Haltiwanger (2001). Sectoral job creation and destruction responses to oil price changes. *Journal of Monetary Economics 48*, 465–512.

De Luca, G. and G. M. Gallo (2004). Mixture process for financial intradaily durations. *Studies in Nonlinear Dynamics and Econometrics 8*, 1–18.

De Luca, G. and P. Zuccolotto (2003). Finite and infinite mixtures for financial durations. *Metron-International Journal of Statistics 61*, 431–456.

De Luca, G. and P. Zuccolotto (2006). Regime-switching Pareto distributions for ACD models. *Computational Statistics and Data Analysis 51*, 2179–2191.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society 39*, 1–38.

Deo, R., M. Hsieh, and C. Hurvich (2010). Long memory in intertrade durations, counts and realized volatility of NYSE stocks. *Journal of Statistical Planning and Inference 140*, 3715–3733.

Deo, R., C. Hurvich, and Y. Lu (2006). Forecasting realized volatility using a long memory stochastic volatility model: Estimation, prediction and seasonal adjustment. *Journal of Econometrics 131*, 29–58.

Diamond, D. W. and R. E. Verrecchia (1987). Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Econometrics 18*, 277–311.

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*, 427–431.

Diebold, F., T. Gunther, and A. Tay (1998). Evaluating density forecasts with application to financial risk management. *International Economic Review 39*, 863–883.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*, 253–263.

Ding, Z., C. Granger, and R. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance 1*, 83–106.

Dionne, G., P. Duchesne, and M. Pacurar (2009). Intraday value at risk (IVaR) using tick-by-tick data with application to the Toronto Stock Exchange. *Journal of Empirical Finance 16*, 777–792.

Dowd, K. (2004). A modified Berkowitz backtest. *Risk 17*, 86–87.

Drost, F. C. and B. J. Werker (1996). Closing the GARCH gap: Continuous time GARCH modeling. *Journal of Econometrics 74*, 31–57.

Drost, F. C. and B. J. Werker (2004). Semiparametric duration models. *Journal of Business and Economic Statistics 25*, 786–818.

Duchesne, P. and M. Pacurar (2008). Evaluating financial time series models for irregularly spaced data: A spectral density approach. *Computers and Operations Research, Special Issue: Applications of OR in Finance 35*, 130–155.

Duchon, J., R. Robert, and V. Vargas (2012). Forecasting volatility with multifractal random walk model. *Mathematical Finance 22*, 83–108.

Dufour, A. and R. Engle (2000a). The ACD model: Predictability of the time between consecutive trades. Unpublished Paper ISMA Center, University of Reading.

Dufour, A. and R. Engle (2000b). Time and price impact of a trade. *Journal of Finance 55*, 2467–2498.

Durbin, J. and S. J. Koopman (1997). Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika 84*, 669–684.

Easley, D. and M. O'Hara (1987). Price, trade, size and information in securities markets. *Journal of Finance 47*, 557–605.

Easley, D. and M. O'Hara (1992). Time and the process of security price adjusment. *Journal of Finance 19*, 69–90.

Eberlein, E. and U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli 1*, 281–299.

Ederington, L. H. and W. Guan (2005). Forecasting volatility. *Journal of Futures Markets 25*, 465–490.

Eisler, Z. and J. Kertész (2004). Multifractal model of asset returns with leverage effect. *Physica A 343*, 603–622.

Eisler, Z. and J. Kertész (2006). Size matters: Some stylized facts of the stock market revisited. *European Physical Journal B 51*, 145–154.

Elder, J. and A. Serletis (2010). Oil price uncertainty. *Journal of Money, Credit and Banking 42*, 1137–1159.

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

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

Engle, R. and T. Bollerslev (1986a). Modelling the persistence of conditional variances. *Econometric Reviews 27*, 1–50.

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

Engle, R. and J. Lange (2001). Predicting VNET: A model of the dynamics of the market depth. *Journal of Financial Markets 4*, 113–142.

Engle, R. and S. Manganelli (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics 22*, 367–381.

Engle, R. and J. Russell (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica 66*, 1127–1162.

Evertsz, C. J. and B. B. Mandelbrot (1992). *Chaos and Fractals: New Frontiers of Science*, Chapter Multifractal Measures. Berlin: Springer.

Falconer, K. J. (1994). The multifractal spectrum of statistically self-similar measures. *Journal of Theoretical Probability 7*, 681–702.

Fama, E. F. (1963). Mandelbrot and the stable Paretian hypothesis. *Journal of Business 36*, 420–429.

Fama, E. F. (1965). The behavior of stock market prices. *Journal of Business 38*, 34–105.

Fama, E. F. (1970). Efficient capital market: A review of theory and empirical work. *Journal of Finance 25*, 383–417.

Fama, E. F. and R. Roll (1971). Parameter estimates for symmetric stable distributions. *Journal of the American Association 66*, 331–338.

Farmer, D., L. Gillemot, F. Lillo, S. Mike, and A. Sen (2004). What really causes large price changes? *Quantitative Finance 4*, 383–397.

Feng, D., G. J. Jiang, and P. Song (2004). Stochastic conditional duration models with "leverage effect" for financial transaction data. *Journal of Financial Economics 119*, 413–433.

Ferderer, J. (1996). Oil price volatility and the macroeconomy. *Journal of Macroeconomics 18*, 1–26.

Fergussen, K. and E. Platen (2006). On the distributional characterization of daily log-returns of a world stock index. *Applied Mathematical Finance 13*, 19–38.

Fernandes, M. and J. Grammig (2005). Nonparametric specification tests for conditional duration models. *Journal of Econometrics 127*, 35–68.

Fernandes, M. and J. Grammig (2006). A family of autoregressive conditional duration models. *Journal of Econometrics 130*, 1–23.

Filimonov, V. and D. Sornette (2011). Self-excited multifractal dynamics. *Europhysics Letters 94*, 46003.

Filipuk, G. and W. Van Assche (2013). Recurrence coefficients of generalized Charlier polynomials and the fifth Painlevé equation. *Proceedings of the American Mathematical Society 141*, 551–562.

Fillol, J. (2003). Multifractality: Theory and evidence: An application to the French stock market. *Economics Bulletin 3*, 1–12.

Finn, M. G. (2000). Perfect competition and the effects of energy price increases on economic activity. *Journal of Money, Credit, and Banking 32*, 400–416.

Fisher, A., L. Calvet, and B. B. Mandelbrot (1997). Multifractality of Deutschemark/US Dollar exchange rates. Cowles Foundation Discussion Papers 1166, Cowles Foundation for Research in Economics, Yale University.

Frisch, U. and G. Parisi (1985). *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*, Chapter Fully developed turbulence and intermittency, pp. 84–88. Amsterdam: North-Holland.

Galluccio, S., G. Galdanelli, M. Marsili, and Y.-C. Zhang (1997). Scaling in currency exchange. *Physica A: Statistical Mechanics and its Applications 245*, 423–436.

Garman, M. (1976). Market microstructure. *Journal of Financial Economics 3*, 257–275.

Gavrishchaka, V. and S. Ganguli (2003). Volatility forecasting from multiscale and high-dimensional market data. *Neurocomputing 55*, 285–305.

Geman, H. and C. Kharoubi (2008). WTI crude oil futures in portfolio diversification: The time-to-maturity effect. *Journal of Banking and Finance 32*, 2553–2559.

Gençay, R. (2001). Scaling properties of foreign exchange volatility. *Physica A 289*, 89–106.

Ghasghaie, S., W. Breymann, J. Peinke, P. Talkner, and Y. Dodge (1996). Turbulent cascades in foreign exchange markets. *Nature 381*, 767–770.

Ghysels, E., C. Gouriéroux, and J. Jasiak (2004). Stochastic volatility duration models. *Journal of Econometrics 119*, 413–433.

Ghysels, E., A. C. Harvey, and E. Renault (1996). *Handbook of Statistics*, Volume 14, Chapter Stochastic volatility, pp. 119–191. North-Holland, Amsterdam.

Giacomini, R. and I. Komunjer (2005). Evaluation and combination of conditional quantile forecasts. *Journal of Business and Economic Statistics 23*, 416–431.

Giot, P. (2005). Market risk models for intraday data. *European Journal of Finance 11*, 309–324.

Glaser, R. E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association 75*, 667–672.

Glosten, L., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and volatility of the nominal excess return on stocks. *Journal of Finance 46*, 1779–1801.

Glosten, L. and P. Milgrom (1985). Bid, ask, and the transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 13*, 71–100.

Gopikrishnan, P., M. Meyer, L. Amaral, and H. Stanley (1998). Inverse cubic law for the probability distribution of stock price variations. *European Journal of Physics B Rapid Communications 3*, 139–140.

Gouriéroux, C., J. Jasiak, and G. L. Fol (1999). Intra-day market activity. *Journal of Financial Markets 2*, 193–226.

Gouriéroux, C. and A. Monfort (1996). *Simulation Based Econometric Methods*. Oxford University Press.

Grammig, J. and K. O. Maurer (2000). Non-monotonic hazard functions and the autoregressive conditional duration model. *Journal of Econometrics 3*, 16–38.

Grammig, J. and M. Wellner (2002). Modeling the interdependence of volatility and inter-transaction duration processes. *Journal of Econometrics 106*, 369–400.

Granger, C. W. and T. Teräsvirta (1999). A simple nonlinear time series model with missleading linear properties. *Economics Letters 62*, 161–165.

Grossman, S. and J. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review 70*, 393–408.

Guillaume, D. M., M. M. Dacorogna, R. R. Davé, U. A. Müller, R. B. Olsen, and O. V. Pictet (1997). From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics 1*, 95–129.

Haas, M. (2005). Improved duration-based backtesting of value-at-risk. *Journal of Risk 8*, 17–36.

Hafner, C. (2005). Durations, volume and the prediction of financial returns in transaction time. *Quantitative Finance 5*, 145–152.

Hamilton, J. D. (1983). Oil and the macroeconomy since World War II. *Journal of Political Economy 91*, 228–248.

Hamilton, J. D. (2003). What is an oil shock. *Journal of Econometrics 113*, 363–398.

Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics 23*, 365–380.

Harris, D. and L. Mátyás (1999). *Generalized Method of Moments Estimations*, Chapter Introduction to the generalized method of moments estimation. Cambridge University Press.

Harvey, A., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *Review of Economic Studies 61*, 247–264.

Hasbrouck, J. (1988). Trades, quotes, inventories, and information. *Journal of Financial Economics 22*, 229–252.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance 46*, 179–208.

Haughton, D. (1988). On the choice of a model to fit data from an exponential family. *Annals of Statistics 16*, 342–355.

Haughton, D. (1989). Size of the error in the choice of a model to fit data from an exponential family. *Indian Journal of Statistics 51*, 45–58.

Hautsch, N. (2001). Modelling intraday trading activity using Box-Cox ACD models. Unpublished Paper, University of Konstanz.

Hautsch, N. (2003). Assessing the risk of liquidity suppliers on the basis of excess demand intensities. *Journal of Financial Econometrics 1*, 189–215.

Hautsch, N. (2006). Testing the conditional mean function of autoregressive conditional duration models. Unpublished Paper, University of Copenhagen.

Hautsch, N. (2008). Testing multiplicative error models using conditional moments tests. Unpublished Paper, University Humboldt.

Hautsch, N. (2012). *Econometrics of Financial High-Frequency*. Berlin: Springer.

Hill, B. (1975a). A simple general approach to inference about the tail of a distribution. *Annals of Mathematical Statistics 3*, 1163–1174.

Hill, B. M. (1975b). A simple general approach to inference the tail of a distribution. *Annals of Statistics 3*, 1163–1174.

Ho, T. and H. Stoll (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics 9*, 47–73.

Ho, T. and H. Stoll (1983). The dynamics of dealer markets under competition. *Journal of Finance 38*, 1053–1075.

Holley, R. and E. C. Waymire (1992). Multifractal dimensions and scaling exponents for strongly bounded random cascades. *Annals of Applied Probability 2*, 819–845.

Hong, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica 64*, 837–864.

Hong, Y. (1997). One-sided testing for conditional heteroskedasticity in time series models. *Journal of Time Series Analysis 18*, 253–277.

Huberman, G. (2001). Familiarity breeds investment. *Review of Financial Studies 14*, 659–680.

Hujer, R. and S. Vuletić (2007). Econometric analysis of financial trade processes by discrete mixture duration models. *Journal of Economic Dynamics and Control 31*, 635–667.

Hujer, R., S. Vuletić, and S. Kokot (2002). The Markov switching ACD model. Unpublished Paper Series: Finance and Accounting 90, University of Frankfurt/Main.

Hurst, H. (1951). Long-term storage capacity of reservoirs. *Trans Am Soc Civil Eng 116*, 770–808.

Hurst, H. (1955). Methods of using long-term storage in reservoirs. *Proc Inst Civil Eng Part I*, 519–577.

Idier, J. (2011). Long-term vs. short-term comovements in stock markets: The use of Markov-switching multifractal models. *European Journal of Finance 17*, 27–48.

Ivanov, P. C., A. Yuen, B. Podobnik, and Y. K. Lee (2004). Common scaling patterns in intertrade times of U.S. stocks. *Physical Review E 69*, 056107.

Jach, A. and P. Kokoszka (2010). Empirical wavelet analysis of tail and memory properties of LARCH and FIGARCH processes. *Computational Statistics 25*, 163–182.

Jansen, D. and C. de Vries (1991). On the frequency of large stock market returns: Putting booms and busts into perspective. *Review of Economics and Statistics 23*, 18–24.

Jasiak, J. (1998). Persistence in intertrade durations. *Journal of Finance 19*, 166–195.

Jiang, Z. Q., W. Chen, and W. X. Zhou (2008). Scaling in the distribution of intertrade durations of chinese stocks. *Physica A 387*, 5818–5825.

Jones, C. M. and G. Kaul (1996). Oil and the stock markets. *Journal of Finance 51*, 463–491.

Kahane, J. P. and J. Peyrière (1976). Sur certaines martingales de Benoit Mandelbrot. *Advances in Mathematics 22*, 131–145.

Kang, S. H., S. M. Kang, and S. M. Yoon (2009). Forecasting volatility of crude oil markets. *Energy Economics 31*, 119–125.

Karanasos, M. (2004). The statistical properties of long-memory ACD models. *WSEAS Tran Bus Econ 2*, 169–175.

Karanasos, M. (2008). The statistical properties of exponential ACD models. *Quantitative and Qualitative Analysis in Social Sciences 2*, 29–49.

Kearns, P. and A. Pagan (1997). Estimating the tail density index for financial time series. *Review of Economics and Statistics 79*, 171–175.

Kendall, M. (1953). The analysis of economic time series. *Journal of the Royal Statistical Society, Series A 96*, 11–25.

Kiyono, K. (2009). Log-amplitude statistics of intermittent and non-Gaussian time series. *Physical Review E 79*, 031129.

Kiyono, K., Z. R. Struzik, N. Aoyagi, S. Sakata, J. Hayano, and Y. Yamamoto (2004). Critical scale-invariance in healthy human heart rate. *Physical Review Letters 93*, 178103.

Kiyono, K., Z. R. Struzik, N. Aoyagi, F. Togo, and Y. Yamamoto (2005). Phase transition in healthy human heart rate. *Physical Review Letters 95*, 058101.

Kiyono, K., Z. R. Struzik, and Y. Yamamoto (2007). Estimator of a non-Gaussian parameter in multiplicative Lognormal models. *Physical Review E 76*, 041113.

Knight, J. and C. Q. Ning (2008). Estimation of the stochastic conditional duration model via alternative methods. *Econometrics Journal 11*, 593–616.

Koedijk, K. G., M. Schafgans, and C. de Vries (1990). The tail index of exchange rate returns. *Journal of International Economics 29*, 93–108.

Koedijk, K. G., P. A. Stork, and C. de Vries (1992). Differences between foreign exchange rate regimes: The view from the tails. *Journal of International Money and Finance 11*, 462–473.

Kolmogorov, A. N. (1941). The local structure of turbulence in incompressible viscous fluids at very large Reynolds number. *Dokl. Akad. Nauk. SSSR 30*, 299–303. Reprinted in *Proc. R. Soc.* London A 434, 9 − 13 (1991).

Kolmogorov, A. N. (1962). A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *Journal of Fluid Mechanics 13*, 82–85.

Kon, S. J. (1984). Models of stock returns- a comparison. *Journal of Finance 39*, 147–165.

Koulikov, D. (2003). Modeling sequences of long memory non-negative covariance stationary random variables. Unpublished Paper, University of Aarhus.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *Journal of Derivatives 3*, 73–84.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica 53*, 1315–1365.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge University Press.

LeBaron, B. (1992). Some relations between volatility and serial correlations in stock market returns. *Journal of Business 65*, 199–219.

LeBaron, B. (2001). Stochastic volatility as a simple generator of apparent financial power laws and long memory. *Quantitative Finance 1*, 621–631.

Lee, J. and R. P. Degennaro (2000). Smooth transition ARCH models: Estimation and testing. *Review of Quantitative Finance and Accounting 15*, 5–20.

Lee, K. and S. Ni (2002). On the dynamic effects of oil price shocks: a study using industry level data. *Journal of Monetary Economics 49*, 823–852.

Lee, S. W. and B. E. Hansen (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory 10*, 29–52.

Leövey, A. (2013). *Multifractal Models: Estimation, Forecasting and Option Pricing*. Ph. D. thesis, University of Kiel.

Leövey, A. and T. Lux (2012). Parameter estimation and forecasting for multiplicative Lognormal cascades. *Physical Review E 85*, 046114.

Li, W. K. and T. K. Mak (1994). On the squared residual autocorrelations in non-linear time series with conditional heteroscedasticity. *Journal of Time Series Analysis 15*, 627–636.

Li, W. K. and P. L. Yu (2003). On the residual autocorrelation of the autoregressive conditional duration model. *Economics Letters 79*, 169–175.

Liesenfeld, R. and J. F. Richard (2003). Univariate and multivariate stochastic volatility models: Estimation and diagnostics. *Journal of Empirical Finance 10*, 505–532.

Liu, R. (2008). *Multivariate Multifractal Models: Estimation of Parameters and Application to Risk Management*. Ph. D. thesis, University of Kiel.

Liu, R., T. di Matteo, and T. Lux (2007). True and apparent scaling: The proximity of the Markov-switching multifractal model to long-range dependence. *Physica A 383*, 35–42.

Liu, R., T. di Matteo, and T. Lux (2008). Multifractality and long-range dependence of asset returns: the scaling behaviour of the Markov-switching multifractal model with Lognormal volatility components. *Advances in Complex Systems 11*, 669–684.

Liu, R. and T. Lux (2014). Non-homogeneous volatility correlations in the bivariate multifractal model. *European Journal of Finance*. in Press.

Liu, Y. P., P. Gopikrishnan, P. Cizeau, M. Meyer, C.-H. Peng, and H. E. Stanley (1999). Statistical properties of the volatility of price fluctuations. *Physical Review E 60*, 1390–1400.

Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica 59*, 1279–1313.

Lobato, I. and N. Savin (1998). Real and spurious long-memory properties of stock market data. *Journal of Business and Economics Statistics 16*, 261–283.

Lobato, I. and C. Velasco (2000). Long memory in stock market trading volume. *Journal of Business and Economics Statistics 18*, 410–427.

Løvsletten, O. and M. Rypdal (2012). Approximated maximum likelihood estimation in multifractal random walks. *Physical Review E 85*, 046705.

Lundbergh, S. and T. Teräsvirta (2002). Evaluating GARCH models. *Journal of Econometrics 110*, 417–435.

Lunde, A. (1999). A generalized gamma autoregressive conditional duration model. Unpublished Paper, Aalborg University.

Lux, T. (1996). The stable Paretian hypothesis and the frequency of large returns: An examination of major german stocks. *Applied Economics Letters 6*, 463–475.

Lux, T. (2001a). The limiting extremal behaviour of speculative returns: An analysis of intra-daily data from the Frankfurt stock exchange. *Applied Financial Economics 11*, 299–315.

Lux, T. (2001b). Power-laws and long memory. *Quantitative Finance 1*, 560–562.

Lux, T. (2001c). Turbulence in financial markets: The surprising explanatory power of simple models. *Quantitative Finance 1*, 641–652.

Lux, T. (2004). Detecting multi-fractal properties in asset returns. *International Journal of Modern Physics 15*, 481–491.

Lux, T. (2008). The Markov-switching multifractal model of asset returns: GMM estimation and linear forecasting of volatility. *Journal of Business and Economic Statistics 26*, 194–210.

Lux, T. and T. Kaizoji (2007). Forecasting volatility and volume in the Tokyo stock market: Long memory, fractality and regime switching. *Journal of Economic Dynamics and Control 31*, 1808–1843.

Lux, T. and L. Morales-Arias (2010). Forecasting volatility under fractality, regime-switching, long memory and Student-*t* innovations. *Computational Statistics and Data Analysis 54*, 2676–2692.

Lux, T. and L. Morales-Arias (2013). Relative forecasting performance of volatility models: Monte Carlo evidence. *Quantitative Finance*. in Press.

Lux, T., L. Morales-Arias, and C. Sattarhoff (2014). A Markov-switching multifractal approach to forecasting realized volatility. *Journal of Forecasting 33*, 532–541.

Madhavan, A. (1992). Trading mechanisms in securities markets. *Journal of Finance 47*, 607–642.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business 26*, 394–419.

Mandelbrot, B. B. (1969). Long-run linearity, locally Gaussian processes, H-spectra and infinite variance. *International Economic Review 10*, 82–111.

Mandelbrot, B. B. (1974). Intermittent turbulence in self similar cascades; divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics 62*, 331–358.

Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. New York: WH Freeman and Co.

Mandelbrot, B. B. (1989). Multifractal measures, especially for the geophysicist. *Pure and Applied Geophysics 131*, 5–42.

Mandelbrot, B. B. (1990). Limit Lognormal multifractal measures. In E. Gotsman and et al. (Eds.), *Frontiers of Physics*. New York: Pergamon. Landau Memorial Conference.

Mandelbrot, B. B. (1997a). *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*. New York: Springer Verlag.

Mandelbrot, B. B. (1997b). Three fractal models in finance: Discontinuity, concentration, risk. *Economic Notes 26*, 171–212.

Mandelbrot, B. B. (1999). A multifractal walk down street. *Scientific Amereican Magazine*.

Mandelbrot, B. B. (2001a). Scaling in financial prices: I. Tails and dependence. *Quantitative Finance 1*, 113–123.

Mandelbrot, B. B. (2001b). Scaling in financial prices: II. Multifractals and the star equation. *Quantitative Finance 1*, 124–130.

Mandelbrot, B. B. (2001c). Scaling in financial prices: III. Cartoon Brownian motions in multifractal time. *Quantitative Finance 1*, 427–440.

Mandelbrot, B. B., A. Fisher, and L. Calvet (1997). A multifractal model of asset returns. Cowles Foundation Discussion Papers 1164, Cowles Foundation for Research in Economics, Yale University.

Mandelbrot, B. B. and H. M. Taylor (1967). On the distribution of stock price differences. *Operations Research 15*, 1057–1062.

Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial Markets 8*, 377–399.

Mantegna, R. N. (1991). Lévy walks and enhanced diffusion in Milan Stock-Exchange. *Physica A 179*, 232–242.

Mantegna, R. N. and H. E. Stanley (1995). Scaling behaviour in the dynamics of an economic index. *Nature 376*, 46–49.

Mantegna, R. N. and H. E. Stanley (1996). Turbulence and financial markets. *Nature 383*, 587–588.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons, New York.

Matia, K., L. A. Amaral, S. P. Goodwin, and H. E. Stanley (2002). Different scaling behaviors of commodity spot and future prices. *Physical Review E 66*, 045103.

McCulloch, J. H. (1986). *Handbook of Statistics*, Volume 14, Chapter Financial applications of stable distributions, pp. 393–425.

McDonald, J. B. and D. O. Richards (1987). Hazard rates and generalized Beta distributions. *IEEE Transactions on Reliability 36*, 463–466.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica 57*, 995–1026.

McInish, T. and R. Wood (1992). An analysis of intraday patterns in bid/ask spreads for NYSE stocks. *Journal of Finance 47*, 753–764.

Meddahi, N., E. Renault, and B. Werker (2006). GARCH and irregularly spaced data. *Economic Letters 90*, 200–204.

Meitz, M. and T. Teräsvirta (2006). Evaluating models of autoregressive conditional duration. *Journal of Business and Economic Statistics 24*, 104–124.

Miccichè, S., G. Bonanno, F. Lillo, and R. N. Mantegna (2002). Volatility in financial markets: Stochastic models and empirical results. *Physica A: Statistical Mechanics and its Applications 314*, 756–761.

Mike, S. and D. Farmer (2008). An empirical behavioral model of liquidity and volatility. *Journal of Economic Dynamics and Control 32*, 200–234.

Mills, T. (1997). Stylized facts of the temporal and distributional properties of daily FTSE returns. *Applied Financial Economics 7*, 599–604.

Mohammadi, H. and L. Su (2010). International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models. *Energy Economics 32*, 1001–1008.

Mostafa, M. D. and M. W. Mahmoud (1964). On the problem of estimation for the bivariate Lognormal distribution. *Biometrika 51*, 522–527.

Müller, U. A., M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Pictet, and J. E. von Weizsäcker (1997). Volatilities of different time resolutions: Analyzing the dynamics of market components. *Journal of Empirical Finance 4*, 213–239.

Müller, U. A., M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, and C. Morgenegg (1990). Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. *Journal of Banking and Finance 14*, 1189–1208.

Muzy, J.-F. and E. Bacry (2002). Multifractal stationary random measures and multifractal random walks with log infinitely divisible scaling laws. *Physical Review E 66*, 056121.

Muzy, J.-F., E. Bacry, and A. Kozhemyak (2006). Extreme values and fat tails of multifractal fluctuations. *Physical Review E 73*, 066114.

Muzy, J.-F., J. Delour, and E. Bacry (2000). Modeling fluctuations of financial time series: From cascade process to stochastic volatility model. *European Journal of Physics B 17*, 537–548.

Nelson, D. B. (1990). ARCH models as diffusion approximations. *Journal of Econometrics 45*, 7–38.

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

Nelson, D. B. and C. Q. Cao (1992). Inequality constraints in the univariate GARCH model. *Journal of Business and Economic Statistics 10*, 229–235.

Newey, W. K. (1985). Maximum likelihood specification testing and conditional moment tests. *Econometrica 53*, 1047–1070.

Nishii, R. (1984). Asymptotic properties of criteria for selection of variables in multiple regression. *Annals of Statistics 12*, 758–765.

Obukhov, A. M. (1962). Some specific features of atmospheric turbulence. *Journal of Fluid Mechanics 13*, 77–81.

Ossiander, M. and E. C. Waymire (2000). Statistical estimation for multiplicative cascades. *Annals of Statistics 28*, 1533–1560.

Pacurar, M. (2008). Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature. *Journal of Economic Survey 4*, 711–751.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business 53*, 61–65.

Peng, L. and W. Xiong (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics 80*, 563–602.

Plerou, V., P. Gopikrishnan, and H. E. Stanley (2005). Quantifying fluctuations in market liquidity: Analysis of the bid-ask spread. *Physical Review E 71*, 046131.

Pochart, B. and J. P. Bouchaud (2002). The skewed multifractal random walk with applications to option smiles. *Quantitative Finance 24*, 303–314.

Politis, M. and E. Scalas (2008). Fitting the empirical distribution of intertrade durations. *Physica A 387*, 2025–2034.

Poon, S. and C. Granger (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature 26*, 478–539.

Prigent, J. L., O. Renault, and O. Scaillet (2001). An autoregressive conditional Binomial option pricing model. In H. Geman, D. Madan, S. Pliska, and T. Vorst (Eds.), *Selected papers from the first World Congress of the Bachelier Finance Society*. Heidelberg: Springer.

Qiu, T., G. Chen, L.-X. Zhong, and X.-R. Wu (2012). Dynamics of bid-ask spread return and volatility of the Chinese stock market. *Physica A 391*, 2656–2666.

Qu, G., W. Chen, and W. Zhou (2007). Quantifying bid-ask spreads in the Chinese stock market using limit-order book data. *European Physical Journal B 57*, 81–87.

Rabemananjara, R. and J. Zakoian (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics 8*, 31–49.

Rachev, S. and S. Mittnik (2000). *Stable Paretian Models in Finance*. Willey, New York.

Racheva, B. and G. Samorodnitsky (2003). *Handbook of Heavy Tailed Distributions in Finance*, Chapter Long range dependence in heavy tailed stochastic processes. Elsevier: Amsterdam.

Rahman, S. and A. Serletis (2012). Oil price uncertainty and the Canadian economy: Evidence from a VARMA, GARCH-in-Mean, asymmetric BEKK model. *Energy Economics 34*, 603–610.

Razali, N. M. and Y. B. Wah (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics 2*, 21–33.

Reiss, R. and M. Thomas (1997). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and other Fields*. Birkhäuser, Basel.

Riedi, R. H. (2002). *Long Range Dependence: Theory and Applications*, Chapter Multifractal processes, pp. 625–715. Birkhäuser.

Roberts, H. (1959). Stock market patterns and financial analysis: Methodological suggestions. *Journal of Finance 44*, 1–10.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance 39*, 1127–1139.

Rosenblatt, R. F. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics 23*, 470–472.

Rotemberg, J. J. and M. Woodford (1996). Imperfect competition and the effects of energy price increases. *Journal of Money, Credit, and Banking 28*, 549–577.

Russell, J. R. and R. F. Engle (2005). A discrete-state continuous-time model of financial transactions prices and times: The autoregressive conditional multinomial-autoregressive conditional duration model. *Journal of Business and Economic Statistics 23*, 166–180.

Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics 21*, 449–469.

Sattarhoff, C. (2010). GMM estimation of multifractal random walks using an efficient algorithm for HAC covariance matrix estimation. Working paper, University of Hamburg.

Schmitt, F., D. Schertzer, and S. Lovejoy (1999). Multifractal analysis of foreign exchange data. *Applied Stochastic Models and Data Analysis 15*, 29–53.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics 6*, 461–464.

Segnon, M. and T. Lux (2012). Assessing forecast performance of financial duration models via density forecasts and likelihood ratio test. Working Paper, University of Kiel.

Shephard, N. (1996). *Time Series Models in Econometrics, Finance and other Fields*, Chapter Statistical aspects of ARCH and stochastic volatility models, pp. 1–67. Chapman & Hall, London.

Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika 84*, 653–667.

Sorriso-Valvo, L., V. Carbone, P. Veltri, G. Consolini, and R. Bruno (1999). Intermittency in the solar wind turbulence through probability distribution functions of fluctuations. *Geophysical Research Letters 26*, 1801–1804.

Stacy, E. (1962). A generalization of the gamma distribution. *Annals of Mathematical Statistics 33*, 1187–1192.

Stein, E. M. and J. C. Stein (1991). Stock price distributions with stochastic volatility: An analytic approach. *Review Financial Studies 4*, 727–752.

Stoll, H. (1978). The supply of dealer services in markets. *Journal of Finance 33*, 1133–1151.

Stoll, H. and R. Whaley (1990). Stock market structure and volatility. *Review of Financial Studies 3*, 37–71.

Sun, W., S. Rachev, F. J. Fabozzi, and P. S. Kalev (2008). Fractals in trade duration: Capturing long-range dependence and heavy tailedness in modeling trade duration. *Annals of Finance 4*, 217–241.

Tauchen, G. E. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica 51*, 485–506.

Taylor, S. J. (1986). *Modelling Financial Time Series*. Wiley.

Teichmoeller, J. (1971). Distribution of stock price changes. *Journal of the American Statistical Association 66*, 282–284.

Tél, T. (1988). Fractals, multifractals, and thermodynamics. *Zeitschrift für Naturforschung 43*, 1154–1174.

Tse, Y. K. (1998). The conditional heteroscedasticity of the Yen-Dollar exchange rate. *Journal of Applied Econometrics 13*, 49–55.

Ureche-Rangau, L. and Q. de Morthays (2009). More on the volatility trading volume relationship in emerging markets: The Chinese stock market. *Journal of Applied Statistics 36*, 779–799.

Vassilicos, J., A. Demos, and F. Tata (1994). *Applications of Fractals and Chaos*, Chapter No evidence of chaos but some evidence of multifractals in the foreign exchange and the stock market. Berlin: Springer.

Veredas, D., J. Rodriguez-Poo, and A. Espasa (2001). On the intraday seasonality and dynamics of a financial point process: A semic-parametric approach. Unpublished Paper, Université Catholique de Louvain.

Wang, S. P., A. M. Hu, and Z. X. Wu (2012). The impact of oil price volatility on China's economy: An empirical investigation based on VAR model. *Advanced Materials Research 524-527*, 3211.

Wang, T., J. T. Wu, and J. Yang (2008). Realized volatility and correlation in energy futures markets. *Journal of Futures Markets 28*, 993–1011.

Wei, Y., Y. Wang, and D. Huang (2010). Forecasting crude oil market volatility: further evidence using GARCH-class models. *Energy Economics 32*, 1477–1484.

Weisberg, S. (2001). The delta method in arc. School of Statistics, university of Minnesota, St. Paul MN 55108-6042.

Weron, R. (2002). Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A: Statistical Mechanics and its Applications 312*, 285–299.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica 64*, 1067–1084.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica 50*, 1–25.

White, H. (1994). *Estimation, Inference, and Specification Analysis*. Cambridge University Press.

White, H. (2000). A reality check for data snooping. *Econometrica 68*, 1097–1126.

Whittle, P. (1962). Gaussian estimation in stationary time series. *Bulletin de l'Institut International de Statistique 39*, 105–129.

Working, H. (1934). A random difference series for use in the analysis of time series. *Journal of the American Statistical Association 29*, 11–24.

Yerel, S. and A. Konuk (2009). Bivariate Lognormal distribution model of cutoff grade impurities: A case study of magnesite ore deposit. *Scientific Research and Essay 4*, 1500–1504.

Zhang, M., J. Russell, and R. Tsay (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics 104*, 179–207.

Affirmation

I hereby affirm that I have completed my doctoral thesis entitled, "Multifractal Models, Intertrade Durations And Return Volatility" entirely on my own and unassisted, and that I have specially marked all of the quotes I have used from other authors as well as those passages in my work that are extremely close to the thoughts presented by other authors, and listed the sources in accordance with the regulations I have been given.

June 17, 2015 Mawuli Kouami Segnon

Mawuli Kouami Segnon

Schmiedekoppel 5, 24106 Kiel (Germany)
Tel: ++49 431 657 271 40/++49 152 343 258 55
Email: mawuli.segnon@yahoo.de

Personal Information

Date of birth	December 31, 1985
Place of birth	Kponou, Togo
Personal status	Single

Education

10/2010-06/2015	Ph.D. Candidate in Quantitative Economics, University of Kiel Ph.D. -Programme, Quantitative Economics Dissertation: "Multifractal Models, Intertrade Durations and Return Volatility"
04/2010-09/2010	Informatics studies at the Christian-Albrechts-University of Kiel
10/2005-02/2010	Studies of Business Administration, Statistics and Econometrics at the Christian-Albrechts-University of Kiel, Degree: Diplom Thesis: "Specification of Random Coefficients in Panel Probit Models"
10/2004-07/2005	Studies of the German language at the Christian-Albrechts-University of Kiel
05/2004-07/2004	Volkshochschule Flensburg, language courses
09/2000-07/2003	High School Tokoin, Togo

Research Interests

International financial markets, energy markets, monetary economics, monetary policy, modeling and forecasting financial time series, market microstructure, derivative Products, risk management and multifractal models

Research Projects

1. Modeling and forecasting CO_2 emission allowance spot prices in the presence of structural breaks.

2. Multifractal modeling of inflation uncertainty and application of Granger-causality tests to provide some statistical evidence on the nature of the relationship between average inflation and nominal uncertainty in Euro area.

3. Use ARMA, VAR, BVAR and its extensions and DSGE models to provide point- and density forecasts for the US output, inflation and interest rate using a measure of uncertainty.

Publications

1. Co-editor of **lecture notes for Econometrics** in the Bachelor program at the University of Kiel (2009)

2. **Multifractal Models: Their Origin, Properties and Applications in Finance** (2013), Working Paper No 1860, Kiel Institute for the World Economy, Forthcoming in *Handbook on Computational Economics and Finance*, Oxford University Press (with Thomas Lux).

3. **Forecasting the Price of Gold**, *Applied Economics* (2015) (in press), (with Hossein Hassani, Emmanuel S. Silva und Rangan Gupta).

Working Papers

- **Assessing Forecast Performance of Financial Duration Models via Density Forecasts** (2012), *Working Paper*, University of Kiel, Germany (mit Thomas Lux).

- **Forecasting Intra-day Value-at-Risk Using Markov-Switching Multifractal Duration Model** (2013), *Working Paper*, University of Kiel, Germany (with Thomas Lux).

- **Currency Options Pricing with the Poisson Multifractal Diffusion Model** (2014), *Working Paper*, University of Kiel, Germany (with Thomas Lux).

- **Forecasting Crude Oil Price Volatility: Evidence from Historical and Recent Data** (2014), *Working Paper*, University of Kiel, Germany (with Thomas Lux und Rangan Gupta), submitted for publication to "Energy Economics" (in revision).

- **Forecasting Home Sale in the Four Census Regions and the Aggregate US Economy Using Singular Spectrum Analysis** (2014), *Working Paper*, University of Pretoria, South Africa (with Hossein Hassani, Zara Ghodsi und Rangan Gupta), submitted for publication to "Computational Economics" (in revision).

Conference and Presentations

Brown Bag Workshop, University of Kiel, July 2012
Research seminar in "Financial Economics", University Kiel, July 2013
BIRS Workshop "Multifractal Analysis: From Theory to Applications and Back", the Banff Centre in Banff, Alberta, Canada, February 2014
Research seminar in "Financial Economics", University Kiel, December 2014

Working Experience

12/2007-03/2010 Student Assistant at the Chair of Statistics and Econometrics at the University of Kiel (Do administrative tasks. Data gathering and preparation for research, literature research.)

07/2011-02/2013 Student Assistant at the Kiel Institute for the World Economy, Research field: "Financial Markets and Macroeconomic Activity" (Data gathering and preparation for research, literature research.)

10/2012-02/2014 Teaching Assistant for Introduction to Macroeconomics at the University of Kiel (Help students to better understand the lecture, to solve their exercisesand prepare them for the exam.)

Scholarships

10/2010-12/2012 *Post-graduate scholarship*, Christian-Albrechts-University Department of Economics Doctoral Programme "Quantitative Economics"

01/2013-06/2015 *Post-graduate scholarship*, Evangilisches Studienwerk e.v. Villigst

Languages and Computer Skills

Languages French (native), German, English (Proficient)

Computer MS-Office (Excel, PowerPoint), Matlab, LateX, Beamer (LateX), Eviews, Gauss, R, C++ (Basis knowledge)

References

Prof. Dr. Thomas Lux
Institute for Economics, University of Kiel
Chair of Monetary Economics and International Financial Markets
Olshausenstrasse 40, 24118 Kiel, Germany,
Tel: ++49 431 880 3661
E-mail: thomas.lux@bwl.uni-kiel.de

Prof. Dr. Markus Haas
Institute for Quantitative Business and Economics Research, University of Kiel
Chair of Empirical Economic Research
Olshausenstrasse 40, 24118 Kiel, Germany,
Tel: ++49 431 880 5505
E-mail: markus.haas@qber.uni-kiel.de

Kiel, June 15, 2015